



HAL
open science

Mécanismes de raisonnement possibiliste pour l'aide à la décision et l'interprétation de scènes

Imene Khanfir Kallel

► To cite this version:

Imene Khanfir Kallel. Mécanismes de raisonnement possibiliste pour l'aide à la décision et l'interprétation de scènes. Traitement du signal et de l'image [eess.SP]. Ecole Nationale des Ingénieurs de Sfax, Université de Sfax, 2019. tel-02868499

HAL Id: tel-02868499

<https://hal.science/tel-02868499>

Submitted on 15 Jun 2020

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

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



RAPPORT DE SYNTHÈSE

Présenté à

L'École Nationale d'Ingénieurs de Sfax

En vue de l'obtention de

HABILITATION UNIVERSITAIRE

Génie Electrique

Par

Dr. Ing. Imène KHANFIR KALLEL

Doctorat en génie électrique

**Mécanismes de raisonnement possibiliste pour
l'aide à la décision et l'interprétation de
scènes**

Soutenu le 05 Avril 2019, devant le jury composé de :

Pr. Abdennaceur Kachouri	(Université de Sfax/ENIS, Tunisie)	<i>Président</i>
Pr. Imed Riadh Farah	(Université de la Manouba/ISAMM, Tunisie)	<i>Rapporteur</i>
Pr. Jean Paul Haton	(Université de Lorraine/LORIA, France)	<i>Rapporteur</i>
Pr. Nabil Derbel	(Université de Sfax/ENIS, Tunisie)	<i>Examineur</i>
Pr. Basel Solaiman	(IMT Atlantique/ITI, France)	<i>Examineur</i>
Pr. Eloi Bossé	(McMaster University/CEED, Canada)	<i>Examineur</i>

Remerciements

Je remercie Monsieur **Abdennaceur Kachouri**, Professeur de l'Ecole Nationale d'Ingénieurs de Sfax (ENIS), laboratoire LETI, pour l'honneur qu'il me fait en présidant mon jury d'habilitation. J'éprouve pour lui un grand respect depuis que j'étais étudiante dans ses cours et je lui adresse ma reconnaissance pour son soutien.

J'exprime ma sincère gratitude à Monsieur **Imed Riadh Farah**, Professeur en Informatique à l'Institut Supérieur des Arts Multimédias de La Manouba (ISAMM) et Responsable de l'équipe de recherche Géomatique Décisionnelle (GeoDec) au sein du laboratoire RIADI, pour m'avoir fait l'honneur de rapporter ce travail et pour ses remarques précieuses qui ont contribué à améliorer ma synthèse.

Je suis très honorée que Monsieur **Jean Paul Haton**, Professeur de l'Université de Lorraine à Nancy, LORIA/INRIA, membre senior de l'Institut Universitaire de France (IUF) et créateur de la chaire « Modélisation informatique des processus perceptifs et cognitifs », ait accepté d'évaluer mon travail et de le rapporter. Je tiens aussi à le remercier pour m'avoir soutenue et pour avoir cru en ma compétence scientifique.

J'exprime ma profonde reconnaissance à Monsieur **Nabil Derbel**, Professeur de l'Ecole Nationale d'Ingénieurs de Sfax (ENIS), laboratoire CEM-Lab, pour avoir accepté d'examiner ce travail. C'est avec un grand plaisir et une grande fierté que je travaille avec lui depuis ma Thèse de Doctorat dont il était Directeur. Je tiens à lui exprimer des remerciements très particuliers pour son bienveillant soutien et ses conseils pertinents. Je lui suis très sincèrement reconnaissante pour les encouragements et la confiance qu'il m'a accordée notamment dans le co-encadrement des thèses.

Je voudrais remercier tout particulièrement et profondément Monsieur **Basel Solaiman**, Professeur de l'École nationale supérieure Mines-Télécom Atlantique Bretagne Pays de la Loire et Responsable du Département Image et Traitement de l'Information, sans qui ce travail n'aurait pas été possible. Il a été mon guide scientifique durant toutes mes années de recherche, depuis mon projet de PFE d'ingénieur en 2002. Par sa générosité et son immense compétence et savoir il répondait à toutes mes questions et sollicitations, me transmettant ainsi la plupart de mes connaissances actuelles. Je le remercie aussi pour m'avoir accueillie dans son équipe pour plusieurs séjours scientifiques de courtes durées. Il a en outre accepté d'examiner la présente synthèse de mes travaux, qui a bénéficié de ses nombreuses remarques. Il est l'exemple scientifique et humain que j'essaye de suivre et pour tout cela je lui dois un immense merci.

J'exprime toute ma gratitude à Monsieur **Éloi Bossé**, Professeur adjoint de l'Université McMaster, pour avoir accepté d'examiner ce travail. Durant cette dernière année, il fut la personne qui a su m'apporter des solutions et m'indiquer la voie à suivre lorsque j'ai été confronté à des difficultés scientifiques. J'ai beaucoup d'admiration pour sa modestie, sa facilité à transmettre les connaissances, sa très grande compétence et son exceptionnelle générosité. En l'espace d'une petite année, je lui dois déjà beaucoup. Notre collaboration a déjà commencé à donner ses fruits et je souhaite qu'elle dure très longtemps.

Ma respectueuse reconnaissance s'adresse également à Madame **Najoua Essoukri Ben Amara**, Professeur de l'Ecole Nationale d'Ingénieurs de Sousse et Responsable du Laboratoire de recherche LATIS. Je suis sensible à l'honneur qu'elle m'a fait de participer à mon jury.

J'adresse mon chaleureux remerciement à Monsieur **Mohamed Koubaa**, Professeur en physique et directeur de l'Institut Supérieur de Biotechnologies de Sfax (ISBS), pour son amitié précieuse son soutien et ses encouragements.

Mes remerciements s'adressent à Madame **Dorra Sellami**, Professeur à l'Ecole Nationale d'Ingénieurs de Sfax (ENIS) et Madame **Houda Ben Jemaa**, Professeur à la Faculté des Sciences de Sfax (FSS), pour la confiance qu'elles m'ont témoignée en co-encadrement de thèses.

Ce travail d'habilitation est le fruit d'un travail collectif, dont une grande partie a été menée avec mes collègues de l'ENIS et de l'ENST Bretagne, mes étudiants en mastère de l'ISIMS, l'ENIS et l'ISBS et mes doctorants de l'ENIS. Que toute personne, ayant contribué à la réalisation de ces travaux de recherche puisse trouver l'expression de ma reconnaissance. Ma gratitude et mon amitié sont adressées spécialement à Sonda Ammar, Sameh Fakhfakh, Ameni Yengui, Nadia Smaoui, Houda Khmila, Samia Barhoumi, Mouna Medheyoub, Marwa Chakroun, Amal Charfi, Hana Boukhris, Nesrine Keskes et Nour Hamza.

Je dédie ce travail,

A mon cher mari, mon ami et l'amour de ma vie, merci pour ta patience, pour tes sacrifices, pour ton amour, pour ton soutien et pour ton assistance. A mes chers parents, les prunelles de mes yeux, sans vous, je ne serai certainement pas arrivée à achever ce travail. Ma chère maman, merci pour ton amour, ta tendresse et ton soutien. Mon cher papa, tu sais que tu es mon partenaire pour ce travail, ... alors voilà, nous y sommes enfin, ... et au suivant ! A mes chers fils, la joie de mon existence, Midou, Koukou et Didou, vous allez enfin avoir une maman à temps complet, ...bon, pas à ce point quand même, je ne prends pas ma retraite ! ... j'ai hâte de rattraper le temps avec vous. A mes chers beaux-parents, merci pour tout l'amour dont vous ne cessez de me faire preuve, sans votre soutien je n'aurai pas pu y arriver. A mes deux « Hanene »s, mes sources de tendresse et de soutien aux moments difficiles, ma chère belle sœur « Wifek », tu es une vraie sœur et tu es belle. A mon frère « Mohamed Amin », nos deux princesses « Lina et Nour », mes frères « Anis » et « Anouar », Merci.

Sommaire	i
CHAPITRE 0 : INTRODUCTION GENERALE	1
CHAPITRE I : THEORIE DES POSSIBILITES : POSITIONNEMENT, RECOURS, PRINCIPES ET NOTIONS	14
I. INTRODUCTION	17
II. CONCEPT DE DONNEE, INFORMATION, CONNAISSANCE, COMPETENCE :	18
1. Donnée	18
2. Information	18
3. Connaissance	19
4. Compétence, savoir-faire, expertise	20
III. CONTEXTES INCERTAINS	21
1. L'incomplétude	21
2. L'Ambigüité	22
3. Le conflit	22
4. L'incertitude	22
5. L'imprécision	23
6. Liens entre les imperfections	24
7. Positionnement épistémologique de la recherche dans le domaine de prise de décision	24
IV. THEORIES DE L'INCERTAIN	24
1. Théorie des probabilités	25
2. Théorie des fonctions de croyance (Dempster-Shafer)	25
3. Théorie des ensembles flous	26
4. Théorie des possibilités	26
5. Positionnement paradigmatique des théories de l'incertain	26
6. Démarche de la recherche : recours à la théorie des possibilités	28
V. OUTILS DE LA THEORIE DES POSSIBILITES	29
1. Définition d'une distribution de possibilité	29
a) Formalisation du jugement d'expert	29
b) Définition formelle d'une distribution de possibilités	29
c) Mesures de possibilité et nécessité	30
d) Hauteur et inconsistance	32
2. Normalisation d'une distribution de possibilités	33
a) Normalisation ordinale	33
b) Normalisation numérique	34
c) Normalisation par décalage de l'inconsistance	34
3. Mesures d'incertitudes	35
a) Mesure de spécificité S	35
b) Mesure d'incertitude U	36
c) Indice de confiance	36
4. Outils de fusion	38
a) Propagation à travers un modèle analytique	38
b) Combinaison par les opérateurs ensemblistes	38
5. Critères de décision	41
a) Mesure de possibilité (resp. nécessité) maximale	42
b) Indice de confiance maximal	42
c) Moyenne maximale	42
d) Préférence sur le couple $[\Pi, N]$	42
e) Mesure de similarité maximale	43
VI. ETAT DE L'ART DES METHODES D'ESTIMATION D'UNE DISTRIBUTION DE POSSIBILITES	43
1. Les méthodes d'estimation guidées par l'expert	43
a) Approche directe	44
b) Approche indirecte	44
2. Transformation d'une distribution de probabilités en une distribution de possibilités	45
a) Principes de cohérences entre les distributions probabilité/Possibilité	45
b) Estimation de distribution de possibilités à partir de mesures précises de probabilités	47
c) Estimation de distribution de possibilités à partir des intervalles de confiance probabilistes	48
3. Estimation d'une distribution de possibilités à partir d'une fonction d'appartenance	50
4. Estimation de distribution de possibilités à partir d'une distribution de masse	50

VII.	CONCLUSION	50
CHAPITRE II :	MODELISATION POSSIBILISTE DE CONNAISSANCES	52
I.	INTRODUCTION	55
II.	SYNTHESE DES ACTIVITES RELATIVES A CE CHAPITRE	57
1.	Collaborations et encadrements	57
2.	Publications	58
a)	Articles de Revue	58
b)	Conférences internationales à comité de lecture	58
III.	ESTIMATION ET INFERENCE DE DISTRIBUTIONS DE POSSIBILITES POUR LA MODELISATION DE DONNEES TENANT COMPTE DE LEURS ALTERITES	59
1.	Modification apportée à la technique d'inférence de Masson et Denoeux	60
2.	Mise en œuvre d'une stratégie d'évaluation	61
a)	Description de la stratégie	61
b)	Approches à comparer	61
c)	Données utilisées pour les expérimentations	61
d)	Critères d'évaluations	62
e)	Choix d'une mesure de similarité	63
3.	Résultats	65
a)	Evaluation qualitative : critères psycho visuels	65
b)	Evaluation quantitative	67
c)	Discussions	71
IV.	MODELISATION POSSIBILISTE DE CONNAISSANCES CONTEXTUELLES DE VOISINAGE	72
1.	Modes de représentations de connaissances pixelliques	72
2.	1ère Contribution : Modélisations possibilistes de connaissances pixelliques	74
a)	Définition du contexte	74
b)	Problématique	75
c)	Hypothèse 1 : Modélisation possibiliste ponctuelle du pixel	75
d)	Hypothèse 2 : Modélisation possibiliste contextuelle du pixel	77
3.	Résultats et discussion	78
a)	Contexte des simulations	78
b)	Tests et évaluation sur une image de synthèse	80
c)	Tests et évaluation sur des images médicales optiques	84
4.	2ème Contribution (en cours de développement) : Modélisation contextuelle spatiale locale de deuxième ordre possibiliste d'un pixel ; Notion de possibilité des possibilités	87
a)	Description de l'idée	87
b)	Simulation sur un échantillon	89
V.	SELECTION D'ATTRIBUTS EN MODELISATION POSSIBILISTE	90
1.	Problématique	90
2.	Cadre de l'étude	91
3.	Outils techniques	92
a)	Indice de Shapley	92
b)	Mesure d'incertitude	93
4.	Description de la contribution	93
5.	Expérimentations et discussion	98
VI.	CONCLUSION	101
CHAPITRE III :	AIDE A LA DECISION DANS UN CONTEXTE POSSIBILISTE	103
I.	INTRODUCTION	105
1.	Base de cas et de connaissances	105
a)	Base de cas	105
b)	Base de connaissance	106
2.	Processus d'association de cas ou de connaissances à des décisions	107
II.	SYNTHESE DES ACTIVITES RELATIVES A CE CHAPITRE	108
1.	Collaborations et encadrements	108
2.	Publications	109
III.	MESURE DE SIMILARITE	110
1.	Notion de similarité dans un contexte possibiliste	111

2.	Propriétés d'une mesure de similarité possibiliste	112
3.	Etat de l'art sur les mesures de similarité en théorie de possibilités	112
a)	Mesure de similarité « Information Closeness »	112
b)	Mesure de similarité déduite de l'information de divergence	113
c)	Mesure de Sangüesa :	113
d)	Distance de Minkowski et ses dérivés	113
e)	Mesure de similarité delta δ	114
f)	Mesure de similarité « Info Affinity »	114
4.	Analyse de l'état de l'art	114
a)	Simulation comparative sur des exemples numériques	114
b)	Critique	114
c)	Motivation	115
5.	Contribution : Proposition d'une nouvelle mesure de similarité « Info-spécificity »	116
a)	Principe	116
b)	Définition	116
6.	Méthodologie d'évaluation	117
a)	Facteurs influant sur l'évaluation	117
b)	Données utilisées pour les expérimentations	117
c)	Approches d'estimation de mesures de similarité à comparer	118
d)	Description de la procédure d'évaluation	118
7.	Résultats et discussions	122
8.	Emploi de la mesure InfoSpécificity dans la construction d'un arbre de décision possibiliste	125
a)	Notions préliminaires concernant un arbre de décision	125
b)	Contribution et application	126
IV.	TRAVAUX DANS LE CADRE DE LA NOTION DE FUSION	127
1.	Sélection d'opérateurs de fusion dans un cadre possibiliste	127
a)	Approche de Weiru Liu [Liu 2007]	128
b)	Modification proposée	129
c)	Expérimentations et résultats	130
2.	Analyse multicritère de la décision : MCDA	135
a)	Objectif	135
b)	Contexte	136
c)	Approche	136
V.	FOUILLE POSSIBILISTE	137
1.	Construction d'une carte possibiliste	138
2.	Prise de décision basée sur les cartes possibilistes	139
3.	Application	140
VI.	CONCLUSION	140
	CHAPITRE IV : INTERPRETATION DE SCENE	142
I.	INTRODUCTION	144
II.	SYNTHESE DES ACTIVITES RELATIVES A CE CHAPITRE	146
1.	Collaborations et encadrements	146
2.	Publications	147
a)	Articles de Revue	147
b)	Conférences internationales à comité de lecture	147
III.	CONSTRUCTION DE SCENES	148
1.	Scène Biométrique	148
a)	Conception du protocole	148
b)	Conception des deux irisopes	148
2.	Scène d'environnement intérieur	150
a)	Description	150
b)	Schéma fonctionnel	152
IV.	DIFFUSION DE CONNAISSANCES POSSIBILISTES ITERATIVES, APPLICATION A L'INTERPRETATION DE SCENE MEDICALE	153
1.	Problématique	153
2.	Principe	155
a)	Système de prétraitement :	156

b)	Système de décision :	158
c)	Système de diffusion possibiliste :	158
d)	Système d'évaluation et de mise-à-jour :	158
3.	Diffusion classe/classe	159
4.	Diffusion croisée	161
5.	Contribution : Diffusion croisée avec une représentation possibiliste des connaissances dans une clique	165
a)	Description du processus	165
b)	Simulation sur un échantillon	167
6.	Résultats et discussions	170
a)	Estimation initiale des cartes possibilistes	170
b)	Diffusion itérative Classe/Classe des connaissances possibilistes	171
c)	Diffusion itérative Croisée des connaissances possibilistes	175
d)	Effet de l'initialisation des connaissances possibilistes	177
e)	Utilisation de la diffusion croisée pour la segmentation d'images mammographiques	180
f)	Effet de la modélisation possibiliste du contenu informationnel de la clique sur la diffusion croisée	181
V.	FOUILLE POSSIBILISTE D'IDENTITE, INTERPRETATION DE SCENE BIOMETRIQUE : NOTION DU CARNET D'IDENTITE	185
VI.	CONCLUSION	186
CHAPITRE IV :	CONCLUSION GENERALE ET PERSPECTIVES	189
REFERENCES		199

Chapitre 0 : Introduction générale

Dans ce document, je présente une synthèse de mes travaux de recherche menés après la soutenance de ma thèse en février 2010. Mes travaux de recherche s'articulent essentiellement autour du développement de mécanismes de raisonnement dans un univers incertain. Dans mes travaux de recherche, je me suis spécialement intéressée à la théorie des possibilités. J'ai étudié et exploité cette théorie pour d'abord modéliser des connaissances mal-définies pour ensuite employer ces modèles pour interpréter et aider à la prise de décision dans des contextes incertains.

Nous assistons à l'extraordinaire expansion des données accessibles par ordinateur sur toutes sortes d'activités humaines. La disponibilité de ces grands volumes de données de grande diversité et nos capacités limitées pour les traiter efficacement créent un fort besoin de nouvelles méthodologies pour en extraire des connaissances utiles. Il existe également un besoin de méthodologies pour dériver des connaissances plausibles à partir de données petites et indirectement pertinentes, car dans de nombreux domaines pratiques, seules ces données peuvent être disponibles, par exemple détection de fraude, prévention du terrorisme, détection d'intrusion informatique, diagnostic précoce du cancer, etc. Afin de satisfaire ce besoin, les chercheurs ont exploré des idées et des méthodes développées dans l'apprentissage automatique, l'analyse de données statistiques, la fouille de données, la fouille de texte, la visualisation de données, la reconnaissance de formes, etc.

Par exemple, une analyse statistique peut déterminer des distributions, des covariances et des corrélations parmi des variables extraites des données, mais n'est pas capable de caractériser ces dépendances à un niveau conceptuel abstrait comme peuvent le faire les humains et produire une explication causale dégageant ces dépendances. Encore, si une analyse de données statistiques peut déterminer les tendances centrales et la variance de facteurs donnés, elle ne peut pas produire une description qualitative des régularités, ni ne peut déterminer une dépendance aux facteurs non explicitement fournis dans les données [KM 2005]. De même, une technique de classification classique peut créer un rangement des entités et spécifier une similarité numérique entre les entités assemblées dans des catégories identiques ou différentes, mais elle ne peut à elle seule construire des descriptions qualitatives des classes créées et présenter une justification conceptuelle pour inclure les entités dans une catégorie donnée. Les attributs et les méthodes utilisés pour mesurer la similarité doivent être préalablement spécifiés par un analyste de données. Les arbres de décision populaires peuvent représenter une relation entre les variables d'entrée et de sortie, mais leur pouvoir de représentation est très modeste. Ces méthodes peuvent ainsi produire un arbre très complexe même pour une relation conceptuellement simple. De même, les règles d'association, qui sont populaires dans la fouille de données, ont un pouvoir de représentation limité. Les réseaux bayésiens classiques sont très attrayants pour de nombreuses applications, mais ils s'appuient généralement sur la contribution humaine en ce qui concerne leur structure, et ne peuvent déterminer automatiquement que des relations relativement simples entre les attributs ou les concepts [KM 2005].

Les méthodes ci-dessus créent généralement des modèles qui utilisent uniquement les attributs présents dans les données. Elles ne permettent pas d'établir une base de connaissances afin de générer automatiquement d'autres attributs pertinents ou encore de prendre en compte les imperfections qui peuvent entacher les données qui ont servi à la construction de ces modèles. Leur mode de raisonnement s'appuie sur l'hypothèse de la réalité objective des paramètres et donc du modèle qui les définit.

Dans les cas où l'objectif est d'aborder des scènes ou situations compliquées, où les données et les informations sont insuffisantes pour lever complètement les incertitudes, et que l'on n'a pas suffisamment de connaissances *a priori* à injecter à fin de formuler les connaissances de base pour un apprentissage, des outils se basant sur des théories dites de l'incertain, semblent être beaucoup plus adéquates.

Dans ce cadre, mes travaux d'Habilitation contribuent au développement d'approches d'interprétation de scènes et de prise de décisions après modélisation adéquate des données issues de différentes sources. L'objectif est de concevoir un système d'analyse et d'interprétation de connaissances, capable d'effectuer un raisonnement symbolique impliquant la prise en compte des détails factuels de la scène à interpréter, tels que les imperfections. Notre intérêt s'est spécialement focalisé sur les scènes affectées d'ambiguïté.

Dans de tels contextes, l'utilisation de la théorie des possibilités pour la modélisation des connaissances est favorisée du fait que cette théorie permette efficacement de modéliser les connaissances faiblement définies «weak defined knowledge». En effet, la modélisation est assurée par des distributions de possibilités et le niveau de connaissances injecté pour la construction de ces distributions de possibilités, est un faible niveau de connaissance. Un comportement stable et une efficacité supérieure dans des conditions de faibles connaissances ont été appréciés suite à l'étude et les expérimentations menées.

Dans la majorité de nos travaux, nous avons adopté une méthodologie opérationnelle. En effet, les aspects théoriques étudiés et les approches développées, ont été testés et validés sur des scènes réelles. Nous avons soigneusement choisi nos scènes d'application, de façon à varier l'origine de l'ambiguïté d'une part et son niveau d'impact d'autre part. Ainsi, les approches de modélisation possibilistes développées ont été dédiées à l'interprétation de scènes médicales, scènes biométriques et scène d'environnement intérieur. Dans le premier type de scène, l'ambiguïté venait du faible contraste dans les images médicales, relié aux niveaux de gris étant les mesures radiométriques acquises. Cette ambiguïté affecte la localisation ainsi que la reconnaissance des zones suspectes. L'ambiguïté dans une scène biométrique vient de la nature même de l'élément biométrique à modéliser. En effet, entre deux éléments biométriques de même nature mais appartenant à deux individus différents, il existe beaucoup plus de ressemblance qu'il y en a de différence. Dans la troisième scène, qui consiste en un environnement d'intérieur exploré par un système à multi-capteurs (optique, de profondeur et de position) pour guider la navigation, l'ambiguïté réside dans deux aspects : le premier concerne la qualité médiocre des

données acquises. En effet, un défi a été lancé par l'emploi de capteurs à bas prix fournissant ainsi des données de très modeste qualité, pour formuler des connaissances à haut niveau. Le deuxième aspect d'ambiguïté vient du fait d'avoir à formuler une décision en se basant sur plusieurs sources de données. Une autre forme d'ambiguïté concerne à la fois les trois scènes, est reliée à la disponibilité même des données. En effet, comme dans la majorité des applications d'ingénierie, les données acquises forment des échantillons de petites tailles. Cette limite est souvent engendrée par la contrainte temps réel, mais dans certains cas c'est une limite d'ordre physique.

Mes collaborations et mes participations à des encadrements ont été à l'origine de contributions sur le développement de certaines approches de modélisation et de d'interprétation de scènes pour la prise de décision, ainsi que de stratégies, scénarios et même dispositifs de test pour la validations de ces approches.

Dans cette introduction, après un bref rappel de mes travaux de thèse, je donne les points importants de mon activité de recherche menée depuis l'année 2011.

Travaux de thèse

Le point de départ de mes activités de recherche a été ma thèse de doctorat encadrée par Pr. Dorra Sellami et sous la direction de Pr. Nabil Derbel, professeurs de l'ENIS.

Dans cette thèse, on s'est intéressé à la vérification des personnes par reconnaissance de l'iris, réputé être l'élément biométrique externe le plus efficace vis-à-vis de son utilisation pour l'identification. L'utilisation de l'iris pour une vérification d'identité nécessite de suivre une démarche, résumée en cinq étapes, à savoir l'acquisition, la localisation de l'iris, sa normalisation, l'analyse de la texture pour l'extraction de la signature et enfin la décision. Dans ce cadre, nous avons d'abord, conçu une approche originale de traitement d'image pour délimiter l'iris dans une image d'œil. Nous avons utilisé une technique de segmentation basée sur un seuillage d'histogramme pour détecter le contour interne et une autre technique basée sur un balayage radial pour détecter le contour externe de l'iris. Les résultats expérimentaux, ont montré l'efficacité et la robustesse de la méthode sur les 756 images de la première version de CASIA. En effet, le taux de bonne détection du contour interne est de 100%, celui du contour externe est de 98,54%. Une stratégie de correction des déformations géométriques, de la non uniformité de l'éclairage et du faible contraste de la texture de l'iris, a été ensuite conçue afin d'avoir une représentation uniforme de l'iris, s'avérant être nécessaire afin que l'appariement entre les images soit significatif. Cet appariement ne s'applique pas directement aux images, mais plutôt à des représentations numériques de ces images qui se construisent sur la base d'une analyse des textures respectives des iris à apparier. Pour ceci, nous avons conçu deux approches visant à spécifier de façon unique un iris : la première se base sur la redondance d'information générée par une analyse par l'ondelette de Meyer de la texture de l'iris. Cette technique génère une signature iridienne de 6426 bits en 1,07 secondes (Le temps d'exécution est celui d'une

simulation sur le logiciel Matlab 7.01, effectuée par une machine classique de microprocesseur pentium 4 et 512Mo de RAM. Cette information donnée pour une évaluation relative et non absolue), et permet un taux d'identification de 91% et un taux d'authentification de 100%. La deuxième approche consiste en une quantification de la texture de l'iris par une analyse fractale. La signature déduite correspond à une représentation de la variation des dimensions fractales calculées sur 256 blocs de l'image. Celle-ci permet de générer une signature iridienne de 255 bits, en uniquement 0,63 secondes (dans les mêmes conditions d'exécution citées précédemment) et de réaliser un taux d'identification de 84% et un taux d'authentification de 96%.

La décision sur l'identification d'un candidat est basée sur l'évaluation de la distance de la signature du candidat par rapport aux signatures stockées dans la base de référence (base de cas). Les candidats légitimes sont, enfin, reconnus par leurs numéros d'identification personnels (NIP) respectifs. L'évaluation de la décision, ainsi prise, est effectuée via quatre critères de mesure de performances : distribution des distances, aire sous la courbe ROC, taux de reconnaissances à taux d'erreurs égales et position du NIP. Les nouvelles approches proposées ont prouvé leurs efficacités et ont manifesté les meilleures performances vis-à-vis de celles étudiées et implémentées de la littérature. Particulièrement, la technique de " Démodulation par l'ondelette de Meyer " a engendré les meilleurs résultats selon tous les critères d'évaluation utilisés.

Les perspectives qui ont été envisagées à l'issu des travaux menés dans le cadre de ma thèse de doctorat, ont été les points de départ pour mes travaux de recherches actuels. En effet, nos perspectives, dont certaines étaient explicitement citées et d'autre implicitement comprises, consistaient à :

- ✓ Creuser dans d'autres théories afin de proposer d'autres modèles pour représenter les connaissances incluses dans une texture d'iris. Des modèles probabilistes Markoviens et des modèles flous ont été précisément indiqués pour exploration. On trouve ici une orientation vers les théories de l'incertain.
- ✓ Etudier et concevoir des systèmes biométriques multimodaux. Il s'agit d'une application directe des systèmes multi-sources, impliquant des notions de modélisation et de fusion. Notons que les systèmes multi-sources présentent l'un des contextes type nécessitant l'usage de la théorie des possibilités.
- ✓ Concevoir et construire une base authentique d'images d'iris de façon à contrôler les conditions de prise de vues et s'orienter vers la réalisation d'un système d'identification biométrique par reconnaissance d'iris commercialisable. Il s'agit ici effectivement d'une construction de scène.
- ✓ La création et l'emploi d'une base de connaissances dans une procédure de prise de décision ont également été implicitement indiqués parmi nos perspectives.

Ces perspectives nous ont guidés vers l'exploration des axes de recherche sur les quels ont été menés nos travaux actuels, à savoir la modélisation, l'aide à la décision et l'interprétation de scènes dans un contexte incertain. Le schéma de la Figure 0.1 illustre, en quelque sorte, l'évolution qu'a suivi nos idées et nos recherches depuis les travaux de thèse.

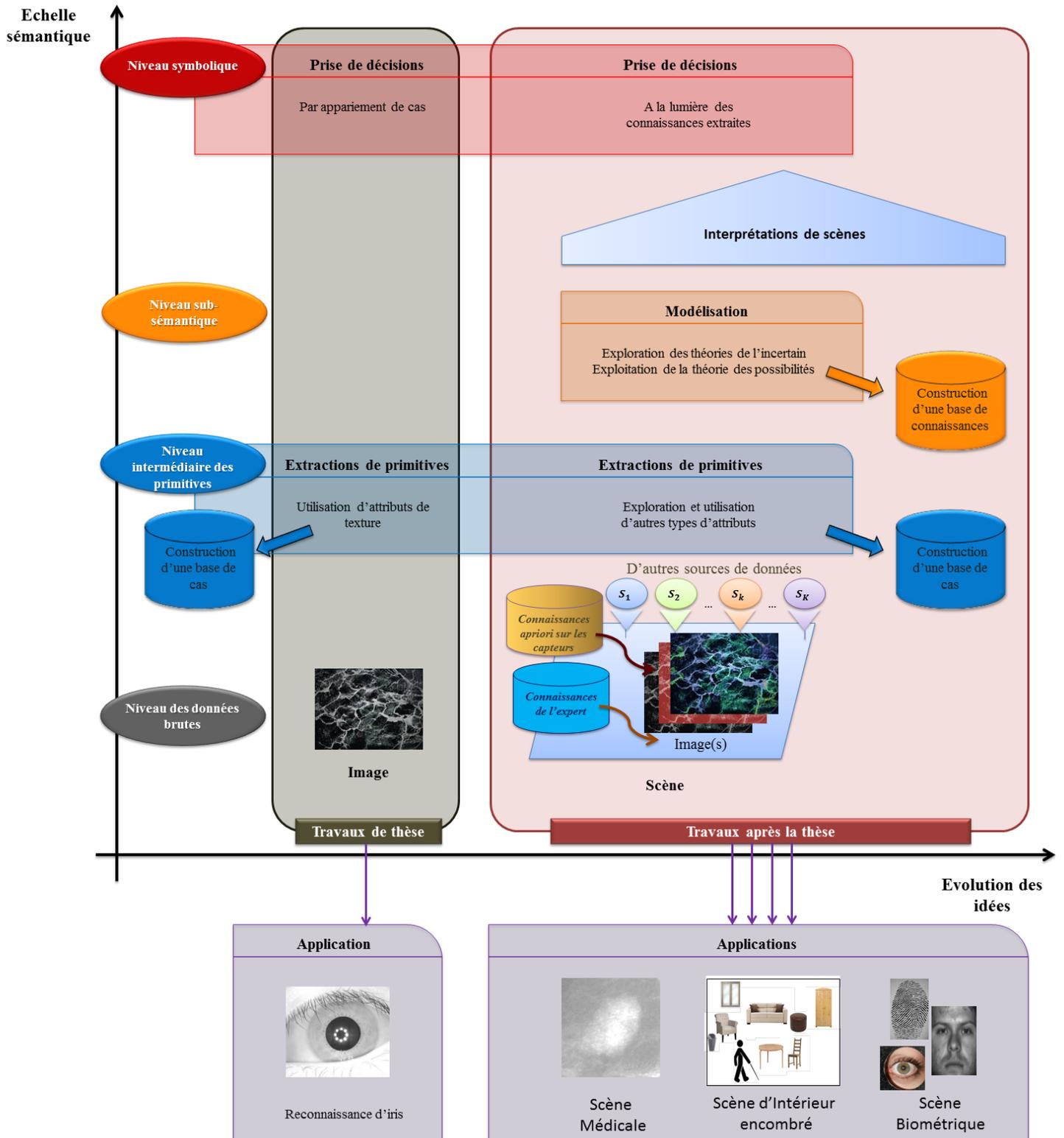


Figure 0.1 Evolution des idées depuis les travaux de thèse

Travaux sur la modélisation de connaissances

Le sujet de la modélisation des informations a toujours suscité beaucoup d'intérêt et a été étudié sous différents axes de recherche comme un moyen prometteur pour combattre le « fossé sémantique ». En effet, les informations contenues dans des données radiométriques (ou brutes) sont une approximation de la scène, du monde réel, observée par des capteurs. Pour en tirer une interprétation, il est nécessaire de les représenter sous une forme numérique ou symbolique permettant leur traitement. Une représentation immédiate et classique de ces informations, dite de bas niveau, fournit un modèle très limité pour comprendre la situation du monde réel. C'est ce qu'on appelle le « fossé sémantique » entre les informations de bas niveau d'une scène et ses connaissances de haut niveau (la compréhension du monde réel). Une formulation plus élaborée de ces données, faisant intervenir en plus de leurs valeurs, leurs paramètres et des relations qui les relient, permet une représentation plus sémantique. Cependant, une représentation universelle et valable pour tout type de contexte n'existe pas. D'une part, les données sont toujours affectées par des détails reliés à leurs conditions de collecte et au dispositif utilisé pour leurs captures. D'autre part, leurs valeurs n'ont de sens valable qu'une fois reliées à leur contexte sémantique d'exploitation. L'objectif de nos travaux reliés à la problématique de modélisation, est de concevoir des approches possibilistes de modélisation qui peuvent être efficaces dans des contextes d'incertitude fréquemment rencontrés dans des applications nécessitant des prises de décision.

Ce thème sur la modélisation de connaissances est articulé autour de 4 volets de recherche. Les deux premiers volets visent à la construction du modèle alors que les deux autres sont entretenus dans l'objectif de raffiner ou améliorer la modélisation dans un cadre possibiliste.

Volet 1. Estimation et inférence de distributions de possibilités pour la modélisation de données tenant compte de leurs altérités (disponibilités, natures, quantité, imperfection)

L'estimation d'une distribution de possibilités est l'une des problématiques les plus recherchées dans la littérature. Ceci n'est pas étrange, vu l'importance qu'occupe la notion de distribution de possibilités dans la théorie. Globalement, il y a deux classes d'approches pour la construire. L'une se base particulièrement sur les connaissances de l'expert. Elle ne peut, de ce fait, être complètement objective. L'autre classe d'approches consiste à l'approximation ou l'inférence de cette distribution à partir d'autres fonctions caractérisant les connaissances disponibles, telles que une distribution de probabilité, une fonction d'appartenance ou encore une fonction de masse. L'inférence la plus utilisée dans la littérature est la transformation d'une distribution de probabilité en une distribution de possibilités. Celle-ci nécessite qu'ait d'abord une distribution de probabilité fiable, ce qui n'est pas toujours réalisable, puisque les données disponibles sont souvent du type incomplet, tant en quantité qu'en qualité. Une alternative à l'utilisation des valeurs de probabilité, pour la construction de distributions de possibilités

épistémiques (à partir d'observations du monde réel), consiste en l'utilisation des intervalles de confiance probabilistes. Cette idée existe déjà dans la littérature. Nous lui apportons une légère modification afin de l'adapter aux cas où l'on ne dispose que de très peu de données : des échantillons de petites à très petites tailles. La modification consiste en l'utilisation des intervalles de confiance simultanés de Sison and Glaz. Notre apport se situe également dans la stratégie conçue pour l'évaluation des approches d'estimation des distributions de possibilités et la validation de la modification apportée sur les cas des échantillons de petite taille.

Volet 2. Modélisation possibiliste de connaissances contextuelles de voisinage,

Concevoir un modèle possibiliste, qui consiste en une distribution de possibilités, pour une classe thématique, est usuel. Sa conception se fait en se basant sur les informations apportées par l'ensemble des observations des éléments de la classe thématique à modéliser. En revanche, quand il s'agit de construire un modèle pour l'un des éléments de cette classe, la question n'est plus usuelle. La connaissance reliée à un élément isolé est généralement réduite aux données radiométriques mesurées ou capturées ou observées au niveau de cet élément. Ces données brutes se situent à la base de l'échelle sémantique de l'information. Transformer ces données abstraites en une représentation conceptuelle nécessite l'injection d'autres connaissances. L'idée à ce niveau est de faire contribuer le voisinage immédiat de l'élément à modéliser pour apporter ces connaissances supplémentaires à injecter à fin de concevoir son modèle. Cette initiative prend sa légitimité du fait qu'un élément observé dans une scène donnée n'est jamais isolé, il est lié physiquement à son voisinage. De cette liaison physique découle la dépendance des connaissances. Ainsi, le contenu informationnel d'un élément est enrichi de celui de son voisinage modéré d'un ordre de crédibilité lié au degré d'éloignement de chaque voisin, puis employé pour la construction d'une distribution de possibilité modélisant l'élément en question.

Cette modélisation possibiliste de connaissances contextuelles de voisinage, est exploitée dans une scène image pour résumer des connaissances pixelliques pour une classification pixellique. L'élément est donc un pixel et son voisinage est du type spatial.

Volet 3. La notion de possibilité des possibilités

Dans un problème de prise de décision donné, les décisions élémentaires représentent des alternatives d'un ensemble de discours. Pour une prise de décision dans un contexte possibiliste un ordonnancement préférentiel est à accomplir sur l'ensemble des alternatives selon leurs possibilités d'occurrence. Accorder une possibilité d'occurrence à une alternative peut être vu comme une projection de cette alternative dans un espace sémantique supérieur à celui des observations, à savoir celui des possibilités. Si on considère que cet espace des possibilités soit un espace immédiat pour estimer que l'occurrence de l'alternative est possiblement vraie, l'idée ici est d'envisager un autre espace sémantique encore plus profond qui relativise le premier, dans lequel, la possibilité attribuée à cette alternative apparaît elle-même comme une alternative dans

un autre ensemble de discours, celui des possibilités d'occurrence. Dans ce cas la connaissance reliée à l'alternative du monde réel sera une possibilité sur sa possibilité d'occurrence.

Cette notion est en cours de développement. Les premiers tests révèlent déjà un grand potentiel.

Volet 4. Sélection d'attributs en modélisation possibiliste

La problématique de sélection d'attributs ne cesse de susciter l'intérêt des chercheurs dans les domaines en relation avec le traitement de l'information dans son sens général. La motivation de cet intérêt se doit essentiellement à la réduction de la complexité du traitement d'une part et la restriction de la construction de modèles aux attributs les plus performants uniquement, d'autre part. C'est la même motivation qui nous a entraînés sur ce terrain, avec la différence que nos modèles représentent des connaissances incertaines. Notre apport consiste à diffuser le raisonnement approximatif apporté par la théorie des possibilités sur la procédure de sélection des attributs, et ce par l'utilisation de critères de préférences intégrant à la fois une mesure d'incertitude et le formalisme de Shapley. Selon notre approche, un attribut est sélectionné s'il montre son intérêt à la fois dans la représentation et la séparation des classes. Une stratégie originale de gestion de conflits entre les deux tests (représentation/séparation) a été conçue.

Travaux sur la prise de décision

Les systèmes d'aide à la décision ont l'ambition d'assister, reproduire ou remplacer le raisonnement humain. Une modélisation d'une situation de prise de décision effectuée par l'humain, consiste à considérer un univers composé de M décisions élémentaires (classes thématiques, hypothèses, etc.) connues a priori, sur lequel les connaissances disponibles sont de nature ambiguë (on sait que ça doit impérativement être l'une de ces décisions à prendre, mais on ne sait pas laquelle). La prise de décision s'effectue en attribuant chaque élément issu du monde réel à une décision élémentaire parmi les différentes décisions élémentaires envisagées. L'élément issu du monde réel est représenté par un ensemble de données collectées à partir d'une ou plusieurs sources d'information, qui peuvent être entre autre un ou plusieurs capteurs mesurant une quantité physique caractérisant cet élément. L'ensemble des valeurs prises par ces données, compose l'univers du contenu informationnel. Il s'agit d'attribuer, à la lumière des connaissances disponibles (en se basant sur le contenu informationnel le caractérisant), chaque élément du monde réel à l'une et une seule des classes thématiques. Ce contexte d'incertitude est idéalement traité si nous disposons de connaissances probabilistes liées aux différentes distributions de probabilités (a priori, de vraisemblance et a posteriori). Néanmoins, les connaissances disponibles sont de nature ambiguë décrivant des degrés d'harmonie ou d'appartenance partielle des éléments du monde réel aux différentes classes thématiques. Ce contexte, pouvant être considéré comme un modèle d'un raisonnement humain pour une prise de décision, est bien adapté à l'application de la théorie des possibilités, et c'est dans ce contexte que sont menés nos travaux sur la problématique de prise de décision.

Nos travaux ont touché, d'une part, à la conception de systèmes de prise de décision appropriés à des applications notamment de classification et de segmentation et d'autre part à l'étude et développement des concepts de similarité et de fusion, étant des notions directement liées aux processus de prise de décision.

Volet 5. Mesures de similarité : conception d'une nouvelle mesure de similarité dans un contexte possibiliste « info-specificity »

Dans un processus de prise de décision, le raisonnement par similarité est une forme de raisonnement par analogie, consistant naturellement à interpréter un nouveau problème en le comparant avec un problème similaire déjà rencontré et résolu. Ce type de raisonnement est une copie du comportement humain pour résoudre les problèmes de la vie quotidienne en faisant appel à son expérience. Dans le contexte possibiliste pour la prise de décision précédemment décrit, la résolution d'un problème consiste à accorder une décision à un élément du monde réel, sachant que les décisions élémentaires ou classes thématiques sont représentées par des modèles possibilistes (des distributions de possibilités). Dans une optique d'imiter ce type de raisonnement, la quantification du degré de similarité entre deux entités que ce soit de même ou de différentes granularité (élément-élément, élément-classe, classe-classe) est d'une importance primordiale. Cette quantification se fait au moyen des mesures de similarités.

Nous avons eu recours à l'utilisation des mesures de similarité, pratiquement dans la quasi-totalité de nos travaux. Cette fréquence d'utilisation nous a révélé certaines faiblesses des mesures usuelles, une fois employées dans des contextes possibilistes. Cette déduction a motivé nos recherches pour concevoir une nouvelle mesure de similarité répondant d'une part aux propriétés fondamentales de la similarité et s'adaptant, d'autre part, aux spécificités sémantiques apportées par l'approximation possibiliste du monde réel. Nous l'avons appelée « Info-specificity ».

Volet 6. Fusion de connaissances

La fusion d'informations consiste à combiner des informations issues de plusieurs sources afin d'améliorer la prise de décision. La fusion peut se définir aussi comme étant une combinaison d'informations imparfaites et hétérogènes afin d'obtenir des informations plus complètes qui permettent une meilleure prise de décision. La qualité de la fusion dépend étroitement de la qualité des sources fournissant l'information, la qualité de l'information fournie, et le mode et les opérateurs assurant la combinaison. Toutes ces notions sont des objets d'études et de développement courant nos travaux de recherche.

Dans ce cadre, une stratégie de sélection d'opérateurs de fusion dans un cadre possibiliste a été conçue et expérimentée sur une variété de données synthétiques et réelles. Nous avons également mis en place une première conception pour un système dynamique incrémental de fusion multimodale, qui prend en compte la qualité tant des capteurs que de l'information de chacun des

éléments intervenant dans la fusion. Ce système assure une analyse multicritère de la décision. Dans ce même sens, nos travaux actuels sont condensés sur l'étude de la qualité de l'information et l'élaboration de métriques permettant son évaluation en se basant sur des concepts alliés à l'incertitude.

Volet 7. Mise en place de systèmes d'aide à la prise de décision dans des contextes incertains

Une prise de décision possibiliste commune, consiste à définir une règle de décision permettant d'identifier, à partir de la connaissance des distributions de possibilités des décisions élémentaires, celle la plus vraisemblable. Il est à noter qu'un apport essentiel de la théorie des possibilités réside dans la définition de deux mesures ensemblistes, à savoir la mesure de possibilité et la mesure de nécessité, qui ont contribué à la formulation de ces règles de décision. Cette action de prise de décision est généralement l'étape qui clôture toute procédure d'un système d'aide à la prise de décision. Dans nos travaux menant à la prise de décision deux procédures ont été développées puis appliquées dans différentes applications.

La première consiste en une fouille possibiliste. Il s'agit d'une approche qui s'intègre dans le processus d'extraction des connaissances possibilistes à partir des données observées du monde réel. Si la construction des distributions de possibilités est considérée l'approche possibiliste basique pour la représentation de connaissances possibilistes, la fouille possibiliste est destinée pour découvrir de nouvelles connaissances en se basant sur ces distributions de possibilités. Dans nos travaux, cette approche est basée sur la construction de « cartes possibilistes ». Une carte possibiliste relative à une décision élémentaire est définie comme une représentation conceptuelle de la décision élémentaire en question, prenant en compte toutes les distributions de possibilités conditionnelles à cette décision élémentaire, relatives à tous les ensembles de sources d'information. Cette approche a été employée pour une classification d'échos ultrason acquises sur des états du sol, en trois classes : sol plat, escaliers descendants et escaliers ascendants. Nous l'avons nommée fouille temporelle. La prise de décision est effectuée sur la base d'estimation de similarités sur les cartes possibilistes correspondant respectivement aux classes et aux échantillons tests. L'emploi des cartes possibilistes a également servi pour des interprétations de scènes.

La deuxième procédure proposée pour une aide à la prise de décision est une modification apportée au concept connu des arbres de décision. Il s'agit d'une transformation de l'algorithme populaire C4.5, au niveau de la phase de partitionnement. L'approche a été utilisée dans la reconnaissance d'objets dans un environnement domestique.

Travaux sur la construction et interprétation de scène

Une scène peut être définie en tant qu'un sous-ensemble d'un environnement au sein duquel sont considérés de multiples composants, reliés à la fois physiquement et sémantiquement. Les recherches effectuées sur le concept de scène concernent principalement sa définition, sa

modélisation, sa reconnaissance et son interprétation. Notre intérêt s'est orienté surtout sur l'interprétation de scènes en se basant sur un raisonnement possibiliste.

Courant mes travaux de recherche, trois types de scènes ont été considérés pour supporter l'expérimentation des processus conçus que ce soit de modélisation, prise de décision ou pour l'interprétation. La première est une scène médicale comportant des données imagées collectées à partir d'un patient en vue d'être diagnostiquées aux lumières des connaissances apportées par un médecin. La deuxième est une scène biométrique comportant des acquisitions graphiques d'éléments biométriques de l'individu en vue de son identification aux lumières des connaissances de l'expert en caractéristiques physiologiques uniques de l'être humain. La troisième consiste en une scène d'un intérieur meublé en vue d'une navigation sécurisée destinée à venir en assistance à des personnes ayant une défaillance visuelle. Des scènes du premier type nous étaient fournies par des experts dans le domaine médical. Pour les deux autres types, le manque de disponibilité d'informations nécessaires au développement de nos travaux, nous a incités à construire nos propres scènes.

Volet 8. Construction de scènes

Nos travaux dans ce volet étaient assistés par des compétences dans des domaines de la conception électronique et développement de solutions informatiques, auxquelles nous avons fait appel pour construire des dispositifs multicapteurs : deux dispositifs pour l'acquisition de l'iris et un dispositif d'acquisition d'environnement intérieur. Notre apport dans ces travaux se situe au niveau de l'étude des besoins, le choix des capteurs et le design technique d'une part, et de la mise en place des cahiers de charge respectifs de ces dispositifs, d'autre part. Les dispositifs réalisés ont servi correctement pour nous fournir les scènes visées. Deux brevets sont en vue, dont l'un est déposé et l'autre est en cours de préparation.

Volet 9. Interprétation de scène par diffusion de connaissances

L'incertitude que nous désirons maîtriser et élucider pour aboutir à des décisions conformes et cohérentes avec la réalité, ne vient pas de la scène elle-même, mais plutôt elle résulte de la manière de la voir et percevoir. En effet, deux fossés existent entre la réalité et sa compréhension par un système automatique. Le premier fossé est dit « sensoriel » et il est dû à la restriction de la scène à des présentations générées par les systèmes d'observations et de captures de la scène. Le deuxième fossé est dit « sémantique » et il concerne le manque de connaissances a priori nécessaires pour extraire la sémantique contenue dans l'image. Un processus d'interprétation de scène consiste à apporter des solutions aux problèmes liés à ces deux fossés.

La première solution que nous proposons, concerne une interprétation d'une scène médicale par propagation de connaissances possibilistes. L'objectif visé était la classification sémantique des pixels de l'image représentative de la scène en des régions représentant des classes thématiques.

La deuxième solution proposée, concerne une interprétation d'une scène biométrique par fouille possibiliste. L'objectif visé était l'extraction de nouvelles connaissances sur les relations

entre les différentes modalités biométriques intervenant dans la scène interprétée. Le résultat de l'interprétation est modélisé en une structure conceptuelle que nous avons appelée « Carnet d'identité biométrique »

Organisation du rapport

Dans ce cadre, ce rapport s'articule autour de quatre chapitres. Le premier chapitre est voulu être une référence aux notions nécessaires au suivi et la bonne compréhension des trois autres chapitres. Nous commençons par présenter le cadre de nos travaux et justifier le recours à la théorie des possibilités pour l'étude de l'incertitude. Ensuite, les principaux outils, notions et concepts reliés à cette théorie sont présentés. Les chapitres deux, trois et quatre sont destinés à la présentation des trois axes principaux autour desquels sont menées mes activités de recherche à savoir, respectivement : la modélisation possibiliste de connaissances, la prise de décision et l'interprétation de scène. Une méthodologie pour la présentation des travaux de recherche réalisés dans chaque thématique a été adoptée au sein des trois derniers chapitres :

- Une synthèse des activités relatives à la thématique du chapitre est d'abord présentée. Cette synthèse résume les collaborations effectuées, les productions scientifiques publiés et les projets de recherche encadrés.
- Une ou deux contributions sont présentées d'une façon détaillée, de sorte à exhiber la méthodologie de recherche aussi bien dans la conception de l'approche que dans son évaluation et validation expérimentales. Les expérimentations sont souvent conduites sur des données synthétiques d'abord, de façon à contrôler l'environnement, puis sur des benchmarks pour comparer les résultats avec ceux des autres approches de la littérature et enfin sur des données réelles généralement acquises par nos propres moyens quand c'est possible, afin de faire face aux conditions d'acquisition les plus communes.
- D'autres contributions sont brièvement présentées tout en référant le lecteur aux travaux dans lesquels elles sont détaillées. Certaines d'entre elles, sont encore en cours de développement ou à un stade de simulation expérimentale.

Finalement, nous clôturons ce rapport par la présentation des objectifs atteints dans ce travail et l'évocation de nouvelles perspectives de recherches futures.

Chapitre I: **Théorie des possibilités : Positionnement,
Recours, Principes et Notions**

CHAPITRE I: THEORIE DES POSSIBILITES : POSITIONNEMENT, RECOURS, PRINCIPES ET NOTIONS	14
I. INTRODUCTION	17
II. CONCEPT DE DONNEE, INFORMATION, CONNAISSANCE, COMPETENCE :	18
1. <i>Donnée</i>	18
2. <i>Information</i>	18
3. <i>Connaissance</i>	19
4. <i>Compétence, savoir-faire, expertise</i>	20
III. CONTEXTES INCERTAINS	21
1. <i>L'incomplétude</i> :	21
2. <i>L'Ambigüité</i>	22
3. <i>Le conflit</i>	22
4. <i>L'incertitude</i>	22
5. <i>L'imprécision</i>	23
6. <i>Liens entre les imperfections</i>	24
7. <i>Positionnement épistémologique de la recherche dans le domaine de prise de décision</i>	24
IV. THEORIES DE L'INCERTAIN	24
1. <i>Théorie des probabilités</i>	25
2. <i>Théorie des fonctions de croyance (Dempster-Shafer)</i>	25
3. <i>Théorie des ensembles flous</i>	26
4. <i>Théorie des possibilités</i>	26
5. <i>Positionnement paradigmatique des théories de l'incertain</i>	26
6. <i>Démarche de la recherche : recours à la théorie des possibilités</i>	28
V. OUTILS DE LA THEORIE DES POSSIBILITES.....	29
1. <i>Définition d'une distribution de possibilité</i>	29
a) <i>Formalisation du jugement d'expert</i>	29
b) <i>Définition formelle d'une distribution de possibilités</i>	29
c) <i>Mesures de possibilité et nécessité</i>	30
d) <i>Hauteur et inconsistance</i>	32
2. <i>Normalisation d'une distribution de possibilités</i>	33
a) <i>Normalisation ordinale</i>	33
b) <i>Normalisation numérique</i>	34
c) <i>Normalisation par décalage de l'inconsistance</i>	34
3. <i>Mesures d'incertitudes</i>	35
a) <i>Mesure de spécificité S</i>	35
b) <i>Mesure d'incertitude U</i>	36
c) <i>Indice de confiance</i>	36
4. <i>Outils de fusion</i>	38
a) <i>Propagation à travers un modèle analytique</i>	38
b) <i>Combinaison par les opérateurs ensemblistes</i>	38
5. <i>Critères de décision</i>	41

a)	Mesure de possibilité (resp. nécessité) maximale.....	42
b)	Indice de confiance maximal.....	42
c)	Moyenne maximale.....	42
d)	Préférence sur le couple [II, N].....	42
e)	Mesure de similarité maximale.....	43
VI.	ETAT DE L'ART DES METHODES D'ESTIMATION D'UNE DISTRIBUTION DE POSSIBILITES.....	43
1.	<i>Les méthodes d'estimation guidées par l'expert.....</i>	43
a)	Approche directe.....	44
b)	Approche indirecte.....	44
2.	<i>Transformation d'une distribution de probabilités en une distribution de possibilités.....</i>	45
a)	Principes de cohérences entre les distributions probabilité/Possibilité.....	45
b)	Estimation de distribution de possibilités à partir de mesures précises de probabilités.....	47
c)	Estimation de distribution de possibilités à partir des intervalles de confiance probabilistes.....	48
3.	<i>Estimation d'une distribution de possibilités à partir d'une fonction d'appartenance.....</i>	50
4.	<i>Estimation de distribution de possibilités à partir d'une distribution de masse.....</i>	50
VII.	CONCLUSION.....	50

I. Introduction

Le raisonnement est un « processus cognitif » qui permet de lier, au moyen de règles logiques ou conventionnelles, deux ou plusieurs propositions pour en tirer une conclusion ou prendre une décision. Ce processus est alimenté par l'intelligence se définissant comme étant une aptitude à traiter de l'information de manière adaptative. Pour traiter l'information, l'humain opère selon le processus suivant:

- 1- d'abord il emploie ses organes sensoriels pour capter l'information
- 2- comme il est impossible de capter l'information dans son intégralité, pour diverses raisons, le cerveau humain tente de la compléter en se basant sur des *aprioris* que ce soient innés ou acquis durant la vie.
- 3- cette information ainsi reconstruite, est ensuite transformée en connaissance pour mettre à jour le système ou processus
- 4- Ce système est à appliquer à de nouvelles situations pareilles ou semblables. Il est actif et interactif de façon à permettre sa mise à jour, sa correction et son adaptation, si nécessaire, à cette nouvelle situation. C'est ce qu'on appelle l'apprentissage.

Automatiser le raisonnement nécessite d'une part sa modélisation et d'autre part la modélisation de tout concept y intervenant. C'est le défi de l'Intelligence Artificielle (IA). Le problème majeur rencontré dans les systèmes de l'IA réside dans la manière de modéliser et simuler ce traitement intelligent de l'information. En effet, un système intelligent devrait pouvoir accomplir chacune des quatre tâches. Selon [Rob 2017] deux grands paradigmes émergent dans les systèmes de raisonnement informatique : L'approche « symboliste représentationaliste » reliant le fait de penser à la manipulation de symboles et représentant la connaissance par des symboles d'une façon adéquate et organisée. Et l'approche « naturaliste » qui présume que la cognition est une émergence dans l'évolution de l'histoire naturelle du vivant et que la cognition est incarnée, située et sociale se basant sur des notions *apriori*, universelles et individuelles.

Le résultat du raisonnement dépend étroitement du degré de fidélité des modèles conçus, à la réalité. Une perception pragmatique du monde réel (objets, événements, concepts, systèmes) à modéliser, cède au fait que ses représentations, sous forme de données (mots, nombres, images, sons...), sont toujours entachées de diverses formes d'incertitudes, qu'il conviendrait de les prendre en considération lors de la modélisation. C'est pourquoi, l'intelligence artificielle, de par ses préoccupations de représentations fidèles et adéquates du monde réel, a été amenée à s'intéresser à différents cadres de traitement de l'incertitude dans l'information, allant de la théorie des probabilités, aux théories encore plus récentes d'intérêt telles que celles des possibilités, des fonctions de croyance, et des probabilités imprécises.

Nous présentons dans la première partie de ce chapitre les concepts de base reliées à l'information dans son sens usuel, aux imperfections qui l'affectent et les théories qui les traitent. Notre intérêt s'oriente vers la théorie des possibilités dans la deuxième partie de ce chapitre. Nous détaillons ses notions de base avant de présenter les approches de base permettant l'estimation des distributions des possibilités.

II. Concept de donnée, information, connaissance, compétence :

L'être humain perçoit des données. Il interprète des informations selon ses connaissances. Puis, il fait des liens avec d'autres connaissances mémorisées et les enrichit pour développer des compétences [RD 2018]. Il est courant que l'emploi de ces termes, à savoir : « donnée », « information », « connaissance », « compétence », se fasse au détriment de leurs sens originels. Il en résulte une certaine confusion. Dans ce qui suit, nous proposons une clarification.

1. Donnée

La donnée est une notion abstraite typée. Par définition, une donnée est un élément brut, qui ne porte pas de sens en lui-même tant qu'il n'a pas encore été interprété ou mis en contexte. Elle est le résultat direct d'une mesure. Elle peut être collectée par un outil de supervision, par une personne ou être déjà présente dans une base de données. La donnée peut se présenter sous diverses formes : numériques, symboliques, textuelles, logiques, linguistiques, graphiques, sensorielles, etc., pour représenter le monde réel (objets, évènements, ...).

On distingue entre données quantitatives, telles que : âge, poids, etc., et données qualitatives, telles que : nom, immatricule, etc.

Exemple de données :

- 38.5
- 10:30
- Imène

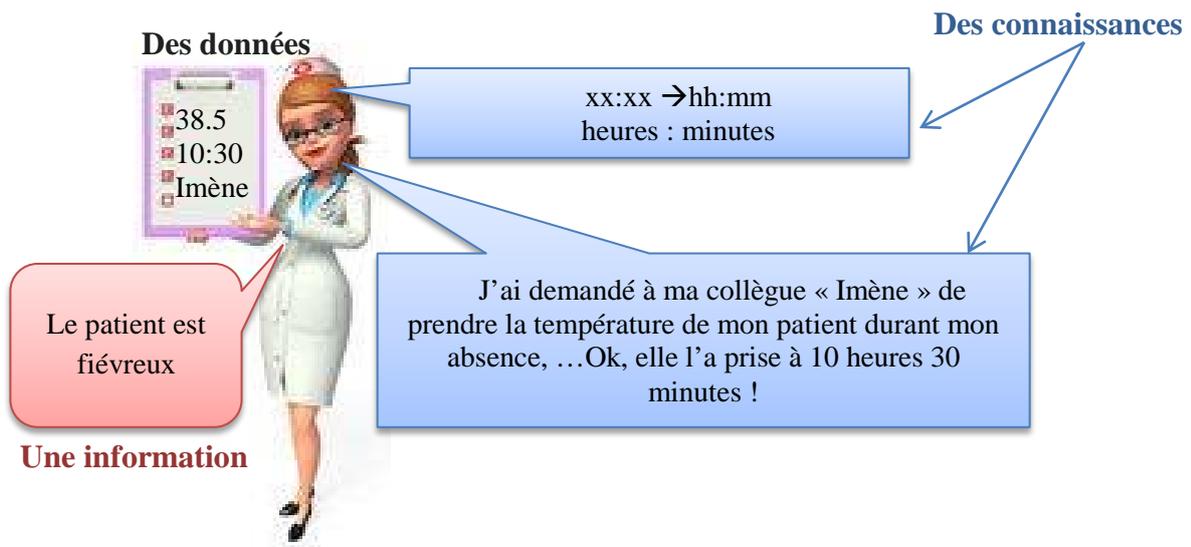
Le raisonnement basé sur des données dont le sens porté par leurs valeurs n'est pas déterminant, est simpliste. Les traitements associés à de tels raisonnements sont très abondants, mais il ne s'agit généralement pas de traitements les plus compliqués. Ils deviennent plus compliqués lorsqu'il est nécessaire de différencier les données en fonction de leur sens, quand ces données deviennent des informations.

2. Information

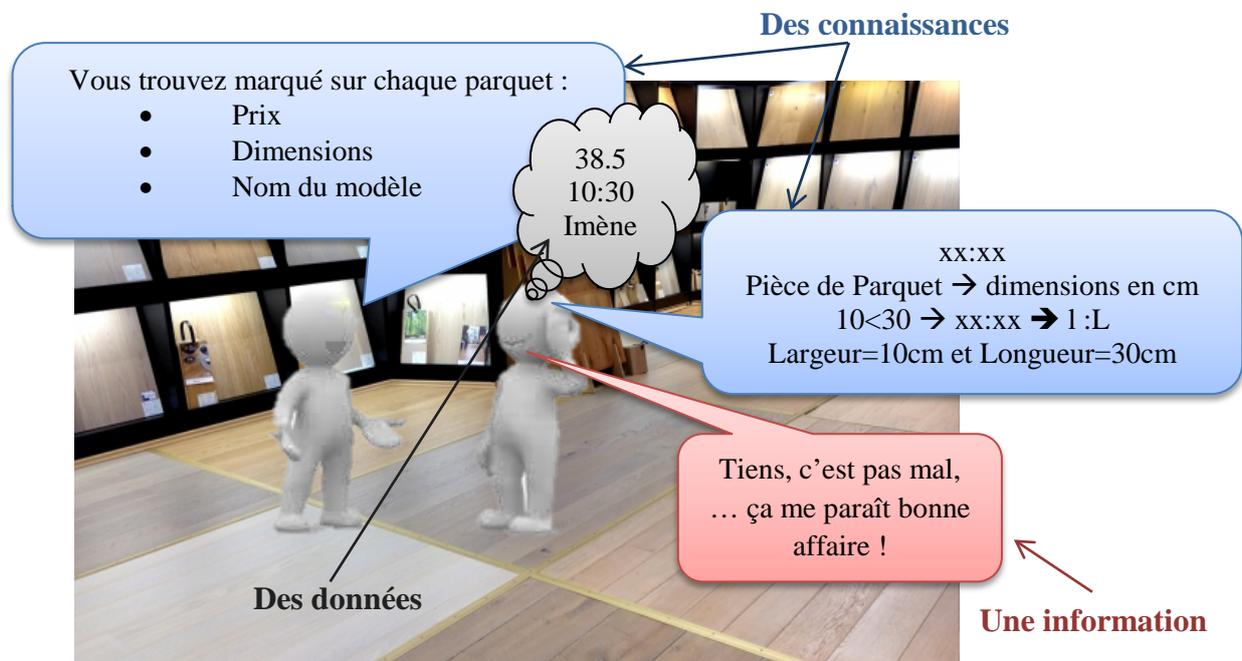
L'information est aussi une notion abstraite, mais d'un niveau d'abstraction supérieur à celui de la donnée. Une information est par définition une donnée interprétée. En d'autres termes, la mise en contexte d'une donnée, crée une valeur ajoutée (un sens) pour constituer une information. La perception de l'information se fait alors selon un modèle interprétatif lié aux connaissances de l'individu.

On dit alors que l'information est l'image du couple (donnée, modèle d'interprétation).

Si on met les données de l'exemple précédent dans un contexte donné, on retrouve l'information associée à l'ensemble (données, modèle d'interprétation). Les deux Figure 1.1-a et Figure 1.1-b illustrent l'obtention de deux informations différentes à partir des mêmes données.



(a) Contexte 1



(b) Contexte 2

Figure 1.1 Obtention d'informations distinctes à partir des mêmes données confrontées à différents contextes

Une information permet de prendre une décision (d'échelle locale ou à petite échelle) sur une action à mener.

3. Connaissance

La connaissance est aussi une notion abstraite, d'un niveau d'abstraction supérieur à celui de l'information. On peut considérer la connaissance comme une information comprise, c'est-à-dire assimilée et utilisée, qui permet d'aboutir à une action.

La connaissance est le résultat d'une réflexion sur les informations antérieures, dont a été sujet l'individu, analysées en se basant sur :

- ses expériences, ses idées, ses valeurs, les avis d'autres personnes consultées pour l'occasion
- sa propre expertise et celle de ses équivalents

La connaissance à la différence de l'information est partagée et s'appuie sur un référentiel collectif. « Nonaka et Takeuchi [Non 1991 et NT 1997], deux experts du « knowledge management », distinguent deux formes de connaissance : une connaissance tacite (ou implicite) et une connaissance explicite.

La connaissance tacite ou implicite : c'est la connaissance que possèdent les individus. Elle n'est pas formalisée et elle est difficilement transmissible. Ce sont les compétences, les expériences, l'intuition, les secrets de métiers, les tours de main qu'un individu a acquis et échangés lors d'échanges internes et externes au système. La connaissance tacite se transmet par imitation et imprégnation. On le sait sans le savoir. On met en œuvre des pratiques sans vraiment s'en rendre compte.

La connaissance explicite : C'est la connaissance formalisée et transmissible sous forme de documents réutilisables. Ce sont les informations concernant les processus, les projets, les utilisateurs, les intervenants, etc. La connaissance explicite se transmet par des documents formalisés et normalisés.

La connaissance permet aux décideurs de confronter les informations au contexte de l'évènement et à d'autres contextes externes à l'évènement afin d'avoir une meilleure connaissance et une interprétation élargie des phénomènes mis en lumière par ces informations. Elle se développe selon un processus itératif illustré par la Figure 1.2.

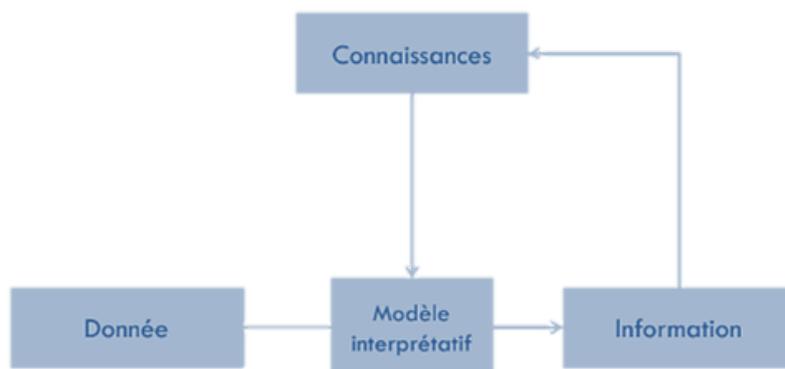


Figure 1.2 Processus itératif de développement de connaissance par confrontation des informations au contexte de l'évènement [RD 2018]

Les décisions prises à ce niveau seront d'échelle intermédiaire, autrement dit : à un niveau stratégique.

4. Compétence, savoir-faire, expertise

L'utilisation des connaissances dans le contexte d'une activité s'appelle du savoir-faire ou des compétences. La conceptualisation des savoir-faire et des compétences et leur association à l'expérience dans des situations relatives à une activité bien déterminée, forment ce que l'on appelle l'expertise.

L'expertise consiste, non pas à énoncer des vérités, mais à mesurer des écarts par rapport à des normes conventionnellement admises, constituant un cahier des charges normatif. La qualité d'un expert vient alors de sa capacité à interpréter des mesures, des observations qui

permettent d'évaluer les éventuels écarts par rapport au cahier des charges selon un référentiel valide.

Cette étape ultime de la démarche permet aux décideurs :

- ✓ d'aboutir à un état d'esprit général de discernement final sur le contenu (informations, connaissances) et de jugement de bon sens et à de sages décisions
- ✓ de s'adapter à de nouvelles situations et de lancer les actions d'adaptation de l'événement, des intervenants, des processus et des outils
- ✓ et, au final, de provoquer des changements de cap de l'événement afin d'anticiper les grands changements à venir et l'évolution d'un système

III. Contextes incertains

Dans l'optique d'un traitement efficace de l'information (dans son sens courant), la gestion des incertitudes est une nécessité : ce besoin est mis en évidence dans multiple travaux de recherche. Fondamentalement, deux sources d'incertitudes doivent être prises en compte :

La première source résulte de la variabilité naturelle des phénomènes à modéliser : c'est l'incertitude dite « aléatoire ».

La seconde source apparaît lorsqu'on est en présence d'une information « imprécise, vague (ambiguë) voire incomplète » i.e. cette incertitude, dite « épistémique », est liée à notre méconnaissance de phénomènes et apparaît dans des situations dites « d'ignorance partielle ». Dans ces situations, les données d'entrée sont trop limitées que ce soit en qualité ou en quantité. Typiquement, on est souvent confronté au cas suivant : les coûts et délais d'étude ne permettent d'obtenir qu'un nombre limité de valeurs d'un paramètre.

Autre que l'incertitude, d'autres types d'imperfections peuvent surgir. Il est important de savoir appréhender le type exact de l'imperfection à laquelle on doit faire face, à chaque niveau du traitement pour pouvoir exploiter convenablement l'information dans les niveaux suivants, de façon à contrôler sa propagation et de pouvoir en tirer profit. En effet, déceler la nature de l'imperfection rencontrée permet de faire le choix judicieux de l'outil mathématique approprié pour une représentation adéquate de l'information incertaine ou imparfaite.

Cependant, cette appréhension de l'imperfection, n'est pas aussi simple à cause, entre autre de la difficulté à comprendre ses différents aspects. Dans ce qui suit, on cite quelques aspects majeurs des imperfections dans l'information.

1. L'incomplétude :

Les incomplétudes sont des absences d'informations ou des informations partielles sur certaines caractéristiques apportées par la source. Elles peuvent se mesurer par la différence de la quantité d'information réellement fournie par la source et de la quantité d'information que la source doit fournir.

Elles peuvent être dues à l'impossibilité d'obtenir certains renseignements ou à un problème au moment de l'interception de la connaissance.

Elles peuvent aussi être associées à l'existence d'informations générales sur l'état d'un système, habituellement vraies, soumises à des exceptions que l'on ne peut pas énumérer ou prévoir, selon les cas.

Exemple :

- « Généralement, Ahmed est à son bureau tous les jours », sauf s'il est malade ou si un événement grave survient dans sa famille.

2. L'Ambigüité

Une information est dite ambiguë si plusieurs contenus informationnels peuvent se produire simultanément ou encore si la réalisation partielle de singletons, est possible. Ce type d'information possède, de ce fait, plus qu'une seule interprétation.

Exemple:

- « J'embobine mes fils. » → le groupe « Mes fils » peut se référer soit à « mes enfants » (embobiner signifie dans ce cas, dans son sens figuré, tromper par des paroles séduisantes), soit à mes bobines de fil.

3. Le conflit

Le conflit caractérise deux ou plusieurs informations conduisant à des interprétations contradictoires, conflictuelles et donc incompatibles. Les situations de conflit sont fréquentes dans les problèmes de fusion.

Exemples :

- Les phrases « Tous les hommes sont barbus » et « Quelques hommes ne sont pas barbus » sont contradictoires.
- Les données binaires 011101 et 100010 sont contradictoires.

4. L'incertitude

La survenance d'un événement dans un contexte donné pose parfois une question à laquelle on ne peut pas répondre. L'incertitude présente la différence résultante entre la réponse disponible et la réponse réelle espérée. Elle fait, de ce fait, référence à la qualité de l'information retenue de cette réponse.

Une information est donc dite incertaine, si on ne peut pas, dans un contexte donné, fournir l'information réellement souhaitée.

L'incertitude peut être due à [DP 2008]:

- ✓ La variabilité d'un phénomène qui empêche de prévoir le résultat de la prochaine expérience.

Exemple:

- Quantité de pluie journalière à Tunis : il s'agit d'une réponse probabiliste en fonction de la fréquence observée.
- ✓ Le manque d'information ou incomplétude.

Exemple:

- Date de naissance d'une célébrité; malgré que ça ne soit pas une grandeur aléatoire, les réponses seront plus ou moins précises en fonction de la connaissance acquise par les individus.
- ✓ La présence d'informations contradictoires ou de conflit.

Exemple :

- Identification d'un véhicule impliqué dans un accident : plusieurs témoignages contradictoires existent sans avoir recours à celui réellement impliqué.

Comme il a été indiqué plus haut, on distingue deux catégories d'incertitudes selon leurs sources:

- ✓ Incertitudes aléatoires qui font référence aux incertitudes provenant de la variabilité de phénomènes répétables ou de caractéristiques au sein de populations.
- ✓ Incertitudes épistémiques qui résultent de notre manque de connaissance des phénomènes.

5. L'imprécision

L'imprécision prend la forme d'une disjonction de valeurs mutuellement exclusives. Elle fait donc référence au contenu de l'information. Celle-ci est dite imprécise si elle est insuffisante pour permettre de comprendre une situation donnée.

Exemple :

- L'âge de Sarah est entre 20 et 25 ans.

Les imprécisions sont dues à une difficulté dans l'énoncé de l'information, soit parce que des connaissances numériques sont mal connues, soit parce que des termes du langage naturel sont utilisés pour qualifier, de façon vague une caractéristique du système. Le premier cas est la conséquence d'une insuffisance des instruments d'observation, d'erreurs de mesure ou encore de connaissances flexibles.

Exemples :

- « 2000 à 3000 manifestants » : insuffisance d'instruments d'observation
- « Poids à 1% près » : erreur de mesure
- « La taille d'un adulte est environ entre 1.50 et 2 mètres » : connaissances flexibles

Le second cas provient de l'expression spontanée de connaissances ou de l'utilisation de catégories aux limites mal définies.

Exemple :

- température douce, grand appartement, proche de la plage : expression spontanée de connaissances
- Enfant, adulte, vieillard : catégories aux limites mal définies

6. Liens entre les imperfections

Les différentes formes d'imperfection précédemment citées, ne sont pas indépendantes, en effet :

- ✓ L'imprécision et l'incertitude sont deux notions qui sont fortement liées : Quand l'imprécision de l'information augmente, l'incertitude diminue et vice versa. Ainsi, dans un jeu de pile-ou-face, l'événement imprécis {pile, face, tranche} est un événement certain tandis que l'événement {pile, tranche} qui est plus précis est moins certain.
- ✓ Les incomplétudes entraînent des incertitudes.
- ✓ Les imprécisions peuvent être associées à des incomplétudes et elles engendrent des incertitudes au cours de leur manipulation. Par ailleurs, plus on exige de précision dans un énoncé quelconque, moins il est certain. En d'autres termes, quand l'imprécision de l'information augmente, l'incertitude diminue.
- ✓ L'ambiguïté peut provenir d'une autre imperfection de l'information (incertitude, imprécision, conflit,...), mais pas nécessairement. Par exemple, l'imprécision d'une mesure qui ne permet pas de différencier deux situations.
- ✓ Les conflits peuvent facilement être confondus avec d'autres types d'imperfection ce qui prouve la difficulté de leurs détections.

7. Positionnement épistémologique de la recherche dans le domaine de prise de décision

Les données et/ou informations recueillies des phénomènes ou événements réels, ne sont généralement pas épargnés d'imperfections. Dans les systèmes d'automatisation du raisonnement sur ces événements ou phénomènes, les problèmes sont souvent flous, mal structurés, les entrées sont imparfaites et les objectifs sont couramment obscurs. Face à cette situation « inconfortable », les outils « classiques » des théories classiques et méthodologies statistiques apparaissent trop contraignants pour la modélisation de tels systèmes.

Les systèmes de raisonnement numérique pour l'aide à la décision, touchent à des domaines alliés, d'une façon ou d'une autre au vivant, ce qui rend les problèmes abordés assez complexes, de façon à ne pouvoir être explicitement formulés sans être sur-simplifiés. Ceci préconise de rechercher d'autres outils mathématiques plus flexibles, pouvant s'adapter aux données effectivement disponibles. L'idée serait de favoriser des conceptions souples permettant d'avoir plusieurs versions possibles du système à modéliser. C'est ce qui explique qu'un certain nombre de théories différentes et interdépendantes sont apparues durant ces dernières décennies. Leurs tentatives de théorisation du raisonnement humain, entre autre pour des prises de décision, ont fourni différents modèles offrant différentes vues du même système. Il s'agit des théories de l'incertain.

IV. Théories de l'incertain

Un formalisme simpliste du problème de prise de décision, associé à un événement ou phénomène donné, consiste à trouver, dans l'univers des « décisions » envisageables, une étiquette « décision », à accorder à l'entrée représentée par un ensemble de données/informations. Quand cette entrée est représentée par un ensemble de

données/informations incertaines, le formalisme consisterait alors à un partage de la certitude sur l'ensemble des étiquettes « décisions » envisageables, dites aussi « classes thématiques ».

Les théories de l'incertain, telles que celle des probabilités, des évidences ou encore celle des ensembles flous, ont différentes visions pour le partage de cette certitude, sur l'ensemble des classes thématiques de l'univers des décisions. En effet, l'approche probabiliste consiste à partager une masse totale unitaire de certitude sur l'ensemble des singletons qui sont les étiquettes des classes. La théorie des fonctions de croyance partage cette masse totale sur l'ensemble des parties de l'univers de discours autorisant ainsi de « mixer » l'incertitude probabiliste (aléatoire) avec l'imprécision. Finalement les théories des ensembles flous et des possibilités considèrent que chaque singleton (entrée) va se voir attribué à une classe unique, indépendamment des autres, avec une valeur de croyance entre 0 et 1.

Dans ce qui suit, on propose de brièvement présenter ces théories en insistant sur le type d'imperfection pris en compte par chacune d'elles.

1. Théorie des probabilités

La théorie des probabilités est considérée comme la théorie la plus classique et la plus exploitée pour traiter les informations affectées par l'incertitude. Quand il s'agit de l'incertitude aléatoire, les méthodes probabilistes permettent de la représenter rigoureusement tant que la quantité et la qualité des données statistiques en entrée sont satisfaisantes. Cependant, face à l'incertitude épistémique et l'imprécision, les outils probabilistes apparaissent inadéquats. A titre d'exemple, lorsque la donnée d'entrée ne permet que de juger, que le paramètre incertain A , se trouve entre les bornes $a-\delta$ et $a+\delta$, alors est-il justifié de supposer a priori la forme de la loi de probabilité ? A cette question [FG 1996], répondent que l'on risque soit de sous-estimer soit de sur estimer le résultat final. Les probabilités se trouvent en difficulté face à la représentation de l'imprécision et elles conduisent souvent à confondre les notions de l'incertitude épistémique et de l'imprécision [Blo 2003]. Elles ne sont pas non plus, plus adaptées à la représentation de l'ambiguïté [Van 2003]. Face à ces limites, la théorie des probabilités est devenue l'objet de multiples critiques. Plusieurs spécialistes du domaine du traitement de l'information incertaine préconisent de recourir à d'autres outils mathématiques plus flexibles pour traiter la donnée effectivement disponible, entachée d'imperfections non probabilistes, notamment l'incertitude épistémique et l'ambiguïté.

2. Théorie des fonctions de croyance (Dempster-Shafer)

Appelée aussi théorie de l'évidence [Sha 1976, TY 2011], elle permet de modéliser l'incertitude et l'imprécision à l'aide de trois fonctions: une fonction de masses de croyance m , et deux fonctions ensemblistes ; la fonction de plausibilité Pl et la fonction de crédibilité Cr [GG 1984]. L'avantage principal de cette théorie réside dans le fait qu'elle permette la modélisation de l'information d'une manière souple et flexible car elle modélise le doute ou « l'hésitation » entre plusieurs singletons [GG 1984]. Notons que cette modélisation est réalisée d'une façon indépendante de la connaissance de la distribution de probabilité [TY 2011], mais elle est très dépendante de la fonction de masse employée.

3. Théorie des ensembles flous

La théorie des ensembles flous [Zad 1965, KG 1991, Kau1973 et BD 1995], représente un cadre bien adapté pour la modélisation des informations ambiguës. Rappelons qu'une information est dite ambiguë si plusieurs contenus informationnels peuvent se produire simultanément ou la réalisation partielle de singletons est possible. Notons que dans les deux cas, cette imperfection est totalement différente de l'incertitude. En effet, la réalisation simultanée de plusieurs singletons sort du cadre de base de la définition de l'incertitude où un seul singleton peut se produire. C'est aussi le cas concernant la réalisation partielle d'un singleton car en considérant l'incertitude, un singleton ne peut qu'avoir une réalisation binaire (i.e. il se produit, ou, il ne se produit pas), donc une réalisation partielle sort aussi du cadre de définition de l'incertitude. A ce titre, la théorie des ensembles flous correspond « mieux » à la modélisation et à la description des concepts et des informations manipulés par l'humain qui sont : ambigus, approximatifs, imprécis ou vagues.

Afin de permettre la prise en compte combinée à la fois de l'imprécision et de l'incertitude dans des connaissances, Zadeh a introduit la théorie des possibilités en 1978 [Zad 1978].

4. Théorie des possibilités

La théorie des possibilités [Zad 1978, DP 1988a-c] est une théorie relativement nouvelle consacrée à la gestion de l'incertitude épistémique, c'est-à-dire l'incertitude dans le contexte où la connaissance disponible n'est exprimée que sous une forme ambiguë. Cette théorie a été introduite pour la première fois par Zadeh en 1978 comme une extension des ensembles flous et de la théorie de la logique floue pour exprimer le flou intrinsèque des langages naturels ainsi que des informations incertaines [Zad 1979, Zad 1989]. Dans le cas où les connaissances disponibles sont ambiguës et exprimées sous forme d'une fonction d'appartenance à un ensemble flou défini sur l'ensemble des décisions, la théorie de la possibilité transforme chaque valeur d'appartenance en un intervalle possibiliste de mesures de possibilités et de mesures de nécessités [KS 2017, KAA 2018].

5. Positionnement paradigmatique des théories de l'incertain

La théorie des probabilités imprécises est un cadre général des autres théories de l'incertain étudiées. La cartographie présentée par la Figure 1.3 illustre la relation entre ces modèles mathématiques.

Le choix de la théorie à utiliser pour modéliser les informations disponibles dépend de la nature de l'imperfection qui les entache, ce qui favorise la sélection d'une théorie particulière au profit des autres. Dans le Tableau 1.1, on présente une synthèse pour guider le choix de la théorie à utiliser pour modéliser une information incertaine selon la situation qui se présente.

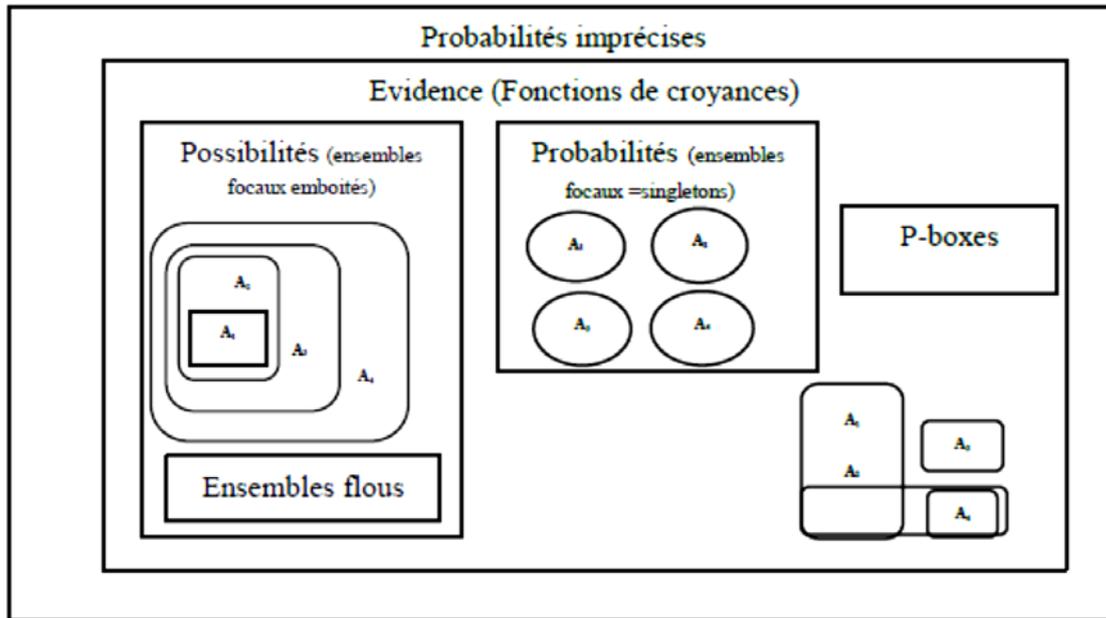


Figure 1.3 Paradigme relationnel des théories de l'incertain

Tableau 1.1 Guide du choix de la théorie de l'incertain

<i>Théories de l'incertain</i>	<i>Connaissances</i>	<i>Imperfections traitées</i>
Probabilités	<ul style="list-style-type: none"> ⇒ Taille importante de l'échantillon est requise ⇒ Connaissance théorique (relative à la loi de distribution de données) 	<ul style="list-style-type: none"> ⇒ Incertitude aléatoire ⇒ Imprécision (de façon implicite)
P-boxes	<ul style="list-style-type: none"> ⇒ Taille de l'échantillon non contraignante ⇒ Modèle paramétrique imprécis en connaissant la loi de distribution des données 	<ul style="list-style-type: none"> ⇒ Incertitude, ⇒ Imprécision, ⇒ Incomplétude
Evidence	<ul style="list-style-type: none"> ⇒ Informations de type quantiles (écart type, médiane, etc) 	<ul style="list-style-type: none"> ⇒ Incertitude ⇒ Imprécision ⇒ Incomplétude ⇒ Conflit ⇒ Ambiguïté (ignorance totale)
Possibilités	<ul style="list-style-type: none"> ⇒ Grande tolérance sur une taille réduite de l'échantillon ⇒ Grandeur fixe mais imprécise à base d'avis d'expert (intervalles de confiance) 	<ul style="list-style-type: none"> ⇒ Incertitude ⇒ Imprécision ⇒ Ambiguïté

6. Démarche de la recherche : recours à la théorie des possibilités

La théorie des probabilités s'avère souvent utile pour représenter des densités de probabilités particulières à condition d'avoir une connaissance *a priori* concernant la loi suivie par la population (à titre d'exemple : La loi uniforme est utilisée pour représenter l'équipartition). Par ailleurs, elle propose aussi des méthodes d'estimation de loi de probabilités, mais pour les performer correctement un échantillon important de données est nécessaire. Ce qui n'est pas toujours disponible en pratique. Elle est en plus, mal adaptée pour représenter les imperfections du type imprécision et incomplétude.

En ce qui concerne la théorie de Dempster-Shafer, elle présente un bon cadre pour représenter l'incertitude, l'imprécision et l'incomplétude de manière efficace. En effet, elle permet de représenter directement l'information de type quantiles ; mais ceci n'est réalisable que lorsqu'on dispose de probabilités d'évènements connus. L'inconvénient majeur de l'application de la théorie des fonctions de croyance réside dans la complexité de construction des fonctions de masse même si on dispose d'une information de type intervalle de confiance qui est beaucoup moins précise que celle du type quantile.

L'incomplétude des données, ou leurs limitations en qualité ou en quantité, empêche de définir de manière univoque un modèle classique (probabiliste ou ensembliste). L'idée serait alors d'utiliser l'information subjective d'expert (jugement) afin de pallier à ce manque de données. Il s'agit de hiérarchiser de manière objective les valeurs d'un paramètre au sein d'un intervalle [Roh 2007]. Il n'est pas rare d'être confronté à un jugement d'expert sur le modèle suivant : « Je suis certain de trouver la quantité incertaine dans les bornes [a, b], mais les valeurs au sein de l'intervalle [c, d] semblent être les plus vraisemblables », car l'expert est souvent bien plus à l'aise « d'exprimer ce qui est possible que ce qui est probable » [BN 1992]. Une telle information peut être rigoureusement représentée dans le cadre de la théorie des possibilités. La Figure 1.4 illustre une représentation du jugement « je suis certain que le paramètre incertain se trouve dans les bornes [1.1,1.8], mais les valeurs au sein de l'intervalle [1.5,1.6] me semblent les plus vraisemblables » fait par l'expert.

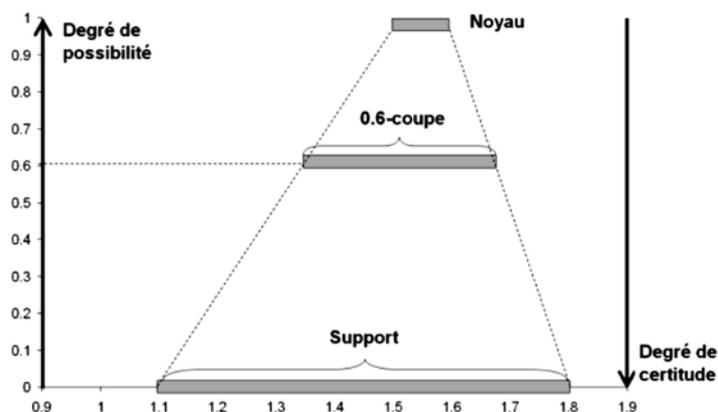


Figure 1.4 « Illustration d'une distribution de possibilité vue comme un emboîtement d'intervalles de confiance sur l'exemple suivant : « je suis certain que le paramètre incertain se trouve entre 1.1 et 1.8, mais les valeurs entre 1.5 et 1.6 me semblent les plus vraisemblables ». », (figure prise de [Roh 2007])

On se propose dans la suite d'expliciter les notions fondamentales liées à cette théorie.

V. Outils de la théorie des possibilités

La théorie des possibilités a été définie à partir de la théorie des sous-ensembles flous. Si la théorie des ensembles flous assouplit le cadre de la théorie classique des ensembles pour pouvoir traiter l'imprécision, la théorie des possibilités propose un cadre permettant de représenter à la fois l'imprécision et l'incertitude, et ce par l'intermédiaire de distributions de possibilités.

1. Définition d'une distribution de possibilité

Avant de donner la définition formelle d'une distribution de possibilités, on propose d'abord de poursuivre le raisonnement sur la façon par laquelle l'expert formule son jugement et comment il aboutit à ce modèle possibiliste.

a) Formalisation du jugement d'expert

Une représentation mathématique de l'information d'expert : « Je suis certain de trouver la quantité dans les bornes $[a, b]$, mais les valeurs au sein de l'intervalle $[c, d]$ semblent être les plus vraisemblables », peut se formuler à partir des deux intervalles $[a, b]$ et $[c, d]$. En effet, $[a, b]$ étant l'intervalle dans lequel l'expert est certain de trouver la « valeur » du paramètre, constitue le « support » du modèle. L'expert peut apporter de l'information supplémentaire en exprimant sa préférence au sein du support, par l'introduction de la notion de degré de possibilité α [Roh 2007]. Selon [Sha 1961], un degré de possibilité peut être vu comme la mesure de la surprise de l'expert s'il apprenait que son jugement était faux. Ainsi, sa surprise serait totale s'il apprenait que la « valeur » du paramètre n'est pas du tout dans l'intervalle $[c, d]$ et c'est ce qui lui accorde un degré de possibilité $\alpha=1$. Ayant le degré maximal de possibilité, cet intervalle est dénommé « noyau ».

Par défaut d'information supplémentaire, les deux intervalles, « support » et « noyau », sont reliés par une fonction affine définissant ainsi un emboîtement d'intervalles hiérarchisés par le degré de possibilité : c'est la distribution de possibilité trapézoïdale associée au paramètre imprécis. La Figure 1.4 en illustre un exemple.

b) Définition formelle d'une distribution de possibilités

Considérons un univers exclusif et exhaustif de discours Ω formé par K alternatives ou hypothèses $x_k, k = 1, \dots, K : \Omega = \{x_1, x_2, \dots, x_K\}$.

- ✓ *L'exclusivité* signifie qu'une seule et même alternative peut se produire en même temps, alors que
- ✓ *L'exhaustivité* fait référence au fait que l'alternative apparaissant, appartienne certainement à Ω .

Dans ce contexte, une distribution de possibilités, notée π , est une fonction définie de Ω dans l'intervalle borné $[0,1]$, qui attribue à chaque alternative x_k une valeur dans $[0,1]$.

$$\begin{aligned} \pi: \Omega &\rightarrow [0,1] \\ x_k &\mapsto \pi(x_k) \end{aligned} \qquad \text{Eq1.1}$$

Cette valeur, $\pi(x_k)$, exprime l'état de connaissance, ou croyance de l'expert, sur le monde réel et représente le degré de possibilité pour que x_k soit l'"unique" occurrence.

De cette définition, on peut donner une interprétation d'une distribution de possibilité unimodale comme un ensemble d'intervalles de confiance emboîtés : des α -coupes de la forme $[x^-; x^+]_\alpha = \{x, \pi(x) \geq \alpha\}$, avec α le degré de possibilité.

De la notion de degré de possibilité découle celle de degré de certitude que « l'intervalle $[x^-; x^+]_\alpha$ contienne x » ayant une valeur égale à $1 - \alpha$. La Figure 1.4 illustre cette interprétation.

Formellement, une telle distribution fournit une double mesure de la vraisemblance d'une occurrence, ou un sous-ensemble d'occurrences (ou événement), à savoir les mesures de possibilité et de nécessité.

c) Mesures de possibilité et nécessité

Il s'agit de deux mesures ensemblistes complémentaires : la mesure de possibilité Π et la mesure de nécessité N . L'utilisation de ces deux mesures dans la théorie des possibilités fait la principale différence par rapport à la théorie des probabilités.

Pour chaque sous-ensemble (ou événement) A , ces deux mesures sont définies comme suit:

$$\Pi(A) = \max_{x_k \in A} (\pi(x_k)) \quad \text{Eq1.2}$$

$$N(A) = 1 - \Pi(A^c) = \min_{x_k \notin A} (1 - \pi(x_k)) \quad \text{Eq1.3}$$

Où A^c désigne le complément de l'événement A .

- ✓ La mesure de possibilité $\Pi(A)$ estime le niveau de cohérence pour l'occurrence de l'événement A , compte tenu des connaissances disponibles encodées par la distribution de possibilités π .

$\Pi(A) = 0$ signifie que l'événement A est un événement impossible, et,

$\Pi(A) = 1$ signifie que l'événement A est totalement possible.

- ✓ La mesure de nécessité $N(A)$ évalue le niveau de certitude pour l'occurrence de l'événement A , impliqué par la distribution de possibilité π .

$N(A) = 0$ signifie que la nécessité de l'occurrence de A est nulle, et

$N(A) = 1$ signifie que l'occurrence de A est totalement nécessaire.

Tenant compte de ces mesures, trois cas extrêmes peuvent être définis :

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

i) Propriétés liées aux mesures de possibilité et de nécessité

Soient A et B deux événements inclus dans Ω , on a :

- ✓ $\Pi(\Phi) = 0$ et $\Pi(\Omega) = 1$;
- ✓ $\forall A, B \subseteq \Omega : \Pi(A \cup B) = \max(\Pi(A), \Pi(B))$;
- ✓ $\forall A, B \subseteq \Omega : \Pi(A \cap B) \leq \min(\Pi(A), \Pi(B))$.
- ✓ $N(\Phi) = 0$ et $N(\Omega) = 1$;
- ✓ $\forall A, B \subseteq \Omega, N(A \cap B) = \min(N(A), N(B))$;
- ✓ $\forall A, B \subseteq \Omega, N(A \cup B) \geq \max(N(A), N(B))$.

Les propriétés suivantes caractérisent les liens de dualité entre les deux mesures d'incertitude Π et N :

- ✓ $\Pi(A) + \Pi(A^C) \geq 1$;
- ✓ $\min\{N(A), N(A^C)\} = 0$;
- ✓ $\max\{\Pi(A), \Pi(A^C)\} = 1$;
- ✓ $\Pi(A) < 1 \Rightarrow N(A) = 0$;
- ✓ $N(A) > 0 \Rightarrow \Pi(A) = 1$.
- ✓ Pour les distributions de possibilité normalisées, $\forall A \subseteq \Omega, N(A) \leq \Pi(A)$

Afin d'élargir cette dernière propriété pour les cas des distributions anormales, plusieurs approches ont été proposées :

ii) Proposition de Yager

Afin de remédier à l'inconvénient ($N(A) > \Pi(A)$), *Yager* [Yag 1986] a proposé de substituer la mesure de nécessité par une nouvelle quantité appelée la *certitude*, *Cert*, qui est définie par :

$$\begin{aligned} \text{Cert}: 2^\Omega &\rightarrow [0,1] \\ A &\mapsto \text{Cert}(A) = \min\{\Pi(A), N(A)\} \end{aligned} \quad \text{Eq1.4}$$

Ainsi, l'occurrence d'un évènement $A \subseteq \Omega$ est caractérisée par l'intervalle possibiliste $[\text{Cert}(A), \Pi(A)]$.

iii) Proposition de Dubois

Comme solution alternative à la proposition de *Yager*, *Dubois* [DP 1987] a proposé de modifier la définition-même de la mesure de nécessité par :

$$\begin{aligned} N_D: 2^\Omega &\rightarrow [0,1] \\ A &\mapsto N_D(A) = \Pi(\Omega) - \Pi(A^c) \end{aligned} \quad \text{Eq1.5}$$

où : $\Pi(\Omega) = h(\pi)$.

Dans cette définition, la valeur de possibilité la plus élevée représente la possibilité de référence par rapport à laquelle le degré de certitude est estimé. Si la distribution de possibilité est normale, alors, la mesure de nécessité de *Dubois* est identique à la mesure de nécessité « classique ».

Ainsi, l'occurrence d'un évènement $A \subseteq \Omega$ est caractérisée par le couple possibiliste $[N_D(A), \Pi_D(A) = \Pi(A)]$.

iv) Proposition de Klir

La propriété de base de la théorie des possibilités « $\Pi(\Omega) = N(\Omega) = 1$ » est non vérifiée avec la proposition de *Dubois* et par conséquent, certaines propriétés associées sont modifiées, comme par exemple :

- ✓ $\Pi(A) + \Pi(A^C) \geq 1$,
- ✓ $\min\{N(A), N(A^C)\} = 0$,

- ✓ $\max \{\Pi(A), \Pi(A^c)\} = 1,$
- ✓ $\Pi(A) < 1 \Rightarrow N(A) = 0,$ et
- ✓ $N(A) > 0 \Rightarrow \Pi(A) = 1.$

Pour cette raison, *Klir* [Kli 1999] a considéré que la proposition de *Dubois* n'est pas adéquate, et comme solution alternative, il a présenté une autre proposition dans laquelle les définitions des mesures de possibilité et de nécessité ont été remplacées par les définitions suivantes :

$$\Pi(A) = \begin{cases} \max_{x \in \Omega} \pi(x) & \text{Si } A \neq \Omega \\ 1 & \text{Si } A = \Omega \end{cases}$$

$$N(A) = \begin{cases} 1 - \Pi(A^c) & \text{Si } \Pi(A) = 1 \\ 0 & \text{Si } \Pi(A) < 1 \end{cases}$$

Une comparaison entre les trois propositions (*Yager*, *Dubois*, *Klir*), détaillée dans [AIS 2012] est résumée dans le Tableau 1.2.

Tableau 1.2 Comparaison entre les propositions de *Yager*, *Dubois* et *Klir* (Tab.2.3 de [AIS 2012])

	Proposition de <i>Yager</i>	Proposition de <i>Dubois</i>	Proposition de <i>Klir</i>
Avantages	- Cert $\leq \Pi$ est toujours respectée.	- $N \leq \Pi$ est respectée ; - Toutes les propriétés de dualité entre la N et Π sont toujours respectées.	- $N \leq \Pi$ est toujours respectée ; - Toutes les propriétés de dualité entre la N et Π sont toujours respectées.
Inconvénients	- Cert n'a pas de sens physique ; - Les autres propriétés ne sont pas toujours respectées.	- $N \in [0, h(\pi) < 1]$; - $\Pi \in [0, h(\pi) < 1]$.	- La valeur de N est toujours égale à zéro, donc elle est perdue comme information utile.

d) Hauteur et inconsistance

La hauteur d'une distribution de possibilité détermine le singleton ayant le degré de compatibilité le plus élevé avec les connaissances disponibles, c'est-à-dire, le singleton qui a la plus grande possibilité à se produire comme une solution compatible avec les connaissances disponibles :

$$h(\pi) = \max_{x \in \Omega} \pi(x) \tag{Eq1.6}$$

La hauteur maximale de la distribution de possibilités π doit être inférieure à 1 pour que la distribution soit appelée inconsistante d'où vient l'appellation du degré d'inconsistance $Inc(\pi) \in [0, 1]$ d'une distribution de possibilités. Cette nouvelle valeur est définie par :

$$Inc(\pi) = 1 - h(\pi) \tag{Eq1.7}$$

La hauteur et l'inconsistance sont deux connaissances qu'on peut déduire d'une distribution de possibilités comme le montre la Figure 1.5.

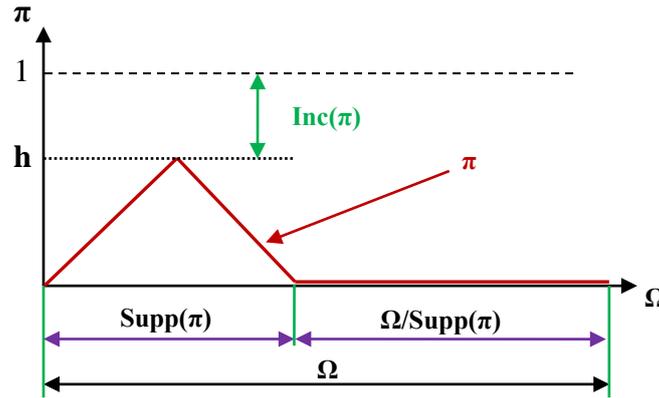


Figure 1.5 Les connaissances véhiculées par une distribution de possibilité (Fig.2.6 de [ALS 2012])

2. Normalisation d'une distribution de possibilités

Dans la modélisation possibiliste, de nombreux travaux de recherche considèrent que les distributions de possibilités sont normalisées. Néanmoins, une distribution non normalisée est fréquemment rencontrée dans la pratique. Par exemple, lorsqu'une source d'information est incapable de caractériser la certitude concernant l'occurrence de tous les singletons, elle génère une distribution non normalisée. Cette source est appelée non consistante. Son inconsistance découle de plusieurs raisons dont nous citons :

- ✓ La non exhaustivité de l'univers Ω .
- ✓ La fusion conjonctive de deux distributions de possibilités non concordantes engendre une distribution de possibilités non normalisée π . En effet, deux distributions de possibilités π_1 et π_2 sur l'univers Ω dites non concordantes lorsque aucune hypothèse singleton n'est considérée comme étant totalement possible simultanément par ces deux distributions de possibilités. La Figure 1.7 démontre ce principe de génération de π utilisant l'opérateur de fusion conjonctive "min". Ce cas est considéré comme un conflit de consistance entre les deux distributions π_1 et π_2 .

Considérons l'univers Ω comportant K singletons $\Omega = \{x_1, x_2, \dots, x_K\}$ présentés par une distribution de possibilités anormale π . Dans la suite, nous décrivons trois méthodes de normalisation de distribution de possibilité très répandues dans la littérature : normalisation ordinale, normalisation numérique, et normalisation par décalage de l'inconsistance.

a) Normalisation ordinale

La normalisation ordinale consiste à décaler, à l'unité, uniquement le degré de possibilité correspondant aux singletons de l'univers Ω dont sa valeur est maximale (égale à la hauteur $h(\pi)$). Les singletons, dont le degré de possibilité est inférieur à $h(\pi)$, gardent leurs valeurs initiales :

$$\pi'(x) = \begin{cases} 1 & \text{Si } \pi(x) = h(\pi) \\ \pi(x) & \text{Sinon} \end{cases} \quad \text{Eq1.8}$$

b) Normalisation numérique

La normalisation de distribution de possibilités utilisant cette méthode repose sur la division de toute la distribution de possibilités par sa valeur maximale $h(\pi)$:

$$\pi'(x) = \frac{\pi(x)}{h(\pi)} = \frac{\pi(x)}{\max_{x \in \Omega} \pi(x)} \quad \text{Eq1.9}$$

c) Normalisation par décalage de l'inconsistance

Le principe de cette méthode, comme son nom l'indique, est de décaler le degré de possibilité de tous les singletons de Ω en leur ajoutant le degré d'inconsistance.

$$\pi'(x) = \pi(x) + \text{Inc}(\pi) \quad \forall x \in \Omega \quad \text{Eq1.10}$$

En appliquant cette méthode de normalisation toutes les valeurs de possibilité sont affectées et non seulement les singletons correspondants à la possibilité maximale.

La Figure 1.7 illustre le comportement de ces méthodes de normalisation.

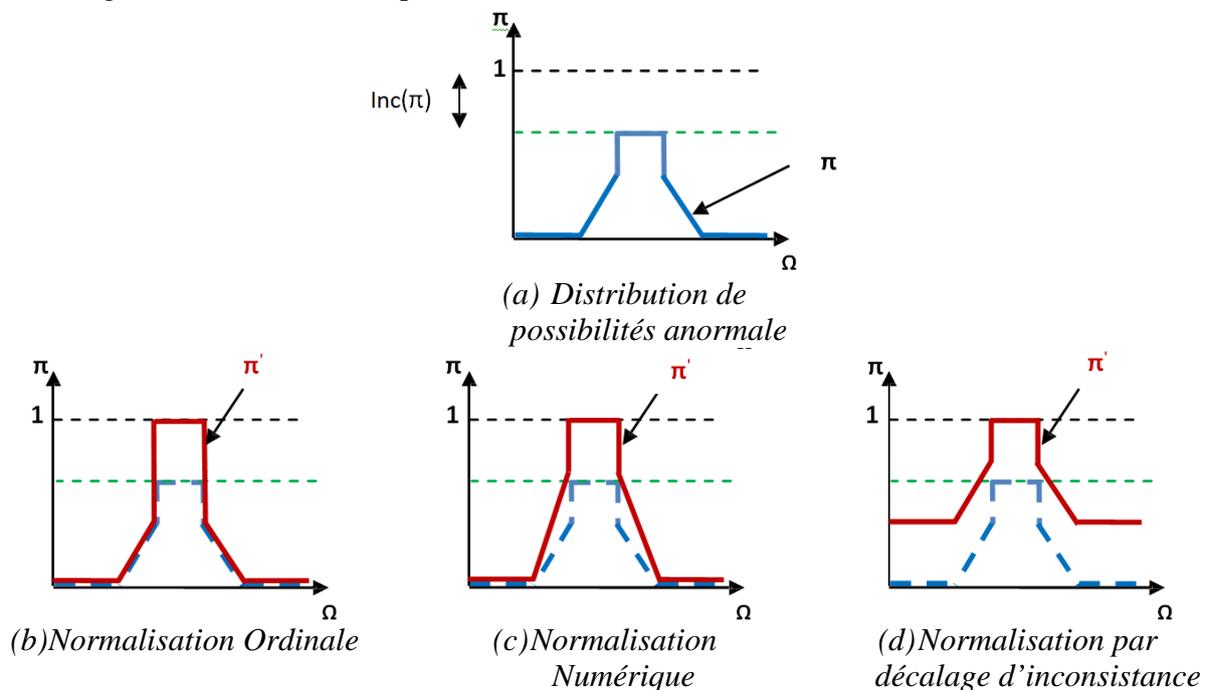


Figure 1.6 Méthodes de normalisation d'une distribution de possibilité anormale (Fig.2.7 de [AIS 2012])

Tableau 1.3 Comparaison entre les méthodes de normalisation (Tab.2.2 de [AIS 2012])

Méthode de normalisation	Domaine d'impossibilité	Ordre des possibilités	Inconsistance
Ordinale	Conservé	Modifié	Modifiée
Numérique	Conservé	Conservé	Modifiée
Décalage de l'inconsistance	Modifié	Conservé	Modifiée

En observant la Figure 1.6 et le Tableau 1.3, nous pouvons remarquer que le processus de normalisation d'une distribution de possibilité entraîne une déformation et perte

d'informations portées par la distribution, et ceci quel que soit la méthode de normalisation appliquée [AIS 2012].

3. Mesures d'incertitudes

Une mesure d'incertitude permet d'évaluer la quantité d'information contenue dans une distribution de possibilité. Plusieurs mesures d'incertitude ont été proposées dans la littérature. [AIS 2012] a étudié le comportement d'une mesure d'incertitude, à la lumière de la notion de normalité d'une distribution de possibilité. Dans ce qui suit, nous présentons brièvement les trois mesures suivantes : la *Spécificité* (Sp) [Yag 1991], la mesure d'incertitude (U) [HK 1982] et l'*indice de confiance* (Ind) [KP 2004], ainsi que l'essentiel des déductions faites par [AIS 2012] sur l'influence de la normalisation sur ces trois mesures.

a) Mesure de spécificité S

Soit Ω l'univers des discours, et $\pi(\Omega)$ l'ensemble de toutes les distributions de possibilités définies sur Ω . La mesure de spécificité, Sp, est une fonction $Sp : \pi(\Omega) \rightarrow [0,1]$, vérifiant les propriétés suivantes :

- 1) $Sp(\pi) = 1$ si et seulement si $\pi(x_0) = 1$, et $\pi(x) = 0$ pour tout $x \in \Omega$ ($x \neq x_0$) ;
- 2) $Sp(\pi_\Phi) = 0$; $\pi_\Phi(x) = 0$ pour tout $x \in \Omega$;
- 3) Si $\pi_1 \leq \pi_2$ alors $Sp(\pi_1) \leq Sp(\pi_2)$.

A la lumière de cette définition, plusieurs expressions de la spécificité ont été proposées dans la littérature. L'une des plus utilisée [GYT 2003], est donnée pour une distribution de possibilité $\pi : \Omega \rightarrow [0,1]$, normale et ordonnée (de telle sorte que : $1 = \pi(x_1) \geq \pi(x_2) \geq \dots \geq \pi(x_N)$), comme suite :

$$Sp(\pi) = \pi(x_1) - \sum_{j=2}^N \omega_j \cdot \pi(x_j) \quad \text{Eq1.11}$$

où : $\{\omega_j\}, j=1..N$ est un ensemble de poids vérifiant les propriétés suivantes :

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

Dans le cas d'une distribution de possibilités anormale π , la mesure de spécificité est donnée par :

$$Sp(\pi) = h(\pi) - \sum_{j=2}^N \omega_j \cdot \pi(x_j) \quad \text{Eq1.12}$$

Supposons que les versions normalisées de cette distribution anormale, selon les trois méthodes de normalisation précédemment indiquées, sont respectivement π_{Ord} , π_{Num} et π_{Dec} . Le Tableau 1.4 donne les expressions de spécificité calculées en fonction de la nature de la normalisation de la distribution de possibilité. Les calculs sont détaillés dans [AIS 2012].

Tableau 1.4 Spécificité en fonction de la nature de la normalisation de la distribution de possibilité

Type de la normalité	Spécificité
Distribution de possibilité anormale	$Sp(\pi) = h(\pi) - \sum_{j=2}^N \omega_j \pi(x_j)$
Distribution de possibilité normalisée (Normalisation ordinale)	$Sp(\pi_{Ord}) = Sp(\pi) + Inc(\pi)$
Distribution de possibilité normalisée (Normalisation numérique)	$Sp(\pi_{Num}) = \frac{1}{h(\pi)} Sp(\pi)$
Distribution de possibilité normalisée (par décalage de l'inconsistance)	$Sp(\pi_{Dec}) = Sp(\pi)$
Distribution de possibilité anormale et Mesure de nécessité de Dubois	$Sp_{Dub}(\pi) = Sp(\pi)$

b) Mesure d'incertitude U

Soient l'ensemble fini de N singletons $\Omega = \{x_1, x_2, \dots, x_N\}$ et une distribution de possibilité $\pi : \Omega \rightarrow [0,1]$ ordonnée sur Ω (i.e. les valeurs de possibilités $\pi(x_n)$ sont disposées dans un ordre décroissant de telle sorte que : $\pi(x_1) \geq \pi(x_2) \geq \dots \geq \pi(x_N)$). La mesure U de π est définie par :

$$U(\pi) = [1 - \pi(x_1)]. \ln(N) + \sum_{n=2}^N ([\pi(x_n) - \pi(x_{n+1})]. \ln(n)) \quad \text{Eq1.13}$$

où $\pi(x_{n+1}) = 0$. La première partie de la somme généralise la mesure U au cas d'une distribution de possibilité anormale.

Si la distribution de possibilité π est anormale de hauteur $h(\pi)$. La mesure d'incertitude U est donnée par :

$$U(\pi) = [1 - h(\pi)]. \ln(N) + \sum_{n=2}^N ([\pi(x_n) - \pi(x_{n+1})]. \ln(n)) \quad \text{Eq1.14}$$

En considérant π_{Ord} , π_{Num} et π_{Dec} , les distributions de possibilité normalisées de cette distribution anormale π , suivant, respectivement, les trois méthodes de normalisation précédemment indiquées, le Tableau 1.5 donne les expressions de mesures U calculées en fonction de la nature de la distribution de possibilité. Les calculs sont détaillés dans [AIS 2012].

c) Indice de confiance

Soit l'événement $A \subseteq \Omega$ ayant les mesures de possibilité $\Pi(A)$ et de nécessité $N(A)$. L'indice de confiance en la réalisation de cet événement, *Ind* est défini par [AIS 2012]:

$$Ind(A) = \Pi(A) + N(A) - 1 \quad \text{Eq1.15}$$

Notons que :

- ✓ $Ind(A) \in [-1,1]$;
- ✓ $Ind(A) = -1$, si $N(A) = \Pi(A) = 0$; et
- ✓ $Ind(A) = +1$, si $N(A) = \Pi(A) = 1$.

Tableau 1.5 Mesure U en fonction de la nature de la normalisation de la distribution de possibilité (Tab.2.6 de [ALS2012])

Type de la normalité	La mesure U
Distribution de possibilité anormale	$U(\pi) = \sum_{n=2}^N (\pi(x_n) - \pi(x_{n+1})) \log_2(n) + [1 - h(\pi)] \log_2(N)$
Distribution de possibilité normalisée (Normalisation ordinale)	$U(\pi_{Ord}) = U(\pi) - Inc(\pi) \log_2(N)$
Distribution de possibilité normalisée (Normalisation numérique)	$U(\pi_{Num}) = \frac{1}{h(\pi)} U(\pi)$
Distribution de possibilité normalisée (par décalage de l'inconsistance)	$U(\pi_{Dec}) = U(\pi) - Inc(\pi) \log_2(N)$
Distribution de possibilité anormale et Mesure de nécessité de Dubois	$U_{Dub}(\pi) = U(\pi)$

Tableau 1.6 Indice de confiance Ind en fonction de la nature de la normalisation de la distribution de possibilité (Tab.2.7 de [ALS 2012])

Type de la normalité	L'indice de confiance Ind
Distribution de possibilité anormale	$Ind(A) = \Pi(A) + N(A) - 1$
Distribution de possibilité normalisée (Normalisation ordinale)	$Ind_{Ord}(A) = \begin{cases} 0 & \text{Si } \Pi(A) = \Pi(A^C) = h(\pi) \\ 1 & \text{Si } \Pi(A) < h(\pi) \text{ et } \Pi(A^C) < h(\pi) \\ 1 - \Pi(A) & \text{Si } \Pi(A) = h(\pi) \text{ et } \Pi(A^C) < h(\pi) \\ \Pi(A) - 1 & \text{Si } \Pi(A) < h(\pi) \text{ et } \Pi(A^C) = h(\pi) \end{cases}$
Distribution de possibilité normalisée (Normalisation numérique)	$Ind_{Num}(A) = \frac{1}{h(\pi)} Ind(A)$
Distribution de possibilité normalisée (par décalage de l'inconsistance)	$Ind_{Dec}(A) = Ind(A)$
Distribution de possibilité anormale et Mesure de nécessité de Dubois	$Ind_{Dub}(A) = Ind(A) + Inc(\pi)$

En considérant π_{Ord} , π_{Num} et π_{Dec} , les distributions de possibilité normalisées de cette distribution anormale π , suivant, respectivement, les trois méthodes de normalisation

précédemment indiquées, le Tableau 1.6 donne les expressions de la mesure d'incertitude Ind , calculées en fonction de la nature de la distribution de possibilité et faisant intervenir les mesures de possibilités et nécessités respectives : $\Pi_{Ord}(A)$, $N_{Ord}(A)$, $\Pi_{Num}(A)$, $N_{Num}(A)$ et $\Pi_{Dec}(A)$, $N_{Dec}(A)$. Les calculs sont détaillés dans [AIS 2012].

Le Tableau 1.7 illustre les effets respectifs des approches de normalisation, ainsi que l'effet de l'application de la proposition de *Dubois* sur une distribution anormale pour garder les propriétés de dualité entre les mesures de possibilité et de nécessité, sur chaque mesure d'incertitude (Sp , U , et Ind). Les détails sont donnés dans [AIS 2012].

Tableau 1.7 Effet des approches de normalisation et de l'adoption de la Nécessité de *Dubois* sur les mesures d'incertitude : (-) signifie que la mesure est changée, (+) signifie que la mesure est conservée

Type de la normalisation	Sp	U	Ind
Normalisation ordinale	-	-	-
Normalisation numérique	-	-	-
Normalisation par décalage de l'inconsistance	+	-	+
Adoption de la Nécessité de <i>Dubois</i>	+	+	-

4. Outils de fusion

La fusion peut se définir comme étant une combinaison d'informations imparfaites et hétérogènes afin d'obtenir des informations plus complètes qui permettent un meilleur raisonnement. De nombreuses approches de fusion d'informations sont proposées dans la littérature [Sol 2001, LKT 2018, PBF 2014, SBP 2015, BS 2016]. Favoriser l'une d'elles par rapport aux autres se base sur la fiabilité, le conflit et l'interdépendance des données fournies par les sources d'informations. [Roh 2007] a classé les approches de fusion d'informations possibilistes en deux catégories : celle qui procède par propagation à travers un modèle analytique, et celle qui se base sur des opérateurs ensemblistes.

Soient π_1 et π_2 deux distributions de possibilités, modélisant respectivement deux informations A_1 et A_2 fournies respectivement par deux sources S_1 et S_2 . Nous proposons dans ce qui suit de décrire la manière de fusionner les deux informations A_1 et A_2 , au sein de chaque catégorie.

a) Propagation à travers un modèle analytique

La fusion se fait selon un modèle analytique défini par la fonction f . La distribution de possibilité modélisant l'information résultante de la fusion des informations A_1 et A_2 , est : $\pi = f(\pi_1, \pi_2)$.

Selon [DP 1988d], le calcul se fait par analyse des α -coupes. Il s'agit d'un calcul classique d'intervalle, pour chaque α -coupe définie par son degré de possibilité α .

b) Combinaison par les opérateurs ensemblistes

Trois modes de combinaison existent :

i) Le mode conjonctif

La fusion conjonctive, $\pi_{\wedge}(\cdot)$, des deux distributions, π_1 et π_2 est définie par [DP 1994] :

$$\forall x \in \Omega, \pi_{\wedge}(x) = \pi_1(x) \otimes \pi_2(x) \quad \text{Eq1.16}$$

où \otimes représente l'opérateur de la fusion conjonctive qui 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-norm : $\pi_{\wedge}(x) = \max(0, \pi_1(x) + \pi_2(x) - 1)$,
- ✓ etc.

Le mode conjonctif généralise l'intersection ensembliste. Il correspond à un consensus entre les informations fusionnées. Il permet une satisfaction simultanée des contraintes induites par le processus de fusion. Le résultat d'un opérateur conjonctif est toujours contenu dans tous les éléments d'informations fournies par les sources. Par conséquent, ce mode de fusion est bien adapté quand les informations sont tout à fait fiables et les sources ne sont pas conflictuelles, dans le sens qu'elles soulignent le même intérêt.

Il est à signaler que l'opérateur conjonctif réduit l'incertitude globale et fournit un résultat plus précis que chacune des informations provenant des sources prises séparément. La Figure 1.7 montre un exemple de fusion conjonctive en utilisant l'opérateur « min ».

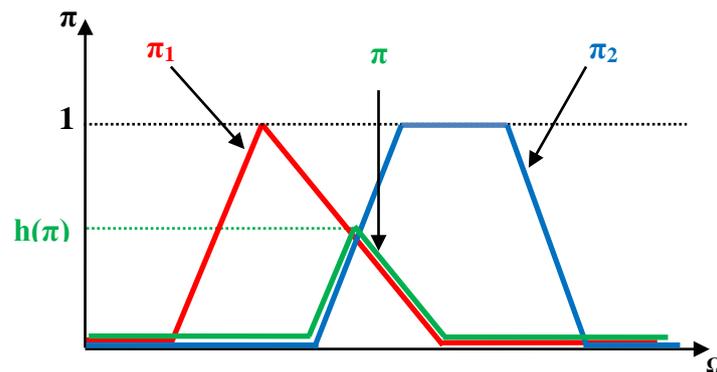


Figure 1.7 Fusion conjonctive par l'opérateur min (Fig.2.2 de [ALS 2012])

ii) Le mode disjonctif

La fusion disjonctive, $\pi_{\vee}(\cdot)$, de deux distributions, π_1 et π_2 , se définit par :

$$\forall x \in \Omega, \pi_{\vee}(x) = \pi_1(x) \perp \pi_2(x) \quad \text{Eq1.17}$$

où \perp représente l'opérateur de la fusion disjonctive. Cet opérateur 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-conorm : $\pi_{\vee}(x) = \min(1, \pi_1(x) + \pi_2(x))$,
- ✓ etc.

Le mode disjonctif, généralise l'union ensembliste. Le résultat d'un tel opérateur disjonctif contient toutes les informations données par les sources. Il fait l'hypothèse qu'au moins l'une

des sources est fiable. Le résultat d'une telle opération est généralement très fiable, mais peut être très imprécis. Ce qui réduit son utilité.

Il est à signaler qu'un opérateur disjonctif augmente l'incertitude globale et fournit un résultat moins précis que chacune des sources prise séparément. La Figure 1.8 montre un exemple de fusion disjonctive en utilisant l'opérateur « max ».

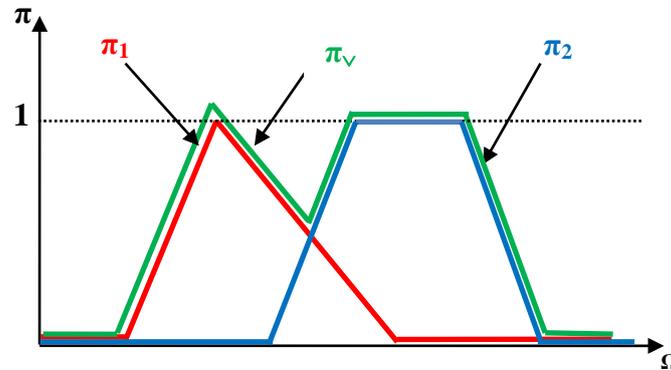


Figure 1.8 Fusion disjonctive par l'opérateur max (Fig.2.3 de [ALS 2012])

iii) Le mode adaptatif ou intermédiaire

Ce mode a été proposé par Dubois et Prade [DP 1992]. Le comportement de l'opérateur adaptatif est intéressant dans le sens où il évolue progressivement d'une combinaison conjonctive à un mode disjonctif, dès que le conflit entre les sources augmente.

La fusion adaptative π_{AD} est donnée par :

$$\pi_{AD} = \max \left[\frac{\min(\pi_1, \pi_2)}{h}, \min[\max(\pi_1, \pi_2), (1 - h)] \right] \quad \text{Eq1.18}$$

Ce mode intermédiaire est lié au degré de consensus h entre π_1 et π_2 , à savoir la borne supérieure de leur intersection (Figure 1.9-a).

Le premier terme $\frac{\min(\pi_1, \pi_2)}{h}$ représente la normalisation de la combinaison conjonctive (intersection) des sources initiales.

Le deuxième terme $\min[\max(\pi_1, \pi_2), (1 - h)]$ représente, quant à lui, l'ensemble des solutions plausibles en dehors de la zone de consensus. Il permet donc de restreindre l'influence du conflit dans le support des deux sources.

Cette règle donne la préférence à la zone de consensus normalisée. En dehors de cette zone, le résultat de la combinaison est supporté de manière équivalente par les deux distributions avec un degré de certitude au moins égal à h [OMB 2003, Ous 2000].

Un exemple de la fusion adaptative est illustré dans la Figure 1.9-b.

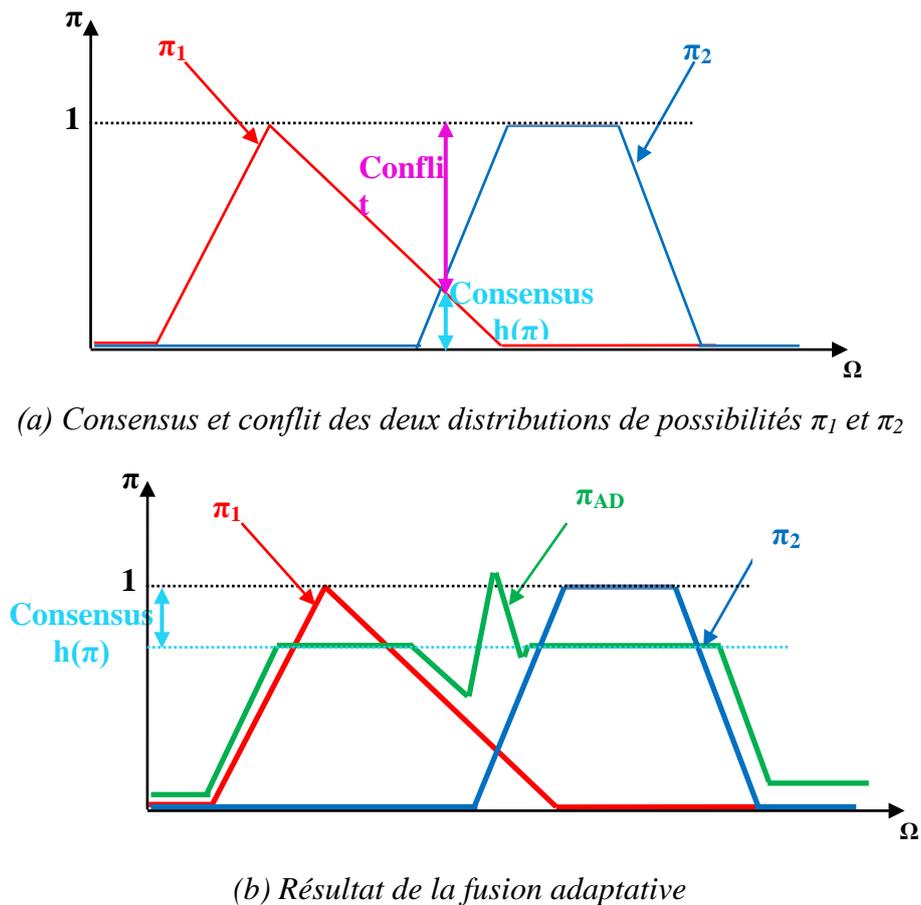


Figure 1.9 Illustration du mode intermédiaire entre combinaison conjonctive et disjonctive avec prise en compte du conflit $1-h$ des sources d'information π_1 et π_2 (modification de la Fig.2.4 de [ALS 2012])

5. Critères de décision

La théorie des possibilités opère dans un contexte d'incertitude, ce qui signifie que la prise de décision est essentielle.

Etant donné un état de connaissance, l'occurrence de l'une des alternatives de l'ensemble de discours Ω lui est certainement reliée. Si l'évidence de l'occurrence et l'unicité de cette alternative sont certaines, grâce aux hypothèses de l'exclusivité et l'exhaustivité de Ω , son identité précise ne l'est pas (l'imperfection est de type incertitudes). Cependant, il est évident que cette alternative soit la plus cohérente avec l'état de connaissance actuel. Un ordonnancement de préférences sur toutes les alternatives de l'univers Ω est alors à établir, selon leurs degrés de cohérences respectifs avec cet état de connaissance. La théorie de possibilités fournit deux mesures ensemblistes permettant d'estimer la cohérence d'une alternative avec un état de connaissance, à savoir la mesure de possibilité et la mesure de nécessité de l'occurrence de l'alternative, étant donné cet état de connaissance. La problématique de prise de décision possibiliste consiste à définir une règle de décision permettant d'identifier l'unique alternative réalisée à travers ce cas, en se basant sur cet ordonnancement donné par les estimations de cohérences respectives de toutes les alternatives avec cet état de connaissance.

Bien que la théorie des possibilités opère dans un contexte d'incertitude, ce qui signifie que la prise de décision est essentielle, peu de travaux analysant cette question existent dans la littérature [AIS 2012, ALS 2012, Ezi 2012]. Une analyse intéressante de cette problématique est présentée dans [Ezi 2012]. Nous proposons dans ce qui suit, d'en présenter un résumé. Considérons pour la suite un ensemble de discours Ω et un cas T .

a) Mesure de possibilité (resp. nécessité) maximale

Cette règle de décision, implique que l'alternative dont la mesure de possibilité (resp. nécessité) étant donné le cas T , est la plus élevée, soit considérée comme la plus compatible avec l'échantillon T .

Notons que cette règle de décision peut être enrichie. En effet, la règle peut être encore plus stricte, en n'acceptant la prise de décision que si la valeur de la mesure de possibilité (resp. nécessité) dépasse un seuil S prédéfini, que l'on appelle seuil possibiliste de décision. Avec cette condition, le cas où aucune des alternatives de l'univers des discours ne répond au critère, est à considérer. Le rejet de toutes les alternatives peut être envisagé, tout en gardant l'hypothèse de l'exhaustivité (du monde fermé). Ceci est possible par l'ajout, à l'univers des discours, d'une $(K+1)^{\text{ème}}$ alternative, dite de rejet,

b) Indice de confiance maximal

Dans [PK 1993] un indice quantifiant le degré de confiance dans la compatibilité de l'alternative C avec le cas T , est défini sur l'intervalle $[-1, +1]$, par :

$$\text{Ind}(C/T) = \Pi(C/T) + N(C/T) - 1 \quad \text{Eq1.19}$$

L'alternative ayant l'indice le plus élevé est celle considérée comme la plus compatible avec le cas T .

c) Moyenne maximale

Le critère de décision basé sur la moyenne des mesures de possibilités et de nécessité a été proposé dans [RNM 2003, PG 2009]. Cette mesure, notée Ψ , est définie sur l'intervalle $[0, 1]$ par :

$$\Psi(C/T) = \frac{\Pi(C/T) + N(C/T)}{2} \quad \text{Eq1.20}$$

Dans ce cas, l'alternative fournissant la mesure de moyenne maximale est celle considérée comme la plus compatible avec le cas T .

d) Préférence sur le couple $[\Pi, N]$

Dans d'autres travaux [DP 1995, DPT88, Dub 1988, ADK 2000], l'ordonnancement est réalisé en considérant la mesure de possibilité et de nécessité comme un couple :

$$[N(C_1/T) ; \Pi(C_1/T)] > [N(C_2/T) ; \Pi(C_2/T)]$$

si et seulement si :

$$\Pi(C_1/T) > \Pi(C_2/T) \text{ et } N(C_1/T) \geq N(C_2/T),$$

e) Mesure de similarité maximale

Dans ce cas, l'alternative la plus similaire au cas T est la plus compatible. La similarité est elle-même quantifiée par des mesures, qui sont étudiées plus loin dans ce manuscrit (Cf. Chapitre 3 du présent manuscrit).

VI. Etat de l'art des méthodes d'estimation d'une distribution de possibilités

L'estimation de distributions de possibilités est l'une des problématiques les plus abordées dans la théorie des possibilités. Un choix approprié de la forme de la fonction de distribution de l'incertitude associée aux connaissances n'est généralement pas une tâche facile [CWC 2000, NB 2007]. Nous proposons de classer les techniques d'estimation des fonctions de distributions de possibilités, que nous avons pu identifier dans la littérature, dans deux catégories : les méthodes guidées par les connaissances d'expert et les méthodes à base d'inférences. La première catégorie regroupe les techniques basées sur la « projection » des connaissances exprimées par les experts, afin de représenter directement, ou indirectement, l'allure générale des distributions des possibilités. En revanche, les méthodes de la deuxième catégorie exploitent des fonctions caractéristiques d'autres théories de l'incertain, telles que une distribution de probabilités, une fonction d'appartenance ou une fonction de masse, afin d'inférer une distribution de possibilités.

Dans la suite, nous présentons brièvement les deux catégories d'estimation de distribution de possibilités. Nous portons cependant un intérêt particulier aux transformations des distributions de probabilités en distributions de possibilités [DP 1982, DS 1985]. Notons qu'une telle transformation représente une certaine cohérence conceptuelle, car les distributions de probabilités et de possibilités s'appliquent sur la même forme d'imperfection de l'information, qui est l'incertitude.

1. Les méthodes d'estimation guidées par l'expert

Le principe général de ces méthodes consiste à réaliser la projection d'une propriété caractérisant une alternative ou classe thématique sous la forme d'une distribution de possibilités en se basant sur les connaissances d'expert. Cette projection est effectuée par un expert qui choisit la forme de la distribution de possibilités qui lui semble la plus adéquate. Cette sélection est réalisée à partir d'un ensemble comportant des formes standards de distributions de possibilités. L'évaluation des degrés de possibilités par ces méthodes dépend donc du savoir-faire de l'expert et de sa capacité à représenter les propriétés caractérisant les classes thématiques d'une manière qui permette de mesurer la force de liaison entre les éléments de la classe et les propriétés considérées. Deux approches sont essentiellement identifiées : l'approche directe et l'approche indirecte.

a) Approche directe

Dans cette approche [AAT 2005, KY 1995], le processus d'estimation de la distribution des possibilités est totalement dépendant des connaissances de l'expert et de son expérience, quant à la définition de l'allure de la distribution, ainsi qu'au choix de ses paramètres. Le processus d'estimation de cette méthode peut être résumé comme suit : tout d'abord, l'expert détermine, selon le problème traité, les formes de la distribution des possibilités qui correspondent au mieux aux propriétés décrivant les classes à caractériser. L'expert choisit, ensuite, d'une façon manuelle, empirique et directe les valeurs des paramètres associés à la distribution des possibilités.

Notons que cette approche représente une évaluation peu robuste des degrés de possibilités, étant donné qu'elle repose uniquement sur l'expérience propre à l'expert qui a réalisé cette mission. Un expert différent aurait pu, très probablement, réaliser d'autres choix, engendrant ainsi une distribution des possibilités d'une toute autre allure pour estimer la même connaissance. En d'autres termes, le résultat de l'estimation des valeurs des paramètres, et donc les degrés de possibilités, sont sujets à une variabilité inter - opérateurs.

b) Approche indirecte

Dans ce type d'approches, on a bien recours aux compétences de l'expert pour une définition approximative de la forme la plus appropriée pour l'estimation de la distribution des possibilités. Il peut avoir recours à des formes standards (comme les fonctions : trapézoïdales, triangulaire, gaussienne, S, Z, etc.) bien détaillées dans [Ezi 2012], ou prédéfinies, décrivant l'allure de la distribution des possibilités relatives à certaines propriétés ambiguës. La distribution des possibilités en question, est par la suite ajustée par action sur les paramètres de la forme standard ou prédéfinie choisie, via un processus automatique d'optimisation [NB 2007, CC 1997, Pun 1981]. Ceci permet d'accomplir cette tâche d'estimation d'une manière plus efficace et surtout plus objective. De tels processus utilisent des mesures floues, considérées comme des fonctions objectives et calculées à partir des contenus informationnels issus du monde réel, évaluant l'information ambiguë contenue dans une alternative. On peut citer les mesures de l'entropie floue [NB 2007, KH 2007, CC 1997], la compacité [PR 1988, TB 1994] ou l'indice floue [NB 2007].

Différents algorithmes d'optimisation ont été proposés dans la littérature pour l'estimation des paramètres associés à une distribution de possibilités. Parmi ces algorithmes nous pouvons citer : les algorithmes génétiques [Mic 1998], le recuit simulé [KGV 1983] et une méthode appelée fast searching [Zah 2006].

Les algorithmes génétiques et l'algorithme du recuit simulé sont utilisés dans divers domaines tels que la robotique, l'industrie automobile, etc. On note également une large utilisation de ces algorithmes en traitement d'images, et spécialement en segmentation, où leur rôle consiste à estimer les paramètres des fonctions d'appartenance des différentes classes de l'image. L'optimisation de ces paramètres s'effectue, pour chacun de ces deux algorithmes, d'une manière stochastique [Zah 2006]. Ceci explique le fait que les résultats obtenus par ces algorithmes, peuvent différer d'une réalisation à une autre.

Ces mesures, ainsi que quelques algorithmes d'optimisation qui les manipulent en vue d'estimer des paramètres des distributions de possibilités, sont détaillés dans [Ezi 2012].

Remarquons que l'approche indirecte peut être considérée comme une technique «intelligente» d'estimation. Elle conduit à des résultats plus précis que l'approche directe grâce aux valeurs optimisées des paramètres, en fonction de mesures liées au contenu informationnel. Par conséquent, les valeurs d'appartenance estimées sont indépendantes de la subjectivité de l'expert humain.

2. Transformation d'une distribution de probabilités en une distribution de possibilités

Établir une relation entre la représentation de l'incertitude probabiliste et la représentation de l'incertitude floue (plus largement possibiliste) a bénéficié d'une attention particulière depuis les propos de Zadeh à cet égard [Zad 1978]. Cette relation ou transformation de la représentation de l'incertitude est une question clé dans le traitement des deux types d'incertitude dans un système donné. Elle se fait dans les deux sens et son défi consiste à utiliser les avantages et les pertinences des deux théories par leurs intégrations. L'intégration des deux théories se fait dans le sens que l'incertitude représentée dans l'une peut être convertie par une transformation justifiable en une représentation équivalente dans l'autre [WK 1992, Mas 2004]. Les représentations probabilistes et possibilistes ne sont pas deux représentations équivalentes d'incertitude [DFM 2004]. Par conséquent, il ne devrait y avoir aucune symétrie entre les procédures, à savoir la conversion et sa réciproque.

- ✓ Transformer une mesure de probabilité en une mesure de possibilité peut être utile dans le cas de la présence de multiples sources d'information, de faible qualité, ou lorsque le calcul avec des possibilités est plus simple qu'avec des probabilités.
- ✓ Transformer une mesure de possibilité en une mesure de probabilité pourrait être d'intérêt dans la prise de décision [Sme 1990].

Dans le présent travail, notre intérêt est focalisé sur les transformations probabilités-possibilités. Plusieurs travaux dans la littérature se sont intéressés à cette question, tels que [CT 1986, DP 1988d, Jum 1993, DP 1994, DP 1993, DPS 1993, Kli 1993] et plus récemment [Mou 2002, MD 2006 et Bou 2007]. Dans ces travaux, des transformations mathématiques ont été présentées et différents aspects de la relation probabilités-possibilités ont été discutés. Nous proposons dans ce qui suit, de présenter quelques principes reliant les distributions de probabilités à celles des possibilités avant de détailler quelques techniques, les plus connues dans la littérature, que nous classons dans deux catégories en se référant à la nature de distribution de probabilités considérée. En effet, on retrouve celles qui sont élaborées en se basant sur des mesures précises de probabilités et d'autres qui se basent sur des intervalles de confiance probabilistes, donnant ainsi plus de flexibilité à la présentation des données.

a) Principes de cohérences entre les distributions probabilité/Possibilité

i) Principe de de Zadeh : cohérence d'ordonnement

Zadeh a défini le principe de cohérence probabilités-possibilités comme "un haut degré de possibilité n'implique pas un degré élevé de probabilité, ni un faible degré de

probabilité n'implique un faible degré de possibilité. Toutefois, si un événement est impossible, il est forcément improbable.» [Zad 1978].

Soit $X=\{x_1, x_2, \dots, x_n\}$ un univers fini de singletons, Zadeh a défini le degré de cohérence entre une distribution de probabilités $p_i=p_i(x_i)$ et une distribution de possibilités $\pi_i=\pi(x_i)$ comme:

$$C_z = \sum_i p_i \pi_i \quad \text{Eq1.21}$$

Zadeh a souligné que la cohérence probabilités-possibilités, définie en Eq1.21, n'est pas une loi précise ou une relation entre les distributions de possibilités et celles de probabilités, mais une formalisation approximative du principe qui dit qu'« une diminution de la possibilité d'un événement tend à diminuer sa probabilité, mais pas vice-versa » [Zad 1978]. Pour la cohérence d'ordonnement, Zadeh a également formulé un principe illustré par:

$$\forall x \in X : \pi(x) \geq p(x) \quad \text{Eq1.22}$$

ii) Principe de Yager : spécificité Maximale

Yager [GYT 2003] a défini le principe de spécificité maximale probabilité-possibilité « la bonne transformation est la transformation qui fournit une distribution de possibilité ayant une valeur de spécificité maximale (i.e. un contenu informatif maximal) ».

Ce principe est le suivant : soit π_1 et π_2 deux distributions de possibilités définies sur un univers de X , on dit que π_2 est plus spécifique que π_1 si :

$$\forall x \in X : \pi_2(x) < \pi_1(x)$$

Ce principe permet de trouver la distribution la plus informative, et la plus conservative d'informations. En effet, la distribution la plus spécifique est celle qui a le plus petit ensemble flou, donc le degré de spécificité d'une distribution π est mesurée par [GYT 2003]:

$$S_p = \sum_i \pi_i \quad \text{Eq1.23}$$

iii) Conservation de la forme

Ce critère permet de garantir que la distribution de possibilité obtenue à la même forme que la distribution de probabilité de départ. La vérification du critère nécessite la conservation forte de la préférence. Il existe deux types de conservation de la préférence :

1) La conservation faible défini par :

$$\forall w, u \in \Omega, p(w) > p(u) \Leftrightarrow \pi(w) > \pi(u) \quad (15)$$

2) La conservation forte définie consécutivement par:

$$\forall w, u \in \Omega$$

- $p(w)=p(u) \Leftrightarrow \pi(w) = \pi(u) \quad (16)$

- $p(w)<p(u) \Leftrightarrow \pi(w) < \pi(u) \quad (17)$

- $p(w)>p(u) \Leftrightarrow \pi(w) > \pi(u) \quad (18)$

iv) Conservation de l'ignorance

Ce critère permet de garantir la conservation du cas de l'ignorance totale lors de la transformation. Le cas de l'ignorance totale est indésirable parce qu'il correspond au cas où on ne peut pas prendre de décision.

Pour la suite, on se pose dans le contexte où on dispose de :

- N observations (X_1, X_2, \dots, X_N) suivant une distribution de probabilités p réparties sur K classes de l'univers $\Omega = \{\omega_1, \omega_2, \dots, \omega_K\}$.
- n_i désigne le nombre d'observations dans la $i^{\text{ème}}$ classe ω_i .
- Le vecteur $n = (n_1, n_2, \dots, n_K)$ est la réalisation d'une variable aléatoire multinomiale, de paramètres $p = (p_1, p_2, \dots, p_K)$, où chaque $p_i = p(\omega_i) > 0$

et
$$\sum_{i=1}^K p_i = 1$$

- π désigne une distribution de possibilités et $\pi_i = \pi(\omega_i)$

b) Estimation de distribution de possibilités à partir de mesures précises de probabilités

i. Transformations de Dubois et Prade

Parmi les différentes transformations proposées dans la littérature la transformation de Dubois-Prade [DP 1993] est la seule transformation qui vérifie deux conditions de qualité. En effet, la distribution de possibilité obtenue par cette transformation représente la distribution la plus spécifique que l'on peut obtenir. De plus, l'inégalité $p(x) \leq \pi(x)$ est toujours vérifiée quel que soit $x \in \Omega$.

Les deux transformations $p \rightarrow \pi$ et $\pi \rightarrow p$ sont définies respectivement par : [DP 1988e]

$$\pi_i = \sum_{j=i}^n p_j, \tag{Eq1.24}$$

$$p_i = \sum_{j=i}^n \frac{(\pi_j - \pi_{j+1})}{j} \tag{Eq1.25}$$

Ces deux transformations ne sont pas réciproques, d'où l'appellation de la transformation asymétrique (ou optimale) de Dubois et Prade.

Dubois et Prade ont défini une autre transformation $p \rightarrow \pi$ symétrique satisfaisant le principe de cohérence de Dubois et Prade défini par [DPS 1993] :

$$\pi_i = \sum_{j=i}^n \min(p_j, p_i) \tag{Eq1.26}$$

ii. Transformation de Klir

La transformation de Klir est basée sur son principe de conservation d'incertitude [Kli 1993] à travers l'équation suivante :

$$U(p) = U(\pi) \tag{Eq1.27}$$

Où U est l'incertitude contenue dans une distribution.

L'incertitude contenue dans une distribution de probabilités correspond à l'entropie de Shanon donnée par :

$$U(p) = \sum_{i=1}^n p_i \times \log_2(p_i) \tag{Eq1.28}$$

L'incertitude incluse dans une distribution de possibilités peut être, selon Klir, définie par :

$$U(\pi) = Sp(\pi) + D(\pi) \quad \text{Eq1.29}$$

Où :

$$Sp(\pi) = \sum_{i=2}^n (\pi_i - \pi_{i+1}) \cdot \log_2(i) \quad \text{Eq1.30}$$

présente l'incertitude d'une proposition qui n'est pas spécifique et :

$$D(\pi) = \sum_{i=2}^n (\pi_i - \pi_{i+1}) \cdot \log_2\left(\frac{i}{\sum_{j=1}^i \pi_j}\right) \quad \text{Eq1.31}$$

est le degré de conflit interne de la mesure $Sp(\pi)$

Klir a considéré le principe de l'invariance d'incertitude par deux échelles :

- ✓ L'échelle de rapport où les transformations $p \rightarrow \pi$ et $\pi \rightarrow p$ sont définies par des transformations normalisées :

$$\pi_i = \frac{p_i}{p_{\max}}, \quad p_i = \frac{\pi_i}{n \cdot \sum_{i=1}^n \pi_i} \quad (5) \quad \text{Eq1.32}$$

- ✓ Les échelles log-intervalle où les transformations $p \rightarrow \pi$ et $\pi \rightarrow p$ sont définies par :

$$\pi_i = \left(\frac{p_i}{p_{\max}}\right)^\alpha, \quad p_i = \frac{\pi_i^{1/\alpha}}{\sum_{i=1}^n \pi_i^{1/\alpha}} \quad (6) \quad \text{Eq1.33}$$

où α est un paramètre qui appartient à l'intervalle ouvert] 0, 1 [.

c) Estimation de distribution de possibilités à partir des intervalles de confiance probabilistes

Une approche classique pour construire une distribution de possibilités consiste à attribuer le vecteur des fréquences observées $f = (f_1, f_2, \dots, f_K)$ où $f_i = n_i / N$ au vecteur de probabilités p et d'appliquer directement l'une des techniques de transformation de Dubois et Prade ou Klir. Cependant, cette approche ne tient pas compte du processus d'échantillonnage permettant d'étendre le résultat d'un échantillon à une population, ou encore du fait que les échantillons des données disponibles puissent être réduits entraînant des informations incomplètes, ce qui rend la distribution de possibilité obtenue par cette méthode insatisfaisante. Une alternative aux mesures de probabilités, permettant plus de flexibilité pour l'extension du résultat d'un échantillon à une population, serait l'emploi des intervalles de confiance (IC) de probabilités comme alternative aux mesures de probabilités [MJ 1997].

Masson et Denoeux [MD 2006] ont suggéré une méthode pour générer une distribution de possibilités à partir des intervalles de confiance (IC) simultanés. Cette approche généralise la transformation asymétrique de Dubois et Prade, pour une proportion multinomiale. Masson a

ensuite proposé une procédure de calcul permettant de diminuer remarquablement le temps de calcul de la méthode précédente.

Soit un ensemble d'intervalles de probabilité simultanés :

$$I = \{ [p_i^-, p_i^+] \mid i = 1, \dots, K \} \tag{Eq1.34}$$

tel que la probabilité P pour que p_i appartienne à $[p_i^-, p_i^+]$ soit supérieur à $(1-\alpha)$. P est appelée probabilité ou taux de couverture de p_i . A un taux de couverture équivalent, le meilleur estimateur est celui de volume le plus faible (de variance la plus faible).

On obtient alors :

$$p(\omega_i) \leq p(\omega_j) \iff p_i^+ \leq p_j^- \tag{Eq1.35}$$

Il est bien évident que l'ordre linéaire donné dans Eq1.35 ne se réalise pas lorsque les intervalles $[p_i^-, p_i^+]$ et $[p_j^-, p_j^+]$ sont en chevauchement. Une solution consiste à définir un ordre partiel \mathcal{P} tel que :

$$\sigma(i) \leq \sigma(j) \iff (\omega_{\sigma(i)} \leq \omega_{\sigma(j)}) \in \mathcal{P} \iff p_i^+ \leq p_j^- \tag{Eq1.36}$$

L'ordre partiel $\wedge \mathcal{P}$ peut être représenté par l'ensemble des permutations $\{\sigma^\ell, \ell = 1, \dots, L\}$.

La distribution de possibilités proposée par [MD 2006] est la solution du système suivant :

A) Pour toute permutation possible σ_ℓ et pour chaque classe ω_i , résoudre le programme linéaire (PL) suivant :

$$\pi_{\sigma^\ell(i)} = \sum_{j, \sigma^{\ell-1}(j) \leq \sigma^{\ell-1}(i)} p_j \tag{Eq1.37}$$

Sous les contraintes :

$$\left\{ \begin{array}{l} \sum_{j=1}^K p_j = 1 \\ p_j^- \leq p_j \leq p_j^+ \mid j = 1, \dots, K \\ p_{\sigma(1)} \leq p_{\sigma(2)} \leq \dots \leq p_{\sigma(K)} \end{array} \right.$$

B) La distribution la plus spécifique dominant toutes les distributions π_{σ_ℓ} est ensuite retenue

$$\pi_i = \max_{\ell=1, \dots, L} \pi_{\sigma^\ell(i)}, \quad i = 1, \dots, K \tag{Eq1.38}$$

Plusieurs méthodes existent pour construire des intervalles de confiance IC pour des proportions multinomiales comme les méthodes de Goodman(1965), de Fitzpatrick & Scott (1987), de Sison & Glaz(1995) [MJ 1997]. Chacune des méthodes présente une faveur d'utilisation.

Masson et Denoeux ont utilisé la méthode de construction des ICs de Goodman [Goo 1965] dans leur proposition. Cette procédure assure l'obtention d'une distribution de possibilités qui satisfait le principe de cohérence.

Masson a introduit des modifications sur la procédure de calcul afin de réduire le temps de calcul. Les détails sont donnés dans [MD 2006].

3. Estimation d'une distribution de possibilités à partir d'une fonction d'appartenance

La théorie des ensembles flous permet de modéliser fidèlement l'ambiguïté des informations. Pour traiter ce type d'imperfections, il faut estimer une fonction d'appartenance.

De nombreuses méthodes ont été proposées dans la littérature pour l'estimation des fonctions d'appartenance [MKK 1998]. Elles peuvent être soit guidées par les connaissances d'expert soit des méthodes à base d'apprentissage [Ezi 2012].

Une fois la fonction d'appartenance est estimée, le passage de l'espace flou à l'espace possibiliste est possible grâce au postulat de Zadeh [Zad 1978]. Ce principe suppose que la fonction d'appartenance qui décrit une classe peut être considérée numériquement égale à la distribution des possibilités de cette classe. La valeur de possibilité $\pi_C(x)$ pour qu'un élément x soit décrit par une caractéristique C est égale à son degré d'appartenance $\mu_C(x)$:

$$\pi_C(x) = \mu_C(x) \quad \text{Eq1.39}$$

4. Estimation de distribution de possibilités à partir d'une distribution de masse

Une fonction ou distribution de masse est la fonction caractéristique de la théorie d'évidence. Le choix d'une telle fonction est basé sur des opinions d'experts qui dépendent totalement de leurs connaissances, ce qui rend son estimation subjective. Pour cela, l'estimation des fonctions de masses présente l'inconvénient majeur de l'emploi de la théorie de l'évidence. Cependant, cette théorie gagne un intérêt croissant pour le traitement des incertitudes notamment dans le domaine de fusion et qualité de l'information [YF 2016, NR 2005]

Supposons qu'on a ℓ alternatives (AI) possibles pour un problème donné d'aide à la décision. Chaque AI^q , ($q=1, \dots, \ell$) est évaluée selon n critères C_i (pour $i=1, \dots, n$).

Considérons un ensemble de r évaluateurs Evl (r experts) présentant des évaluations partielles. Chaque expert Evl^α ($\alpha=1, \dots, r$) va attribuer un degré de satisfaction $x_i^{\alpha,q}$ selon le critère C_i pour l'alternative AI^q .

Si les degrés de satisfactions sont donnés sous formes de groupes disjoints d'intervalles de confiances disjoints ayant de fortes intersections, alors, il s'agit au sens de Shafer de «distributions de masses ». Dubois et Prade proposent dans [DP 1986], une approche pour construire une distribution de possibilités approximant une distribution de masses sur un groupe disjoint d'intervalles ayant de fortes intersections. Les détails de construction des distributions de possibilités π_i^q relatives aux différents critères d'évaluations C_i à partir des distributions de masses, sont explicités dans [Den 2007].

VII. Conclusion

Nous avons présenté dans ce chapitre des concepts de base nécessaires à la compréhension du contexte théorique des axes académiques entamés dans le présent rapport.

Dans la première partie de ce chapitre, nous avons étudié les caractéristiques générales d'une information ainsi que les imperfections pouvant l'affecter. La modélisation d'information tenant compte des imperfections rencontrées, a été abordée. Il s'agit du processus permettant le passage d'une représentation sous forme de données issues de différentes sources de connaissances, vers une représentation mathématique de l'information selon une théorie particulière. Une telle modélisation devrait faciliter le traitement des informations en vue d'obtenir de meilleures performances des décisions ultérieures. Plusieurs théories mathématiques sont utilisées pour une telle finalité. Quatre d'entre elles, connues par leurs capacités à représenter et à gérer des informations imparfaites ont été décrites dans ce chapitre : la théorie des probabilités, la théorie de l'évidence, la théorie des ensembles flous et la théorie des possibilités. D'après notre étude, nous avons trouvé que ces théories ne traitent pas de la même façon l'incertitude dans l'information. L'intérêt a été spécialement focalisé sur la dernière théorie du fait qu'elle présente une meilleure capacité à représenter et à décrire les informations entachées d'ambiguïté, d'imprécision et de conflits, qui figurent naturellement dans les concepts et les informations manipulés par l'homme et intervenant dans les situations de prises de décision.

L'objectif principal du présent travail étant l'étude et le développement d'outils de raisonnement possibiliste pour l'aide à la décision et interprétation de scènes, une large investigation de l'état de l'art des outils, concepts et notions liés à cette théorie a été menée et l'essentiel a été résumé dans ce chapitre.

Un intérêt particulier a été porté à la définition et aux techniques d'estimation des distributions de possibilités, étant les modèles mathématiques adoptés par la théorie des possibilités pour la représentation de l'information incertaine. Ces modèles ne nécessitent que l'injection d'un faible niveau de connaissances pour leur construction. La théorie des possibilités assure ainsi une modélisation facile des connaissances faiblement définies. Ceci nous amène à l'objectif du chapitre suivant à savoir l'étude de la modélisation possibiliste de connaissances incertaines. Nous présentons également nos contributions dans cette thématique faisant l'objet de l'une des problématiques les plus rencontrées au cours de la conception des systèmes d'aide à la prise de décision.

Chapitre II: **Modélisation possibiliste de
connaissances**

CHAPITRE II: MODELISATION POSSIBILISTE DE CONNAISSANCES.....	52
I. INTRODUCTION	55
II. SYNTHÈSE DES ACTIVITÉS RELATIVES À CE CHAPITRE.....	57
1. Collaborations et encadrements.....	57
2. Publications.....	58
a) Articles de Revue	58
b) Conférences internationales à comité de lecture.....	58
III. ESTIMATION ET INFÉRENCE DE DISTRIBUTIONS DE POSSIBILITÉS POUR LA MODÉLISATION DE DONNÉES TENANT COMPTE DE LEURS ALTERITÉS	59
1. Modification apportée à la technique d'inférence de Masson et Denoeux.....	60
2. Mise en œuvre d'une stratégie d'évaluation.....	61
a) Description de la stratégie.....	61
b) Approches à comparer	61
c) Données utilisées pour les expérimentations.....	61
d) Critères d'évaluations.....	62
e) Choix d'une mesure de similarité	63
3. <i>Résultats</i>	65
a) Évaluation qualitative : critères psycho visuels	65
b) Évaluation quantitative	67
c) Discussions	71
IV. MODÉLISATION POSSIBILISTE DE CONNAISSANCES CONTEXTUELLES DE VOISINAGE	72
1. <i>Modes de représentations de connaissances pixelliques</i>	72
2. <i>1^{ère} Contribution : Modélisations possibilistes de connaissances pixelliques</i>	74
a) Définition du contexte.....	74
b) Problématique.....	75
c) Hypothèse 1 : Modélisation possibiliste ponctuelle du pixel	75
d) Hypothèse 2 : Modélisation possibiliste contextuelle du pixel.....	77
3. <i>Résultats et discussion</i>	78
a) Contexte des simulations:	78
b) Tests et évaluation sur une image de synthèse:.....	79
c) Tests et évaluation sur des images médicales optiques:.....	84
4. <i>2^{ème} Contribution (en cours de développement): Modélisation contextuelle spatiale locale de deuxième ordre possibiliste d'un pixel ; Notion de possibilité des possibilités</i>	87
a) Description de l'idée.....	87
b) Simulation sur un échantillon.....	89
V. SÉLECTION D'ATTRIBUTS EN MODÉLISATION POSSIBILISTE	90
1. <i>Problématique</i>	90
2. <i>Cadre de l'étude</i>	91
3. <i>Outils techniques</i>	92

Chapitre 2 : Modélisation possibiliste de connaissances

a)	Indice de Shapley.....	92
b)	Mesure d'incertitude.....	93
4.	<i>Description de la contribution</i>	93
5.	<i>Expérimentations et discussion</i>	98
VI.	CONCLUSION.....	101

I. Introduction

Une scène peut être décrite sous forme d'un ensemble de mesures acquises par un ou plusieurs capteurs et présentées sur divers genre de supports (image, son, vidéo, description linguistique, etc.). La mesure du capteur dépend du phénomène physique acquis par ce capteur au niveau élémentaire seul. Par conséquent, cette mesure est fortement influencée par les sources de bruit liées au processus d'acquisition et à l'environnement. En apparence, l'information délivrée par cette mesure ne porte aucun sens sémantique de haut niveau (il s'agit, juste, d'une mesure physique). Mais, en réalité l'information véhiculée par cette représentation va bien au-delà de la seule donnée des mesures acquises par un capteur en chaque élément (valeur d'une grandeur à un instant ou en une position donnés). En effet, ces mesures radiométriques conduisent à une interprétation de la scène acquise: i.e. un label désignant le sens sémantique, ou thématique, de chaque élément du support de l'information acquise.

Sur une échelle sémantique d'interprétation, quatre niveaux de représentation des connaissances se placent, comme le montre la Figure 2.1, selon la sémantique qu'ils englobent.

- Niveau des données brutes : La représentation de l'élément de la scène par les données brutes du capteur (i.e. mesure physique) se positionne au plus bas niveau de l'échelle.
- Niveau intermédiaire des primitives : Le niveau intermédiaire des primitives consiste à représenter l'élément de la scène par des paramètres déduites des données radiométriques, des interactions locales avec ses voisins, de transformations appliquées sur ces données brutes ou en injectant des connaissances externes.

Ainsi à chaque élément de la scène est associé un vecteur de primitives le représentant et désignant un autre point de vue de la scène.

Parmi les primitives, on peut citer à titre d'exemple :

- ✓ Les valeurs filtrées des grandeurs mesurées
- ✓ Entropie, moyenne, variance locale...de chaque élément de la scène
- ✓ La distance par rapport à des centroïdes de classes, ...
- Niveau sub-sémantique : Le niveau sub-sémantique consiste à injecter des connaissances *a priori* liées aux classes thématiques dans la scène à interpréter pour en établir une modélisation destinée au système d'interprétation qui va suivre la phase de la modélisation. Pour ce niveau, l'élément de la scène est projeté dans un autre espace de connaissances, un espace où son appariement, aux modèles de classes de la scène, s'avère être plus intéressant. La représentation des connaissances élémentaires est alors un vecteur présentant les scores de proximité de l'élément de la scène par rapport aux modèles de chacune des classes thématiques.

A titre d'exemple on peut citer comme modèles :

- ✓ Distributions de probabilités
- ✓ Distributions de masses
- ✓ Distributions de possibilités
- Niveau symbolique ou sémantique : C'est le plus haut niveau de l'échelle représentant l'élément de la scène par des connaissances symboliques telles que étiquettes (ou

labels) d'appartenance à une classe thématique. En ce niveau, l'ensemble des éléments de la scène se transforme en une carte thématique. Elle présente le résultat de l'interprétation finale. Cette interprétation consiste à accorder un label à chaque élément de la scène. Ce label est considéré comme le plus haut niveau sémantique représentant l'élément. La définition de l'ensemble de labels (classes), est fixée *a priori* par l'expert après analyse des structures qui composent la scène.

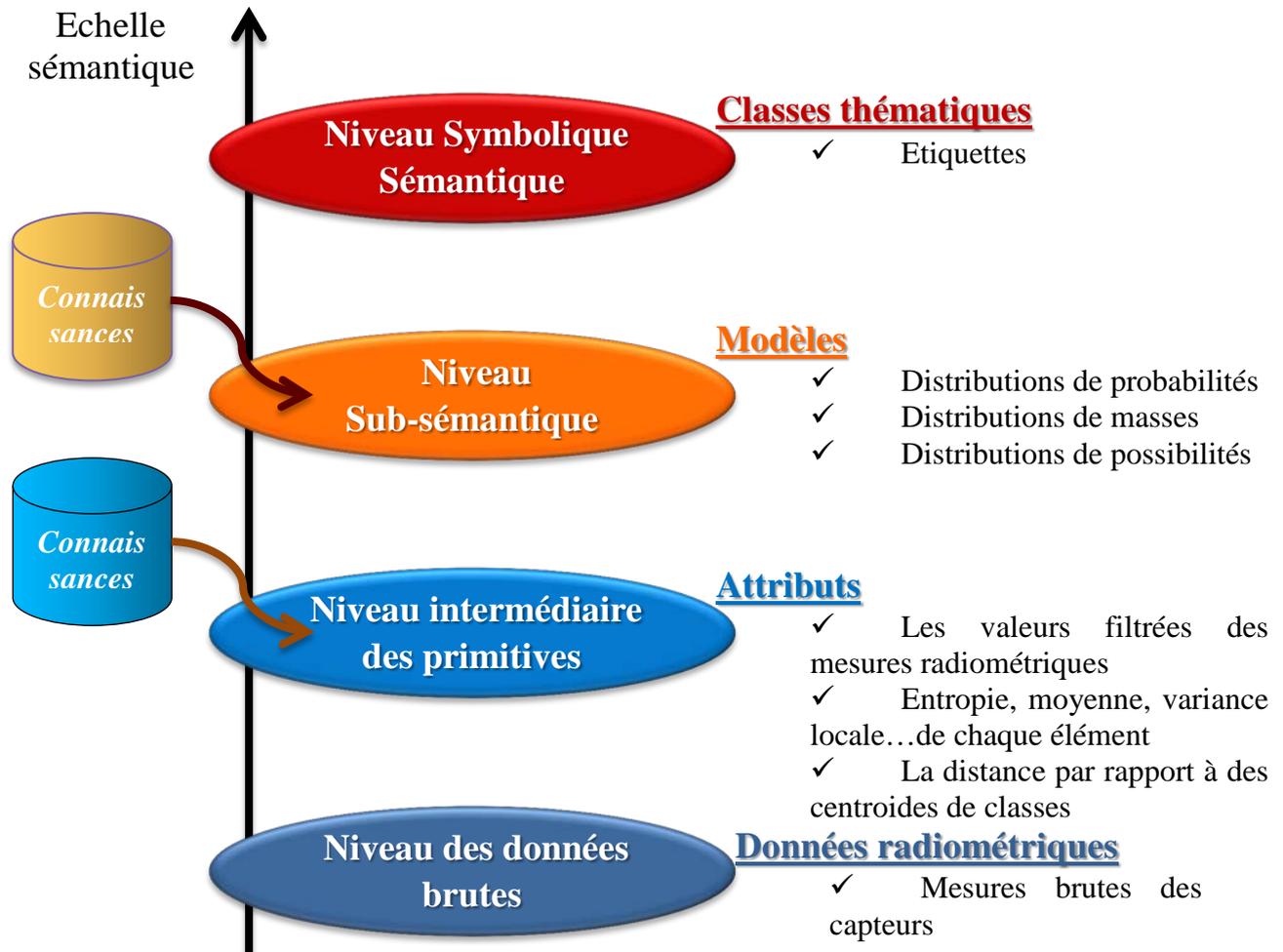


Figure 2.1 Niveaux de représentation des connaissances sur une échelle sémantique d'interprétation

Nous nous intéressons dans ce chapitre à la représentation des connaissances liées aux éléments de la scène au niveau sub-sémantique. Un modèle universel n'existe pas. Le choix doit s'effectuer en se référant à la nature et aux conditions de disponibilité des données. En se situant dans le contexte d'ambiguïté, c'est le modèle possibiliste fourni par les distributions de possibilité que nous adoptons pour représenter les connaissances d'une scène. Dans ce cadre, se situe les travaux présentés par ce chapitre. Une synthèse des activités est d'abord donnée, suivie par la description des contributions effectuées pour modéliser des connaissances dans des circonstances particulières d'ambiguïté. Chaque contribution est évaluée à travers des expérimentations menées sur différents domaines d'applications.

II. Synthèse des activités relatives à ce chapitre

1. Collaborations et encadrements

Mes travaux sur la modélisation possibiliste de connaissances sont essentiellement réalisés en collaboration avec le Pr. Basel Solaiman, Professeur à l'IMT Atlantique (France) et deux membres de son équipe du département iTi : Dr. Shaban Almouahed et Dr. Bassem Alsahwa. Je co-encadre actuellement avec Pr. Solaiman une thèse incluse dans le cadre de modélisation de connaissances issues de multi sources en vue de proposition d'un modèle pour un environnement d'intérieur structuré, sans avoir recours à des connaissances *a priori*.

1. Marwa Chakroun, « Perception de scène par un système multi-capteurs, application à la navigation dans des environnements d'intérieur structuré », thèse en co-tutelle avec iTi de l'IMT-Atlantique, Brest, France. La thèse est sous la direction conjointe de Mme Houda Derbel, Maître de conférences à la faculté de Sciences de Sfax et Mr. Basel Solaiman, Professeur de l'IMT-Atlantique, Brest, France.

J'ai également coopéré avec Pr. Dorra Sellami, Professeur à l'ENIS, dans le cadre de co-encadrement de trois thèses liées respectivement à, la modélisation de l'état du sol en vue de sa classification en trois catégories (sol plat, escaliers ascendants et descendants), la modélisation de connaissances en vue de segmentation du sol pour l'aide à la navigation et la modélisation de connaissances contenues dans une texture d'iris en vue de l'identification:

2. Sonda Ammar, « Analyse et traitement possibiliste de signaux ultrasonores en vue d'assistance des non-voyants », thèse sous la direction de Mme. Dorra Sellami, Professeur de l'Ecole Nationale d'Ingénieurs de Sfax. Cette thèse a été soutenue à l'ENIS, le 4 Mars 2014.
3. Jihen Frikha, « Segmentation du sol par la théorie de possibilité- application à l'autonavigation des malvoyants », thèse sous la direction de Mme. Dorra Sellami, Professeur de l'Ecole Nationale d'Ingénieurs de Sfax. Cette thèse a été soutenue à l'ENIS, le 24 Novembre 2016.
4. Majd Bellaaj, « Application de la théorie de possibilités en reconnaissance d'individus par l'iris », thèse sous la direction de Mme. Dorra Sellami, Professeur de l'Ecole Nationale d'Ingénieurs de Sfax. Cette thèse a été soutenue à l'ENIS, le 07 Décembre 2016.

Tous ces travaux ont eu recours à l'estimation de distributions de possibilités, étant les modèles adoptés pour la représentation des connaissances. Afin de bien maîtriser cette étape essentielle à l'évolution de mes travaux j'ai encadré le projet de mastère suivant :

5. Mariem CHAABENE, “ Etude et comparaison d'approches d'estimation de distributions de possibilités dans des contextes ambigus “. Présenté à l'Institut Supérieur d'Informatique et de Multimédia En vue de l'obtention du diplôme de Mastère De Recherche en Informatique et Multimédia (I2M), Soutenu le 17/06/2014.

Durant ce projet une étude panoramique de l'état de l'art sur les approches d'estimation et d'inférence des distributions de possibilités a été effectuée, suivie d'une contribution sur l'estimation de ces distributions basée sur les intervalles de confiance. Cette

contribution a ouvert le chemin pour une nouvelle recherche sur la représentativité possibiliste dans un contexte de Big Data à partir d'un échantillon de petite taille. Ce travail est en cours de développement en collaboration avec Pr. Basel Solaiman et Dr. Sonda Ammar.

2. Publications

Les résultats de ces travaux ont été publiés à travers 7 articles de Revues et 7 articles de conférences internationales.

a) Articles de Revue

1. **I. Khanfir Kallel**, S. Almouahed, B. Alsahwa and B.Solaiman, "The use of contextual spatial knowledge for Low-quality Image Segmentation", *Multimedia Tools and Applications*, 2018.
2. M. Bellaaj, **I. Khanfir Kallel** and D. Sellami, "Uncertainty Theories Based Iris Recognition System", *Electronic Letters on Computer Vision and Image Analysis* 16(2):29-32; 2017.
3. J. Frikha Elleuch, D. Sellami and **I. Khanfir Kallel**, "Indoor/ outdoor navigation system based on possibilistic traversable area segmentation for visually impaired people", *Electronic Letters on Computer Vision and Image Analysis* 15(1), pp.60-76, 2016.
4. M. Bellaaj, D. Sellami and **I. Khanfir Kallel**, "Possibilistic Modeling of Iris System for High Recognition Reliability", *IJCSNS International Journal of Computer Science and Network Security*, Vol.16 No.7, pp. 34-47, 2016.
5. S. Ammar Bouhamed, **I. Khanfir Kallel**, D. Sellami Masmoudi and B. Solaiman "Feature selection in possibilistic modeling", *Pattern Recognition*, Volume 48 Issue 11, p. 3627-3640, 2015.
6. J. Frikha Eleuch, **I. Khanfir Kallel**, D. Sellami Masmoudi, "New ground plane segmentation method for electronic cane", *Journal of Image and Graphics, (JIG)*, Vol. 1, No. 2, pp. 72-75, 2013.
7. S. Ammar Bouhamed, **I. Khanfir Kallel**, D. Sellami Masmoudi. "New electronic white cane for stair case detection and recognition using ultrasonic sensor", *International Journal of Advanced Computer Science and Applications, (IJACSA)*, Vol. 4, No. 6, pp. 243-256, 2013.

b) Conférences internationales à comité de lecture

1. M. Medhioub, S. Ammar Bouhamed, **I. Khanfir Kallel**, N. Derbel, B. Solaiman and O. Kanoun, "An Electronic Embedded System for Stair Recognition", *Sensorica 2017: Series of Lectures of the University of Applied Sciences Ruhr West*, June, 8th - 9th 2017-Vol. 6.
2. **I. Khanfir Kallel** and Basel Solaiman, "Contextual Possibilistic Modeling Of Pixellic Knowledge For Tumor Segmentation In Mammographic Images", 3rd IEEE international conference on Advanced Technologies for Signal and Image Processing (ATSIP'2017), May 22-24, 2017, Fez, Morocco.
3. M. Bellaaj, J. Frikha Elleuch, D. Sellami, **I. Khanfir Kallel**. "An Improved Iris Recognition System Based on Possibilistic Modeling", *MoMM Proceedings of the*

- 13th International Conference on Advances in Mobile Computing and Multimedia, pp. 26-32, 2015.
4. J. Frikha Elleuch, M. Bellaaj, D. Sellami and **I. Khanfir Kallel**. “Traversable area segmentation approach at indoor environment for visually impaired people”, MoMM Proceedings of the 13th International Conference on Advances in Mobile Computing and Multimedia, pp. 48-53, 2015.
 5. M. Bellaj, **I. Khanfir**, D. Sallemi and B. Solaiman, “Possibility Theory based Approach for Iris Recognition in Biometric Systems”. 1st IEEE International Conference on Advanced Technologies for Signal & Image Processing ATSIP 2014, March 17-19, 2014, Sousse-Tunisia.
 6. M. Chaâbène, **I. Khanfir Kallel**, S. Ammar Bouhamed, M. Bellaj, W. Bouchaala, “CIs based strategy for estimating possibility distribution of small data samples”. 1st IEEE International Conference on Advanced Technologies for Signal & Image Processing ATSIP 2014, March 17-19, 2014, Sousse-Tunisia.
 7. J. Frikha, **I. Khanfir** and D. Sallemi, “A new ground segmentation method for obstacle detection using possibility theory”. 1st IEEE International Conference on Advanced Technologies for Signal & Image Processing ATSIP 2014, March 17-19, 2014, Sousse-Tunisia.

III. Estimation et inférence de distributions de possibilités pour la modélisation de données tenant compte de leurs altérités

L’insuffisance des données, l’incomplétude de l’information et l’indisponibilité de la quantité de mesures nécessaires engendrent une grande difficulté pour faire une analyse d’inférence statistique significative nécessaire à la prise de décision sous l’effet de l’apparition d’incertitudes épistémiques. La théorie des possibilités, permet cependant le traitement de ce type d’incertitude. D’après notre étude de l’état de l’art sur les méthodes d’estimations de distributions de possibilités, il s’est montré que les transformations de probabilités en possibilités conviennent mieux aux situations où l’on dispose de prises de mesures, du fait qu’elles intègrent des données statistiques. Ces données statistiques permettent d’étendre le résultat d’un échantillon à une population quand elles sont d’une quantité assez importante. Cependant, les échantillons des données disponibles sont, dans la majorité des cas, réduits entraînant des informations incomplètes, ce qui génère des distributions de probabilités insatisfaisantes. L’utilisation des intervalles de confiance (IC) de probabilités comme alternative aux mesures de probabilités, présente une solution très intéressante pour permettre plus de flexibilité quant à l’extension du résultat d’un échantillon à une population. Cette alternative a été adoptée par Masson et Denoeux dans [MD 2006], où ils ont pu remédier au manque informationnel des données pour étendre la représentativité du modèle à une population. En revanche, le manque quantitatif des données continue à présenter une limite pour leur approche. En effet, la technique qu’ils ont adoptée pour la construction de

leurs intervalles de confiance simultanés, se base sur une estimation paramétrée nécessitant un nombre relativement important de données pour ajuster ses paramètres.

Dans ce qui suit nous présentons notre contribution pour la modélisation possibiliste à partir d'un échantillon de petite taille, qui consiste, d'abord, en une modification apportée à une technique d'estimation de distribution de possibilités à partir des intervalles de confiance probabilistes, proposée par Masson et Denoeux dans [MD 2006]. Nous proposons, ensuite, une stratégie pour une évaluation objective, aussi bien qualitative que quantitative ainsi qu'une comparaison de cinq approches d'estimation de distributions de possibilités, à savoir celle de Masson et Denoeux, sa modification que nous proposons et trois autres approches basiques, connues dans la littérature et présentées dans le chapitre 1 de ce rapport.

1. Modification apportée à la technique d'inférence de Masson et Denoeux

Masson et Denoeux [MD 2006] se sont basés sur une généralisation de la transformation asymétrique de Dubois et Prade pour une proportion multinomiale, afin de générer une distribution de possibilités à partir des intervalles de confiance simultanés de Goodman [Goo 1965] avec un niveau de confiance $1 - \alpha$. Leur approche permet d'avoir une distribution de possibilité qui satisfait le principe de cohérence. Cependant, les performances baissent spécialement quand on dispose d'un échantillon de données relativement de faible taille [MD 2006], car la construction des intervalles de confiance de Goodman se basent sur un test d'ajustement du Khi deux, qui exige des effectifs d'échantillons suffisamment grands afin de détecter des différences pratiques avec une probabilité élevée dans les calculs [Min 2015].

Masson a ensuite proposé une procédure de calcul permettant d'alléger remarquablement le temps de l'exécution, mais sans remédier à la question de la taille de l'échantillon.

Notre apport touche à ce point précisément. Une étude est alors élaborée pour trouver une approche plus appropriée pour la construction des ICs, étant donné une quantité réduite de l'information disponible.

Suite à une étude bibliographique sur les approches de construction d'intervalles de confiance simultanés, nous avons pu constater que la technique de Sison & Glaz est bien appropriée pour les échantillons de petite taille.

Cette méthode est basée sur l'approximation de distribution de probabilités multinomiale en utilisant l'algorithme de Levin (1981) Levin [Lev 1981]. Elle utilise la relation entre la distribution multinomiale, la distribution de poisson et la distribution de poisson tronquée pour le calcul des Intervalles de confiances simultanés pour des proportions multinomiales.

En considérant la proportion multinomiale de paramètres inconnus p_1, p_2, \dots, p_K , l'intervalle de confiance de Sison et Glaz avec un niveau de confiance de $(1-\alpha)$ est donné par :

$$[p_i^-, p_i^+] = \left[\left[\frac{x_i}{n} - \frac{c}{n}, \frac{x_i}{n} + \frac{c + 2\eta}{n} \right] \right] \quad \text{Eq2.1}$$

où $c \geq 1$ est le nombre entier satisfaisant l'inégalité : $v(c) \leq 1 - \alpha \leq v(c + 1)$ et η est définie pour vérifier l'asymétrie de la distribution multinomiale. Plus de détails sont donnés dans [SG 1995, Cha 2014 et Amm 2014].

2. Mise en œuvre d'une stratégie d'évaluation

a) Description de la stratégie

Notre intérêt dans ce chapitre est focalisé sur la modélisation. La procédure d'évaluation devrait donc mesurer l'efficacité de la représentativité de l'approche à évaluer. La meilleure représentation du comportement d'une population selon une caractéristique donnée est fournie par l'expert. Evaluer la représentativité d'une approche de modélisation se rend alors à l'apparier à cette représentation de référence. La stratégie d'évaluation adoptée est donnée par Figure 2.2.

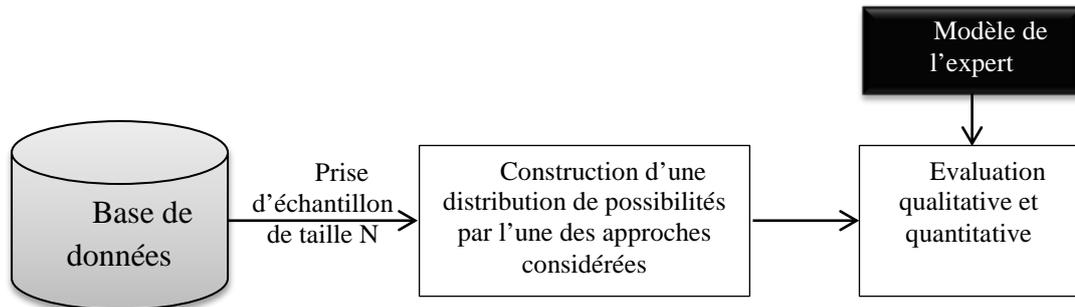


Figure 2.2 Stratégie d'évaluation d'approches de modélisation possibilistes

b) Approches à comparer

Nous n'avons considéré que les méthodes d'estimation de distribution de possibilités à partir d'une distribution de probabilités, vu l'abondance des applications mettant en jeu des données numériques pouvant être présentées sous formes d'histogrammes. Cinq approches sont considérées dont trois consistent à estimer les distributions de possibilités à partir des probabilités simples, à savoir les transformations symétrique et asymétrique de Dubois et Prade et la transformation de Klir, et deux utilisent des probabilités de type intervalle pour estimer des distributions de possibilités, à savoir l'approche de Masson et Dencœur utilisant les intervalles de confiances probabilistes de Goodman et sa version modifiée que nous proposons par le remplacement des Intervalles de confiances de Goodman par les Intervalles de confiances de Sison et Glaz.

c) Données utilisées pour les expérimentations

Trois types de données ont été utilisés pour évaluer les 5 approches d'estimation des distributions de possibilités en fonction de la taille de l'échantillon utilisé pour la modélisation : données synthétiques, une base de données de référence standard 'SVM Guide' [HCL 2003] et des données réelles acquises par notre équipe par un dispositif expérimental d'aide à la navigation (présenté dans le chapitre 4 de ce manuscrit).

❖ Données synthétiques

Une distribution de possibilités aléatoirement construite, a été considérée comme étant le modèle de référence (assimilée à une représentation de l'expert du comportement d'une population donnée selon un attribut donné). Soit « PossExp » la distribution de possibilités de référence répartie sur dix valeurs possibilistes (nombre de classes $K=10$) :

$$\text{PossExp}=[0.3 ; 0.3 ; 0.2 ; 0.55 ; 0.14 ; 0.66 ; 0.36 ; 0.9 ; 0.1 ; 1]$$

Une distribution de probabilités « P », en est déduite par la technique de transformée inverse de Dubois et Prade selon l'équation suivante :

$$P_i = \sum_{j=i}^K \frac{\pi_j - \pi_{j+1}}{j} \quad \text{Eq2.2}$$

« P » est donnée par :

P= [0.035 ; 0.035 ; 0.0215 ; 0.095 ; 0.014 ; 0.131; 0.0475; 0.251; 0.01; 0.351]

P sert de base, aux cinq approches à évaluer, pour estimer leurs modèles possibilistes à partir du nombre d'échantillons à considérer.

❖ **Données de référence 'SVM Guide'**

SVM Guide est une base de données standard contenant 924 éléments répartis équitablement sur deux classes (462 éléments chacune) et observés par 22 caractéristiques. Pour une classe et un attribut, les $\frac{3}{4}$ des mesures (observations) ont servi à la construction des distributions de possibilités de référence, par l'emploi de l'approche symétrique de Dubois et Prade. Du quart restant, sont extraits les échantillons de test, de tailles variables.

❖ **Données ultrasonores**

Un dispositif expérimental multi-capteurs a été conçu au sein de notre équipe dans le cadre de nos travaux pour l'aide à la navigation [Amm 2014, Fri 2016, Cha 2016, Med 2016]. Le dispositif est entre autre équipé d'un capteur ultrason servant à détecter l'état du sol en 'escalier ascendant', 'escalier descendant' ou 'sol plat'. Pour la reconnaissance, plusieurs primitives ont été extraites de multiples présentations de l'écho ultrasonore [Amm 2014]. Nous effectuons les présents tests d'évaluation des approches de modélisation pour une des trois classes et un attribut choisis au hasard. Dans nos travaux ultérieurs, ce n'est pas le hasard qui guidera notre choix, mais la qualité des données brutes à utiliser. L'objectif n'est pas d'en sélectionner les données les plus représentatives ou les plus discriminatives, mais plutôt d'évaluer leur qualité afin d'en prendre conscience lors de l'évaluation des performances des approches qui les utilisent pour la modélisation.

Dans les présentes expérimentations, nous formons notre base de données réelle de 45 acquisitions. Nous employons 30 mesures pour construire un modèle de référence, et nous gardons 15 mesures pour effectuer les prises d'échantillons. Notons que cette base est déjà affectée par le manque de données.

d) Critères d'évaluations

❖ **Critères qualitatifs psycho visuels**

- La proximité visuelle de la forme entre la distribution de possibilités générée et la distribution de référence
- La cohérence entre la distribution de probabilités théorique (ici calculée à partir d'un nombre important de données ; en effet, nous considérons que plus le nombre de données intervenant dans l'estimation d'une distribution de probabilité est important, plus celle-ci s'approche d'une distribution réelle de la population) et la distribution de possibilités inférée. Rappelons que toute distribution de possibilités majorant la distribution de probabilités, satisfait la condition de cohérence.

❖ **Critères quantitatifs**

- Ecart d'incertitude ΔU : quantifie la différence de la quantité d'information incluse dans une distribution de référence et celle contenue dans la distribution de possibilités estimée [HK 1982] :

$$|\Delta U| = |U(\pi(\text{Approche})) - U(\pi(\text{PossExp}))| \quad \text{Eq2.3}$$

Notons que plus $|\Delta U|$ tend vers zéro, plus la distribution de possibilités générée par l'approche conserve la quantité d'information incluse dans la distribution de possibilités de référence.

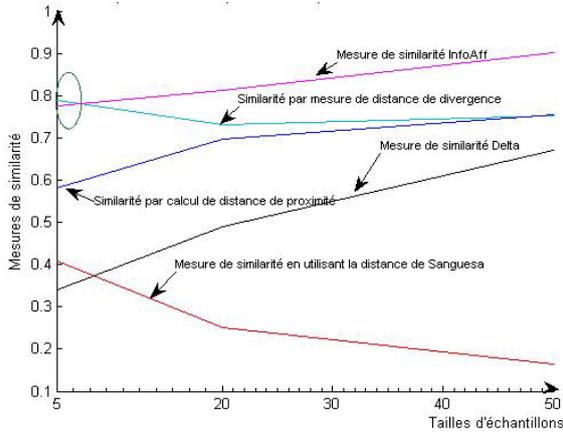
- PSNR (Peak Signal To Noise Ratio) est utile pour mesurer la performance des distributions de possibilités des approches développées par rapport à la distribution de l'expert pour les différentes tailles d'échantillons considérées. Cependant, elle ne peut pas être considérée comme une mesure objective de la qualité visuelle des formes des distributions. Cette mesure dépend étroitement de la mesure de l'PSNR (Mean Squared Error) qui permet de calculer l'erreur produite d'une distribution générée par rapport à une distribution originale. Celle-ci est acceptable dès qu'elle devienne supérieure à 30dB. En général, plus la valeur du PSNR est grande, meilleure est la qualité du signal obtenu. Cela signifie que les déformations créées sur la distribution de possibilités générée, sont difficiles à détecter par le système visuel humain.
- Mesure de similarité: calcule la similarité entre la distribution de possibilités estimée et la distribution de référence. Plusieurs mesures de similarité existent dans la littérature. Certaines sont standard, d'autres sont plus ciblées pour l'appariement de modèles. Dans ce qui suit on présente un test comparatif de quelques mesures de similarité prises de l'état de l'art afin d'en sélectionner la plus appropriée d'entre elles, pour l'appariement des distributions de possibilités.

e) Choix d'une mesure de similarité

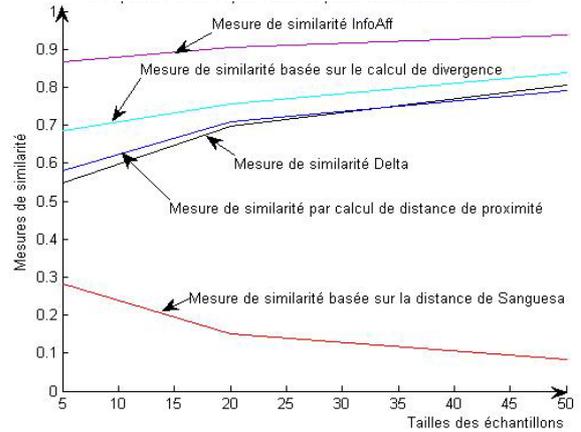
Pour évaluer la similarité entre deux distributions de possibilités, un simple calcul des distances conventionnelles, ne suffit pas. Plusieurs chercheurs ont tenté de trouver des solutions répondant à cet objectif. Quelques mesures se sont ainsi répandues dans la littérature, celles que nous citons dans ce qui suit, pour en sélectionner une pour nos expérimentations, sont décrites et détaillées dans [JBE 2007]. Il s'agit de la mesure de similarité par calcul de distance de proximité des informations, la mesure de similarité basée sur le calcul de la distance de Sanguesa en tant que non-spécificité de l'écart de distance entre 2 distributions, la mesure de similarité par calcul de divergence basée sur l'intégrale discrète de Choquet, la mesure de similarité delta δ et enfin la Mesure de similarité Information Affinity (*InfoAff*).

Le test que nous proposons pour faire le choix de la mesure de similarité à adopter, consiste à étudier le comportement des mesures de similarité, citées ci-dessus, en fonction de la taille d'échantillon pour chacune des approches d'estimation tel que le montre la Figure 2.3. Les similarités sont à chaque fois calculées entre deux distributions de possibilités estimées sur deux prises d'échantillon de mêmes tailles sur une même classe. Dans ce cas, le résultat parfait serait l'identité (similarité égale à l'unité). Notre choix est porté à la mesure de similarité qui permet d'avoir les plus importantes valeurs indépendamment des tailles de

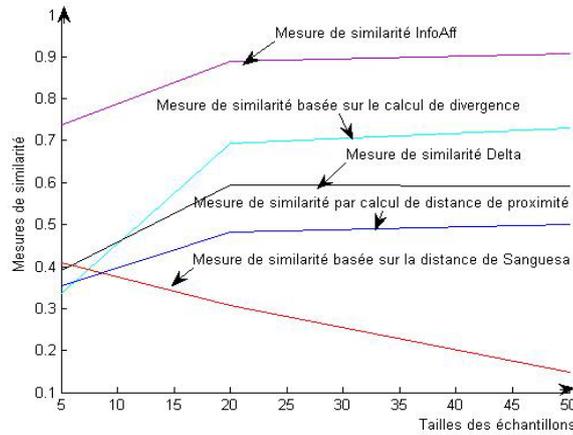
l'échantillon et des approches de modélisation. Le choix optimal correspond à la mesure de similarité « Information Affinity » (*InfoAff*).



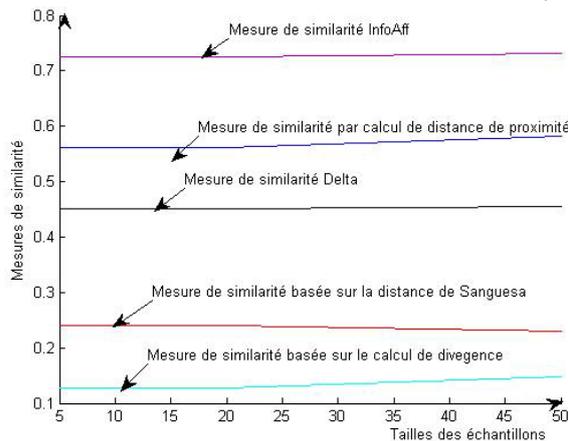
(a) Transformation asymétrique de Dubois et Prade



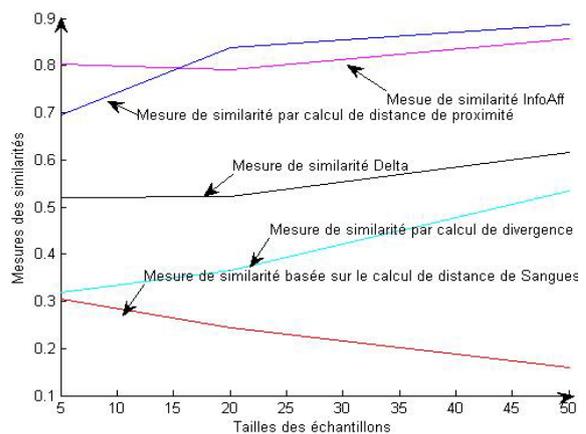
(b) Transformation symétrique de Dubois et Prade



(c) Transformation de Klir



(d) Approche de Masson utilisant les Intervalles de confiance de Goodman



(e) Approche de Masson avec les Intervalles de confiance de Sison et Glaz

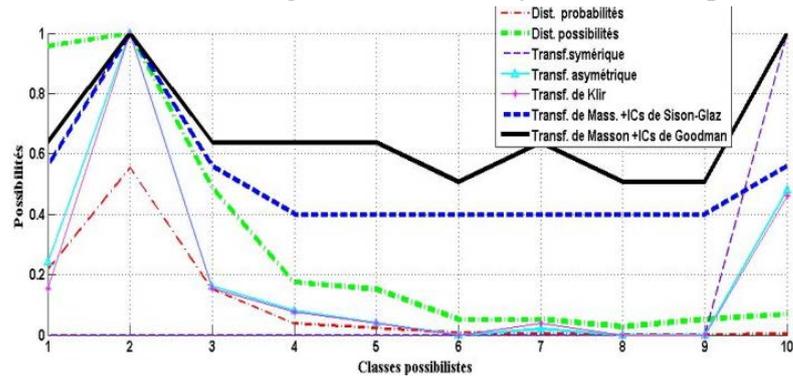
Figure 2.3 Comportement des mesures de similarité en fonction de la taille d'échantillon pour chacune des approches d'estimation possibilistes

3. Résultats

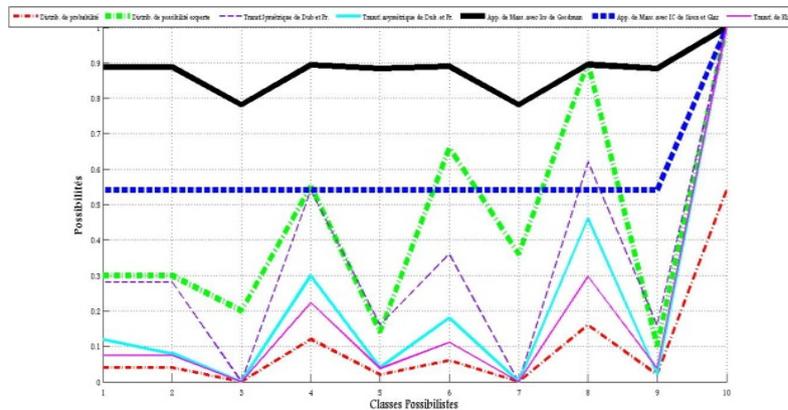
Dans ce qui suit nous effectuons nos tests sur trois gabarits d'échantillons : grand (N=50), moyen (N=20) et petit (N=5). Notons que pour la base de signaux ultrasonores, les approches d'estimation de distributions de possibilités ne sont testées que sur des échantillons de petit gabarit (N=5). Ceci est dû à la disponibilité limitée des données.

a) Evaluation qualitative : critères psycho visuels

Les figures Figure 2.4, Figure 2.5, Figure 2.6 illustrent les distributions de possibilités inférées d'une distribution de probabilités multinomiale, par les cinq approches testées pour, respectivement des gabarits d'échantillons grand (N=50), moyen (N=20) et petit (N=5).

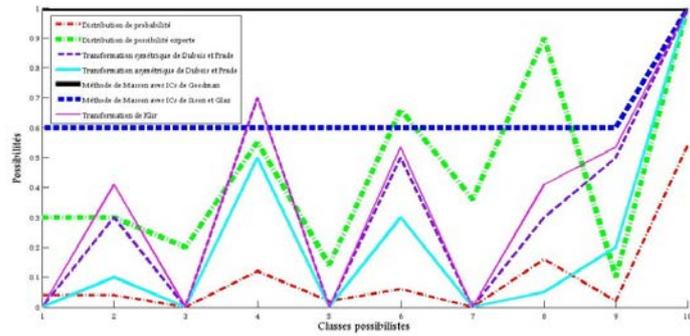


b) Pour la base 'SVM Guide'

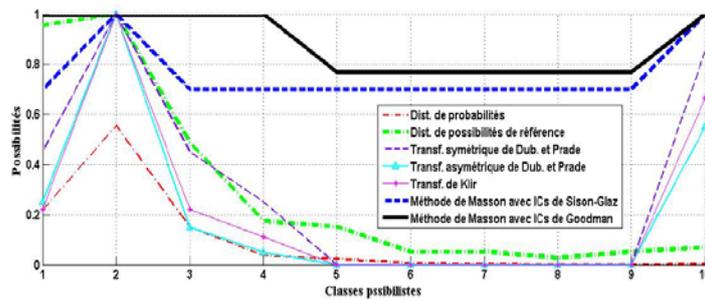


(a) Pour la base synthétique

Figure 2.4 Illustration des distributions de possibilités inférées d'une distribution de probabilités multinomiale, par les cinq approches testées pour des échantillons de grande taille (N=50)

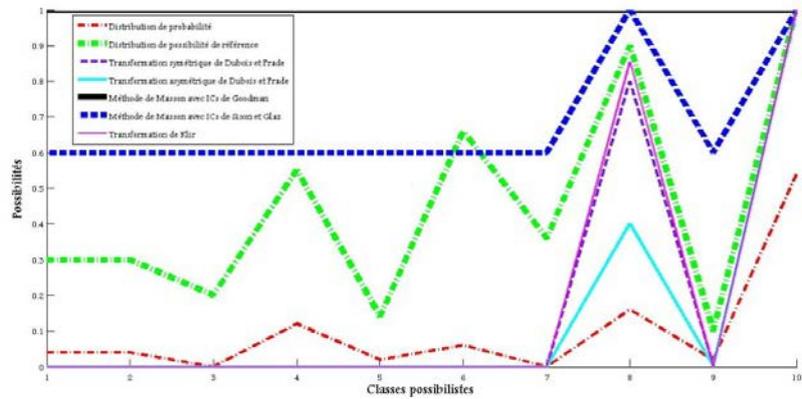


(a) Pour la base synthétique

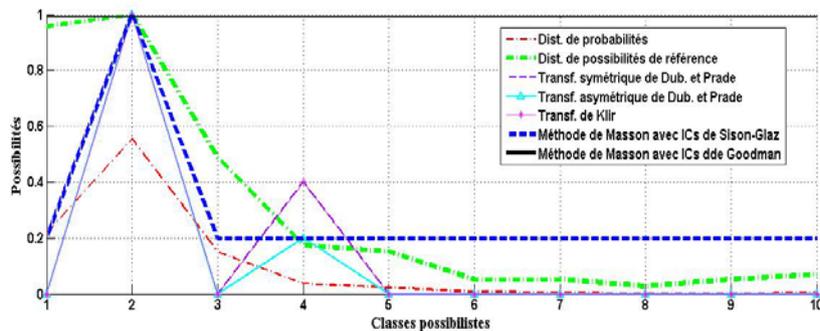


(b) Pour la base 'SVM Guide'

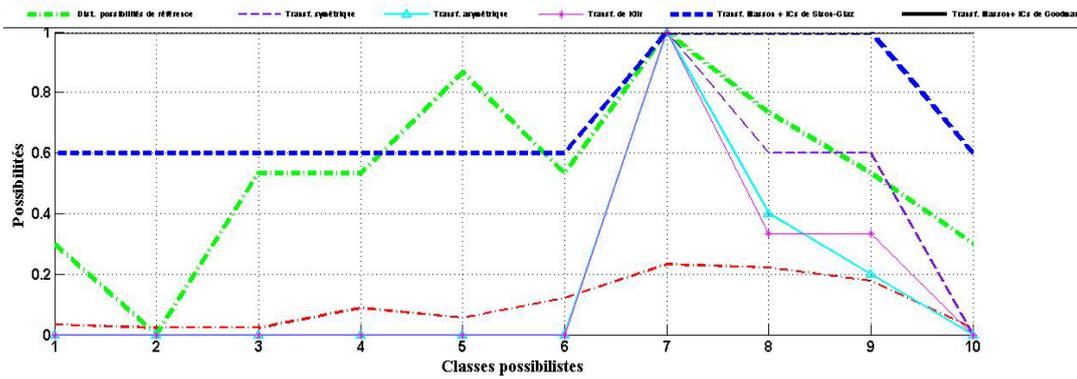
Figure 2.5 Illustration des distributions de possibilités inférées d'une distribution de probabilités multinomiale, par les cinq approches testées pour des échantillons de taille moyenne ($N=20$)



(a) Pour la base synthétique



(b) Pour la base 'SVM Guide'



(c) Pour la base de signaux ultrasonores

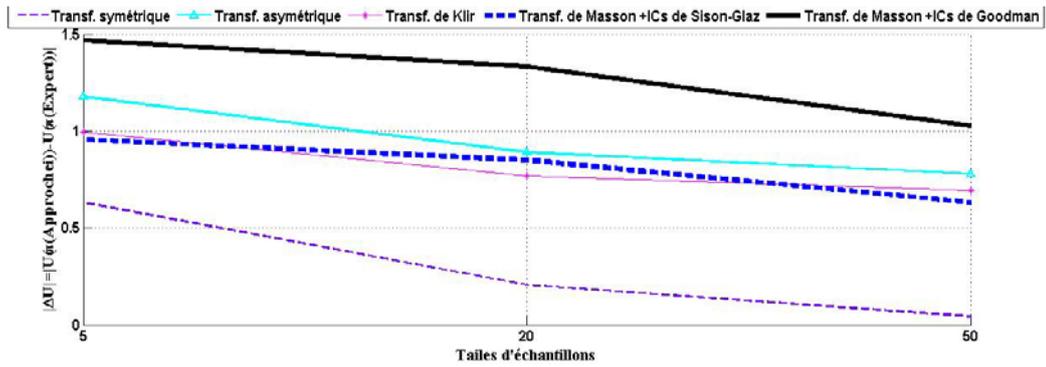
Figure 2.6 Illustration des distributions de possibilités inférées d'une distribution de probabilités multinomiale, par les cinq approches testées pour des échantillons de petite taille moyenne ($N=5$)

Selon nos observations détaillées dans [Cha 2014], nous constatons que les approches basées sur des valeurs de probabilités exactes, possèdent, pour une même taille d'échantillons, des comportements variables en fonction de la variation de supports de données. Cependant, une stabilité relative est marquée par les méthodes basées sur les Intervalles de confiances probabilistes. En effet, appliquées à différentes bases de référence, elles permettent de tirer pratiquement les mêmes conclusions pour un même nombre d'échantillons N . Ceci montre leurs potentiels lors d'un processus d'échantillonnage. Nous constatons que la méthode de Masson basée sur les Intervalles de confiances de Goodman perd progressivement sa sélectivité en diminuant le nombre d'échantillon N . En effet, pour $N=5$, les distributions de possibilités générées présentent un état d'ignorance totale. Par contre la modification proposée utilisant les Intervalles de confiances de Sison-Glaz, génère des distributions de possibilités avec une qualité bien distinguée pour différentes tailles d'échantillons, particulièrement pour un nombre d'échantillon réduit.

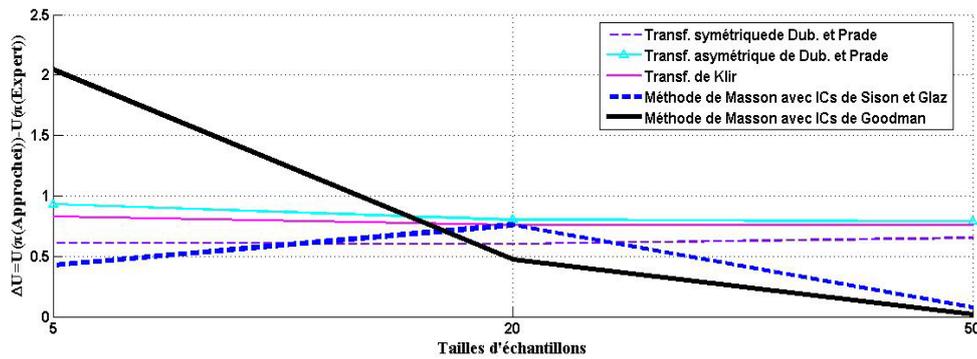
b) Evaluation quantitative

Afin d'évaluer quantitativement la fiabilité des différentes approches, en fonction de la taille des échantillons, nous mesurons: la différence d'incertitude ΔU , le $PSNR$ et la similarité $InfoAff$ dans chaque situation (approche/taille d'échantillon). Pour que les jugements portés aux résultats ne soient pas subjectifs à une prise particulière d'échantillon, nous optons pour une validation croisée qui consiste à varier l'échantillon pour chaque situation (approche/taille d'échantillon) tout en respectant sa taille. On retient ensuite les valeurs moyennes des critères d'évaluation calculés. Cette procédure est appliquée pour chacune des bases de données utilisées, notamment : la base synthétique, la base 'SVM Guide' et la base de signaux ultrasonores.

Les comportements de l'écart d'incertitude, le $PSNR$ et la mesure de similarité $InfoAff$ pour les différentes approches sont respectivement illustrés par les figures Figure 2.7, Figure 2.9 et Figure 2.11 pour les deux bases, synthétique et SVM Guide, et ce, pour les trois gabarits d'échantillons et puis par les figure Figure 2.8, Figure 2.10 et Figure 2.12 pour la base des données ultrason et ce, uniquement pour le petit gabarit d'échantillon.



(a) Pour la base synthétique



(b) Pour la base 'SVM Guide'

Figure 2.7 Ecart d'incertitude ΔU , entre les distributions de possibilités inférées par les cinq approches testées et la distribution de référence, en fonction de la taille de l'échantillon, pour (a) la base synthétique, (b) la base 'SVM Guide'

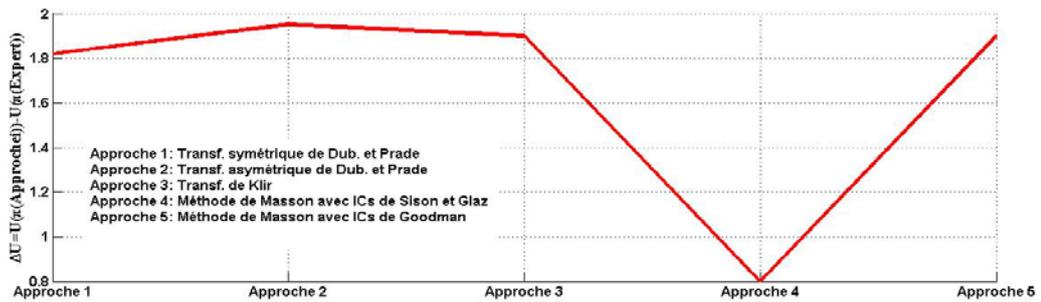
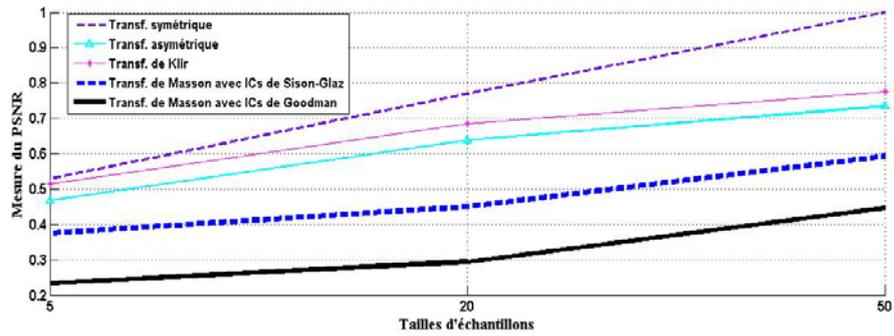
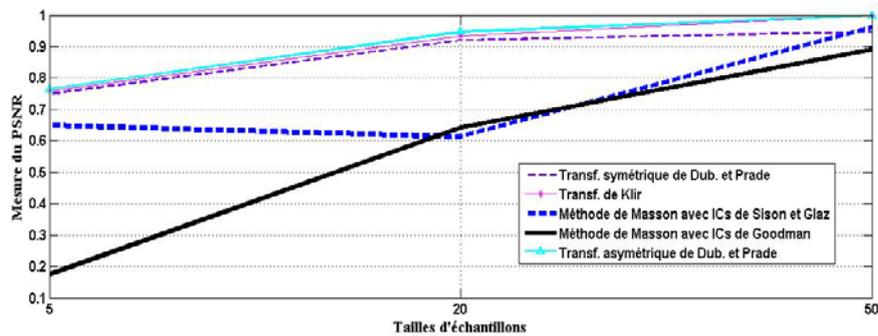


Figure 2.8 Ecart d'incertitude ΔU , entre les distributions de possibilités inférées par les cinq approches testées et la distribution de référence pour le petit gabarit d'échantillons ($N=5$) pris sur la base des données ultrasonores



(a) Pour la base synthétique



(b) Pour la base 'SVM Guide'

Figure 2.9 Rapport du (PSNR), entre les distributions de possibilités inférées par les cinq approches testées et la distribution de référence, en fonction de la taille de l'échantillon, pour (a) la base synthétique, (b) la base 'SVM Guide'

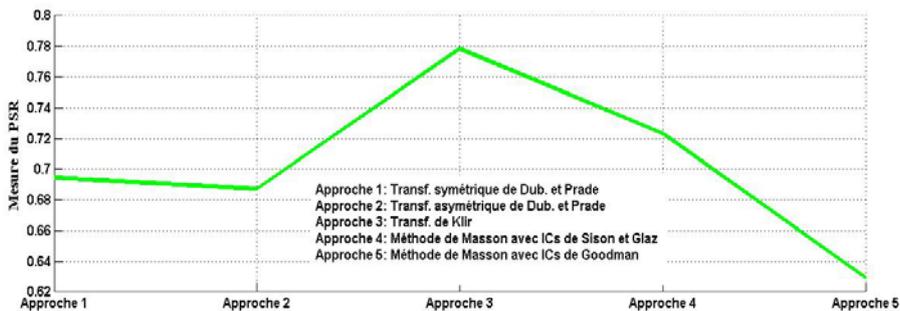
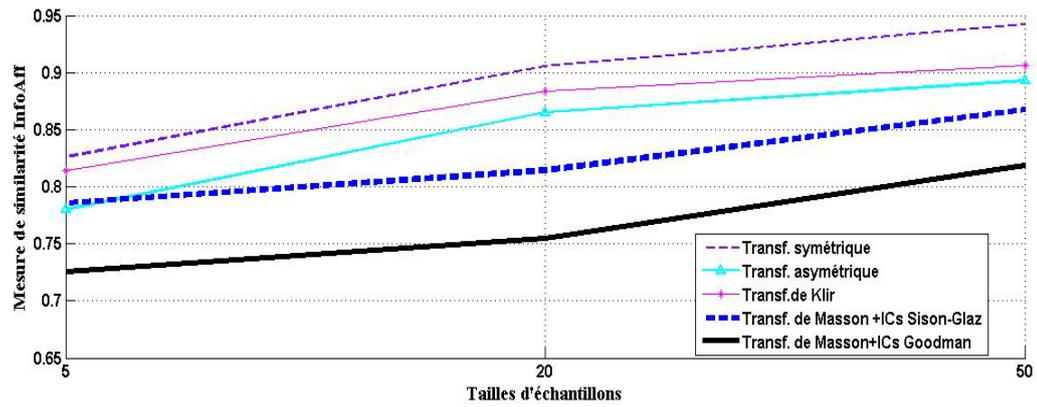
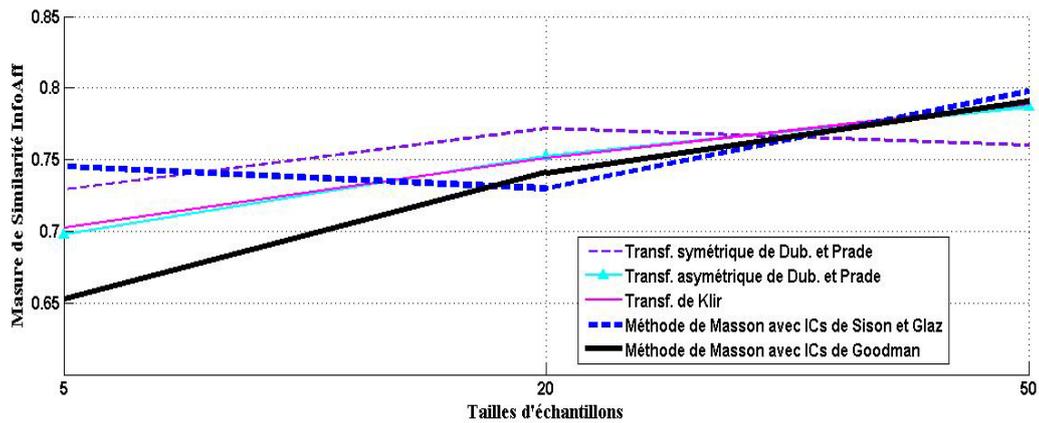


Figure 2.10 Rapport du (PSNR), entre les distributions de possibilités inférées par les cinq approches testées et la distribution de référence pour le petit gabarit d'échantillons ($N=5$)



(a) Pour la base synthétique



(b) Pour la base 'SVM Guide'

Figure 2.11 Mesure de similarité (InfoAff), entre les distributions de possibilités inférées par les cinq approches testées et la distribution de référence, en fonction de la taille de l'échantillon, pour (a) la base synthétique, (b) la base 'SVM Guide'

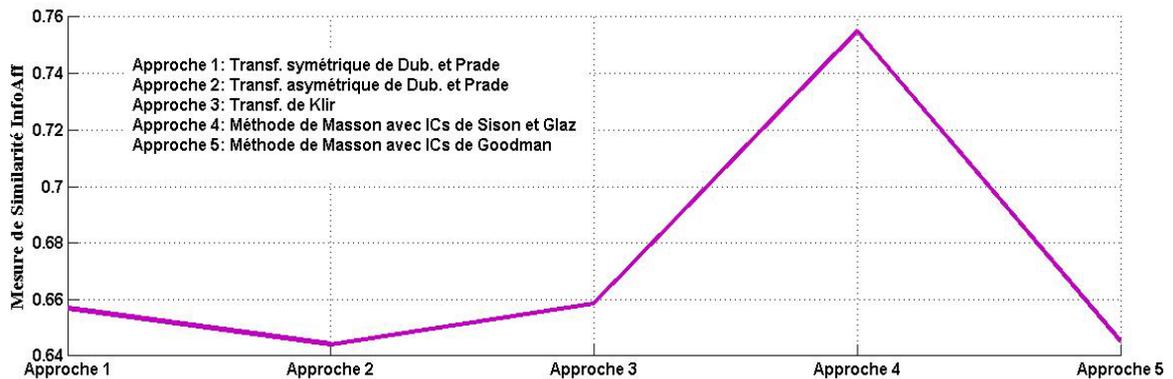


Figure 2.12 Mesure de similarité (InfoAff), entre les distributions de possibilités inférées par les cinq approches testées et la distribution de référence pour le petit gabarit d'échantillons ($N=5$)

c) Discussions

D'après notre évaluation quantifiée, on remarque que les résultats fournis par chaque approche dépendent du support de données qu'on teste. Une différence de résultats est prononcée entre la base synthétique et les autres bases de données qu'on a testées. Ceci peut être dû à la manière même dont on a construit notre base de données. On rappelle en fait qu'on est parti d'une distribution de possibilités de référence, posée arbitrairement, pour en tirer une distribution de probabilités par la transformée inverse de Dubois et Prade symétrique. Toutes les approches sont ensuite appliquées sur des distributions multinomiales déduites de la distribution de probabilités ainsi construite.

C'est ce qui justifie que, pour la base synthétique, la transformation symétrique fournit des distributions de possibilités les plus voisines à la distribution de possibilités de l'expert quel que soit le critère d'évaluation pour les diverses tailles d'échantillons.

Cependant, pour la base 'SVM Guide' et la base des signaux ultrasonores, on aboutit à des résultats attendus selon notre étude théorique.

✓ *Approches basées sur des probabilités simples*

Les transformations de Dubois et Prade et celle de Klir, donnent des distributions de possibilités semblables selon les différents critères pour la base 'SVM Guide', ainsi que des valeurs assez proches, pour la base de signaux ultrasonores, sauf pour le PSNR qui a affiché un meilleur degré pour la transformation de Klir, un degré bien loin des autres obtenus par les transformations de Dubois et Prade.

✓ *Approches basées sur des probabilités en intervalle*

Pour les échantillons de petit gabarit (N=5), l'approche de Masson et Denoeux, basée sur les Intervalles de confiances de Goodman, donne les pires résultats selon les trois critères d'évaluation quantitatives adoptés et ce pour les deux bases de données. Mais en augmentant la taille d'échantillon, les résultats des différents critères s'améliorent beaucoup surtout pour N=50. Ce résultat prouve que les ICs de Goodman perdent leur flexibilité pour des tailles de données inférieures à 50.

Pour les échantillons de petit gabarit (N=5), et pour les deux bases SVM Guide et données ultrasons, l'approche proposée fournit le meilleur degré de similarité et un écart d'incertitude le plus réduit par rapport à la distribution de possibilités de l'expert. Il est vrai qu'elle ne fournit pas le meilleur degré du PSNR, mais quand même un degré proche de celui atteint par la transformation asymétrique pour la base 'SVM Guide' et la transformation de Klir pour la base de signaux ultrasonores.

✓ *Récapitulation pour les échantillons de petit gabarit*

Etant donné que l'objectif principal de notre étude est de conclure sur l'approche la plus adaptée à des petits échantillons, on présente dans le Tableau 2.1, une synthèse des performances réalisées par chaque approche pour la taille N=5. Les scores donnés sont des moyennes des notes attribuées pour chaque approche, sur les trois bases. La note est entre 1 et 5, donnant l'ordre de classement de l'approche selon le critère considéré.

Tableau 2.1 Synthèse des performances d'approches d'estimation de distributions de possibilités pour 5 échantillons

Approches Critères	Transformation symétrique		Transformation asymétrique		Transformation de Klir		Méthode de Masson avec les ICs de Goodman		Méthode de Masson avec les ICs de Sison-Glaz	
	Rang	Score	Rang	Score	Rang	Score	Rang	Score	Rang	Score
Mesure de similarité <i>Info.Aff</i>	Rang	2	Rang	4	Rang	3	Rang	5	Rang	1
	Score	4	Score	2	Score	3	Score	1	Score	5
Mesure de PSNR	Rang	2	Rang	3	Rang	1	Rang	5	Rang	4
	Score	4	Score	3	Score	5	Score	1	Score	2
Ecart d'incertitude ΔU	Rang	2	Rang	4	Rang	3	Rang	5	Rang	1
	Score	4	Score	2	Score	3	Score	1	Score	5
Cohérence	Rang	4	Rang	5	Rang	3	Rang	2	Rang	1
	Score	2	Score	1	Score	3	Score	4	Score	5
Préservation de forme	Rang	2	Rang	4	Rang	1	Rang	5	Rang	3
	Score	4	Score	2	Score	5	Score	1	Score	3
SCORE TOTAL	18		10		19		8		20	

Il résulte de ce bilan que sur les cinq approches étudiées, l'approche proposée ainsi que la transformation de Klir sont les plus appropriées pour modéliser une population à partir d'un échantillon de petite taille.

IV. Modélisation possibiliste de connaissances contextuelles de voisinage

Une scène est composée de portions (morceaux, régions, parties, segments, sections, zones, ...) homogènes et de bords (bordures, contours, limites, ...) faisant percevoir des parcelles, correspondant à des objets sémantiques, à différents niveaux de disposition. L'homogénéité caractérisant une portion se rapporte à une certaine harmonie (ou voisinage) des mesures d'attributs portant sur une granularité sémantique inférieure à celle des portions: on parle des composants élémentaires constituant cette portion. De ce fait, la façon d'exprimer les connaissances liées à un composant élémentaire est d'un impact majeur sur son étiquetage (ou classification). Différentes approches de représentation des connaissances élémentaires, existent. Nous proposons dans ce qui suit, deux approches de modélisation élémentaires possibilistes l'une ponctuelle et l'autre contextuelle.

L'évaluation de nos approches est effectuée sur l'image, étant une certaine vision d'une scène obtenue par l'intermédiaire d'un capteur donné. Le composant élémentaire est alors un pixel et les portions sont les régions de l'image. Les approches que nous proposons correspondent à des modélisations pixelliques que nous comparons à d'autres techniques soigneusement sélectionnées de la littérature. Leur efficacité est jugée vis-à-vis de la classification d'un pixel dans une classe thématique donnée.

1. Modes de représentations de connaissances pixelliques

Une classification pixellique est un problème de prise de décision vis-à-vis de l'attribution du pixel à une classe thématique. La qualité de cette décision dépend de la façon de modélisation des connaissances associées au pixel. La modélisation est liée :

- à la manière de voir le pixel dans une image : en effet le pixel peut être considéré isolement de son voisinage et là on parle d'une modélisation ponctuelle, ou il peut être considéré en tant qu'un membre d'une communauté et là on parle d'une modélisation contextuelle. (isolé ou dans son voisinage)
- au mode de représentation des connaissances qui lui sont associées. En effet, on distingue quatre niveaux de connaissances pixelliques, qu'on peut placer sur une échelle de sémantique croissante comme suit :
 - (i) niveau de données brutes
 - (ii) niveau intermédiaire des primitives
 - (iii) niveau sub-sémantique et
 - (iv) niveau symbolique.

La connaissance liée au pixel au niveau symbolique est déjà explicite, étant une étiquette classe. Alors que les connaissances liées au pixel, dans les trois autres niveaux, sont implicites et nécessitent un traitement afin d'accorder une étiquette thématique au pixel auquel elles sont associées. Plusieurs scénarios peuvent être envisagés afin de parvenir à une classification pixellique, partant de l'un des trois niveaux de connaissances pixelliques. La Figure 2.13, illustre les différents scénarios possibles et envisageables pour passer des données radiométriques brutes issues du capteur, ou de l'un des deux autres niveaux supérieurs, vers la connaissance du plus haut niveau sémantique, à savoir « l'étiquette ».

On se propose de comparer trois modes, correspondant respectivement aux niveaux (i), (ii) et (iii) de représentation des connaissances liées à un pixel, en vue de sa classification. On s'intéresse particulièrement à la modélisation pixellique au niveau sub-sémantique. A ce niveau, plusieurs théories, telles que celles des probabilités ou des évidences, offrent des espaces de projection de connaissances qui permettent de recueillir des résultats élégants mais à condition d'injecter des connaissances *a priori* concernant les modèles optimaux des classes. Notre intérêt s'oriente alors vers la théorie des possibilités, qui permet de modéliser les faibles connaissances connues sous le nom de « ill-defined knowledge ». Nous proposons deux approches possibilistes de modélisation de connaissances pixelliques, l'une ponctuelle et l'autre contextuelle. Des tests d'évaluation sont effectués afin de prouver la supériorité de la théorie des possibilités vis-à-vis de la représentation des connaissances pixelliques dans une image à qualité dégradée.

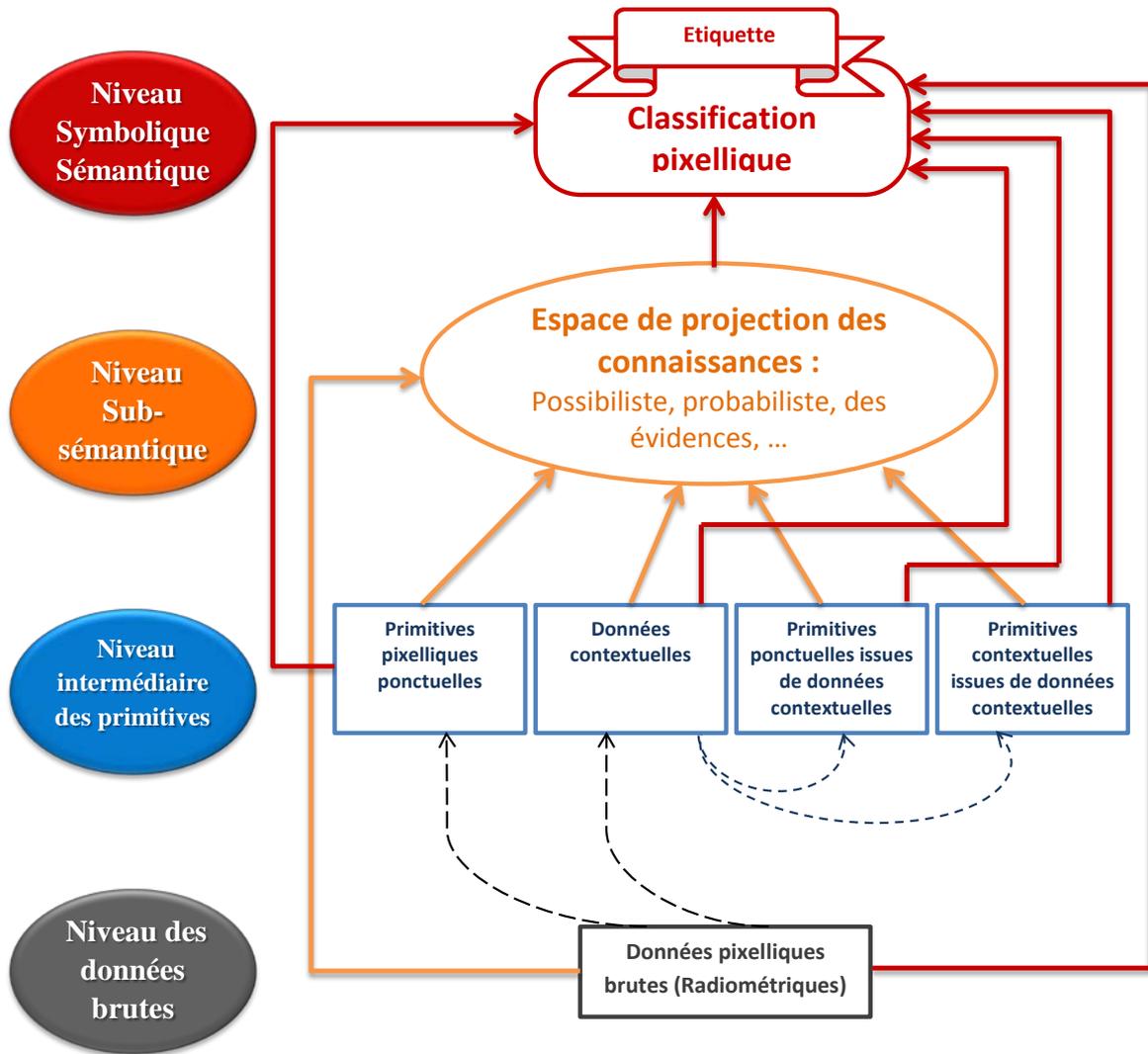


Figure 2.13 Les différents scénarios pour la classification pixellique.

2. 1^{ère} Contribution : Modélisations possibilistes de connaissances pixelliques

a) Définition du contexte

L'attribution d'une étiquette à un pixel candidat est un problème de prise de décision qui peut être formulé de la manière suivante :

On dispose de :

- ✓ Une image Im de $N \times M$ pixels ou objets :

$$Im = \{Im(i, j); i = 1..N \text{ et } j = 1..M\}$$

- ✓ Un univers de K classes :

$\Omega = \{C^1, C^2, \dots, C^K\}$, de telle sorte que chaque objet ou pixel (i, j) doit appartenir à une unique classe de l'univers Ω .

- ✓ Une mesure de similarité qui exprime l'homogénéité régnant au sein d'une classe relativement à l'un ou plusieurs des attributs de l'ensemble des T attributs : $A = \{A_1, A_2, \dots, A_T\}$.

- ✓ En considérant le cadre possibilistes, les connaissances liées à des classes sont directement données sous la forme de distributions de possibilités définies sur l'espace des primitives permettant la description des classes. Ainsi une classe C^k est modélisée par les connaissances possibilistes suivantes: $\pi_{C^k}^{A_t}$, ($t = 1..T$), qu'on notera π_k^t pour simplifier. Ainsi, l'information sub-sémantique utilisée est une distribution de possibilités (sur l'ensemble des classes) et ceci pour chaque attribut.

b) Problématique

L'attribution d'un pixel (i,j) à une classe C^k , se décide selon le degré de similarité qui existe entre eux, ou plus précisément, entre les connaissances qui les caractérisent. Cependant, un pixel et une classe présentent deux granularités différentes, ce qui fait que l'évaluation de la similarité entre ces deux structures devrait prendre en considération ce détail. Cette prise en compte se fait dès la modélisation des connaissances qui caractérisent chaque structure. La distribution de possibilités qu'on propose d'adopter en tant que modèle sub-sémantique est bien appropriée pour modéliser une classe désignant l'information symbolique, ce qui n'est pas très commode pour un pixel. En effet, celui-ci a très couramment été représenté dans l'un des deux niveaux inférieurs, à savoir le niveau numérique et intermédiaire, mais rarement au niveau sub-sémantique. Ceci nous oriente vers la considération du pixel soit ponctuellement soit contextuellement.

Dans ce sens, on considère deux hypothèses pour représenter la connaissance qui caractérise un pixel:

Hypothèse 1 : Les connaissances qui présentent le pixel résumant ses informations ponctuelles. Le pixel est ainsi vu comme un singleton et donc la relation Pixel/Classe est vue comme une relation Singleton/Modèle.

Hypothèse 2 : L'information contextuelle spatiale du pixel est utilisée pour construire un modèle local pour ce pixel. Le pixel est ainsi considéré dans son voisinage et est présenté lui-même par un modèle et donc la relation Pixel/Classe est vue comme une relation Modèle/Modèle

Dans ce qui suit, on propose de détailler pour chacune de ces deux hypothèses, la représentation de la connaissance au niveau du pixel, puis de présenter l'évaluation de la similarité entre le pixel, ainsi modélisé, et une classe.

c) Hypothèse 1 : Modélisation possibiliste ponctuelle du pixel

Le pixel est considéré en tant qu'un singleton isolé de son environnement (i.e. absence d'information spatiale) et dans ce cas la similarité est à estimer entre :

- ✓ un singleton (i,j) → représenté par ses valeurs d'attributs $Im(i,j)$:

$$Im(i,j) = \begin{pmatrix} a_1(i,j) \\ \vdots \\ a_t(i,j) \\ \vdots \\ a_T(i,j) \end{pmatrix} \quad \text{Eq2.4}$$

- ✓ et une classe $C^k \rightarrow$ représentée ou modélisée par une distribution de possibilités π_k^t selon chacun des attributs A_t . Un attribut est considéré comme un observateur ou une source de connaissances pour la classe.

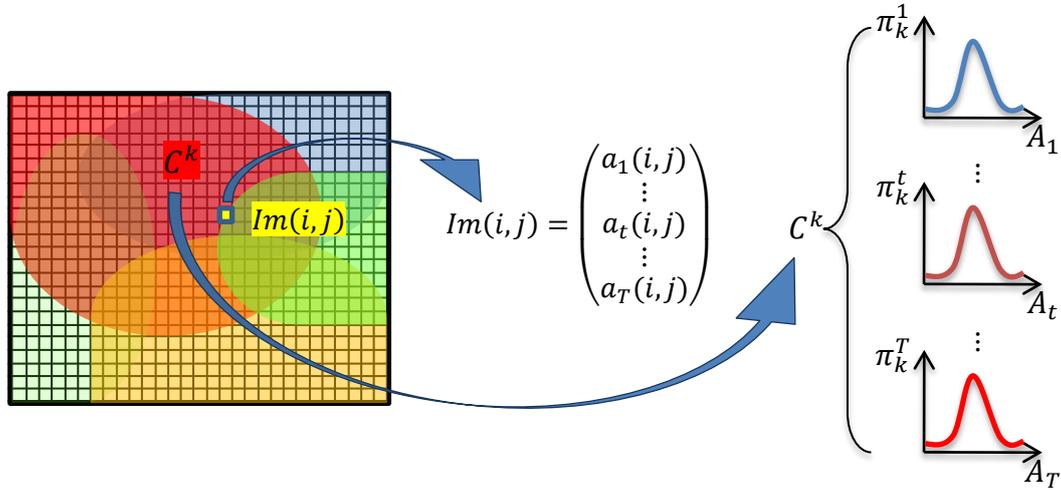


Figure 2.14 Modélisation du pixel par ses valeurs d'attributs $Im(i,j)$, et de la classe par ses distributions de possibilités conditionnelles aux attributs considérés

A ce niveau et tel que représenté par la Figure 2.14, on peut voir que :

Les connaissances liées aux classes thématiques se situent au niveau sub-sémantique du fait qu'elles sont données sous la forme de distributions de possibilités, alors que la représentation des connaissances au niveau du pixel se situe au niveau intermédiaire des primitives, étant donné qu'il est présenté par des mesures d'attributs. L'idée est de faire remonter ces connaissances pixeliques sur l'échelle sémantique en les projetant dans l'espace des distributions de possibilités des classes thématiques (Cf. Figure 2.15). Dans ce cas, la connaissance qui caractérise le pixel, est une matrice donnant pour le pixel sa possibilité d'appartenir à chacune des classes C^k de l'univers Ω et ceci selon chaque attribut t :

$$\begin{aligned} \text{modèle}(i,j) &= \pi_k^t(Im(i,j)) \\ &= \begin{pmatrix} \pi_1^1(Im(i,j)) & \dots & \pi_1^T(Im(i,j)) \\ \vdots & \ddots & \vdots \\ \pi_k^1(Im(i,j)) & \dots & \pi_k^T(Im(i,j)) \end{pmatrix} \end{aligned} \quad \text{Eq2.5}$$

Cette matrice représente les connaissances pixeliques au niveau sub-sémantique. Par ailleurs, chaque ligne de cette matrice exprime la relation $(i,j)/C^k$. A ce niveau, on peut poser une hypothèse/définition supposant que la similarité du pixel avec une classe (Singleton/Modèle) peut être calculée comme étant la possibilité pour que l'attribut mesuré en ce pixel, soit de cette classe :

$$\text{Sim}[(i,j), C^k] = \text{Sim}[Im(i,j), \pi_k^t] = \pi_k^t(Im(i,j)), \text{ pour } t = 1..T \quad \text{Eq2.6}$$

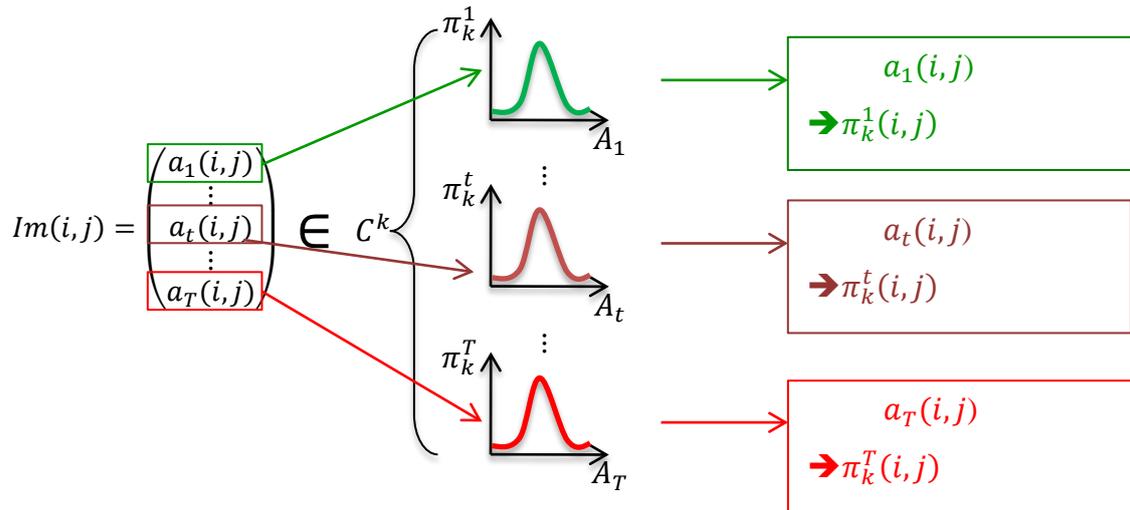


Figure 2.15 Représentation de la connaissance pixellique d'un pixel par un degré de possibilité d'appartenir à une classe

La décision d'attribution à l'une des K classes, est assurée selon la valeur de possibilité maximale.

d) Hypothèse 2 : Modélisation possibiliste contextuelle du pixel

Dans l'hypothèse 1, nous mesurons pour le pixel ses valeurs d'attributs qui ne permettent que de calculer des valeurs de possibilités. En revanche, pour cette deuxième hypothèse, nous considérons le pixel dans son contexte de voisinage et nous utilisons son information spatiale pour lui estimer une distribution de possibilités locale $\pi_{(i,j)}^t$ (dans l'espace de chaque attribut tel qu'illustré par la Figure 2.16.

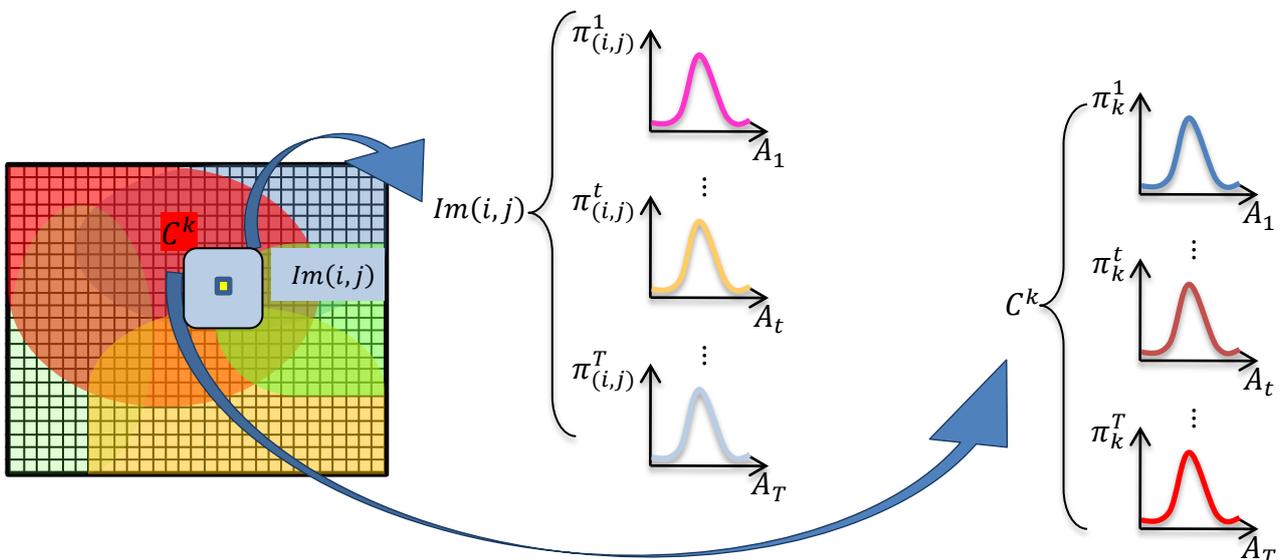


Figure 2.16 La modélisation du pixel et celle de la classe sont assurées par leurs distributions de possibilités conditionnelles aux attributs considérés

Ainsi, les connaissances pixelliques sont représentées au niveau sub-sémantique par une connaissance prenant en considération l'appartenance du pixel à son voisinage spatial. Par conséquent, le calcul de la similarité entre le pixel et les différentes classes peut se faire en

calculant la similarité entre leurs distributions de possibilités respectives. L'estimation de la similarité se fait dans ce cas entre deux connaissances de même nature, à savoir la distribution de possibilités $\pi_{(i,j)}^t$ modélisant le pixel et la distribution de possibilités π_k^t modélisant C^k :

$$Sim[(i,j), C^k] = Sim[\pi_{(i,j)}^t, \pi_k^t] \quad \text{Eq2.1}$$

3. Résultats et discussion

On propose d'évaluer les deux approches de modélisations possibilistes du pixel précédemment proposées au niveau sub-sémantique, par une comparaison de leurs performances à celles d'autres approches de l'état de l'art.

Ces approches sont soigneusement sélectionnées de façon à assurer une comparaison selon deux notions : le niveau sémantique de représentation des connaissances et l'aspect ponctuel/contextuel de l'information employée pour représenter le pixel (cf. Tableau 2.2).

a) Contexte des simulations:

Les approches de modélisation pixelliques à évaluer et comparer, sont :

- Approche 1. Modélisation par le niveau radiométrique suivie d'une classification par seuillage d'histogramme par la technique d'Otsu
- Approche 2. Modélisation par les distances des NG par rapport aux centroïdes des classes, suivie d'une classification par la technique des K-means
- Approche 3. Modélisation bayésienne du pixel, suivie d'une classification basée sur les champs de Markov avec une régularisation assurée par l'algorithme ICM (iterated conditional modes)
- Approche 4. Modélisation bayésienne du pixel, suivie d'une classification basée sur les champs de Markov avec une régularisation assurée par l'algorithme RS (Recuit Simulé)
- Approche 5. Modélisation bayésienne classique
- Approche 6. Modélisation possibiliste ponctuelle selon la 1ère hypothèse: singleton/modèle
- Approche 7. Modélisation possibiliste contextuelle selon la 2^{ème} hypothèse: modèle/modèle

Tableau 2.2 Tableau Caractérisation des approches évaluées selon leur niveau sémantique de représentation et leur aspect ponctuel/contextuel de la formation de l'information pixellique

	Modélisation ponctuelle	Modélisation contextuelle
Niveau Radiométrique	Approche 1	
Niveau Intermédiaire	Approche 2	
Niveau Sub-sémantique	Approches 5 et 6	Approches 3, 4 et 7

Afin d'évaluer ces approches on s'est positionné dans le contexte suivant :

- ✓ Deux types d'images ont été expérimentés : une image de synthèse de façon à considérer différents taux de contrastes et de bruit et trois images médicales optiques avec leurs vérités terrains.
- ✓ L'univers Ω est composé de deux classes : classe « objet » et classe « arrière-plan »
- ✓ L'homogénéité régnant au sein d'une classe est relative à un seul attribut, à savoir le niveau de gris du pixel.

- ✓ L'évaluation de la classification est assurée par le calcul de la distance de Hamming entre les résultats de classifications pixelliques et les images de vérités terrains. Une distance nulle signifie une identité parfaite entre les deux images. Une distance égale à l'unité, signifie une complète différence entre les deux images. Ainsi, moins est la distance, meilleure est la classification. Une matrice de confusion est également employée pour évaluer les 7 approches sur les images optiques médicales.
- ✓ Pour les approches possibilistes :
 - les distributions de possibilités estimées pour les classes thématiques «objet» et «arrière-plan» étaient de la forme trapézoïdale (Cf. Figure 2.17), pour l'image de synthèse, et estimées par prises d'échantillons pour les images optiques selon la technique de Alshwa *et al.* [AAG 2013, ASA 2016]
 - la distribution de possibilités locale estimée pour un pixel est de la forme triangulaire, centrée en la moyenne des niveaux de gris de son voisinage et ouverte de la valeur de l'écart type de son voisinage (Cf. Figure 2.18).

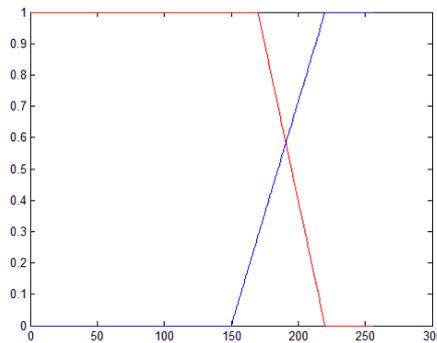


Figure 2.17 Les distributions de possibilités estimées pour les classes thématiques « objet » (en bleu) et « arrière-plan » (en rouge)

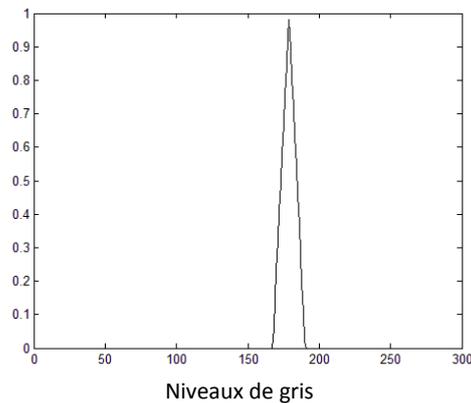


Figure 2.18 La distribution de possibilités estimée pour un pixel

b) Tests et évaluation sur une image de synthèse:

Une «image test», illustrée par la Figure 2.19, est définie de sorte que des indicateurs de qualité sont connus *a priori*. Cette image est formée d'une imagerie de base comportant quatre disques circulaires représentant un objet d'intérêt de différentes tailles.

L'utilisation de plusieurs tailles constitue un premier indicateur permettant d'apporter une première appréciation sur la capacité d'une approche à «localiser/détecter» des objets d'intérêt de différentes tailles.

Sur l'axe horizontal, l'imagerie de base est dupliquée en faisant croître le contraste des objets par rapport au fond de l'image. Cinq niveaux de contraste sont représentés dans cette image. Le contraste est défini comme étant le rapport entre la différence et la somme des niveaux de gris moyens (Objet vs. Fond). Ces cinq niveaux de contraste sont donnés par : 0.3, 0.43, 0.5, 0.56 et 0.6.

Finalement, l'axe vertical permet d'apprécier l'effet du niveau de bruit affectant l'image. En effet, six niveaux de bruit gaussien additif sont introduits dans cette image d'une façon progressive (correspondant à six valeurs croissantes de l'écart type).

L'évaluation est réalisée d'une manière quantitative par le calcul de l'écart entre la vérité terrain et les images étiquetées par chacune des approches considérées. Cet écart est quantifié par la distance de Hamming.

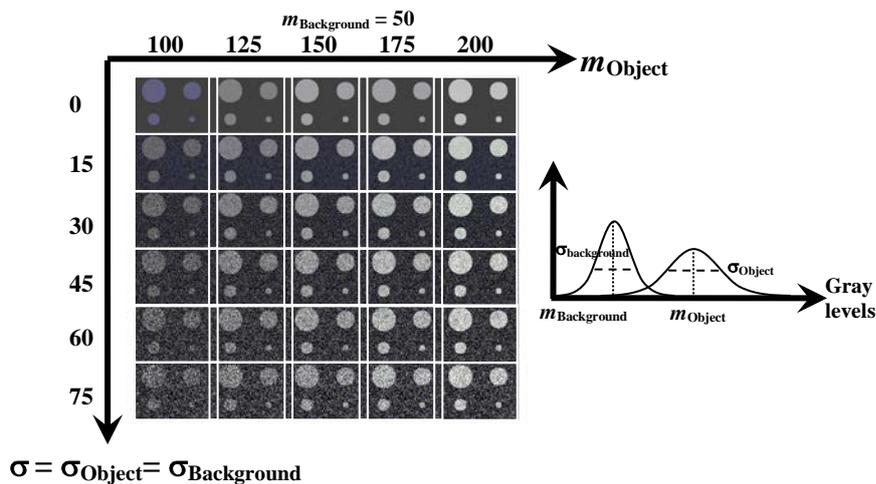


Figure 2.19 Image synthétique utilisée pour les expérimentations

Trois tests sont réalisés sur cette image :

❖ **Test1 : Evaluation selon contraste décroissant**

L'image test est formée par 30 blocs disposés en une matrice de 6 lignes et 5 colonnes. Les blocs de la 5^{ème} colonne, présentent le plus fort contraste de l'image, alors que les blocs de la 1^{ère} colonne présentent le plus faible contraste de l'image (Cf. Figure 2.20-a). Ainsi, en allant sur les colonnes de blocs de la droite vers la gauche, on note une dégradation de la qualité de l'image de point de vue contraste, car plus le contraste est faible, moins on arrive à distinguer les régions de l'image.

La Figure 2.20-b, illustre les distances de Hamming calculées sur des blocs de contraste décroissant (de la 5^{ème} colonne vers la 1^{ère} colonne), entre l'image vérité terrain et les images étiquetées par les 7 approches considérées pour la modélisation pixellique. La Figure 2.20-c illustre l'erreur de segmentation.

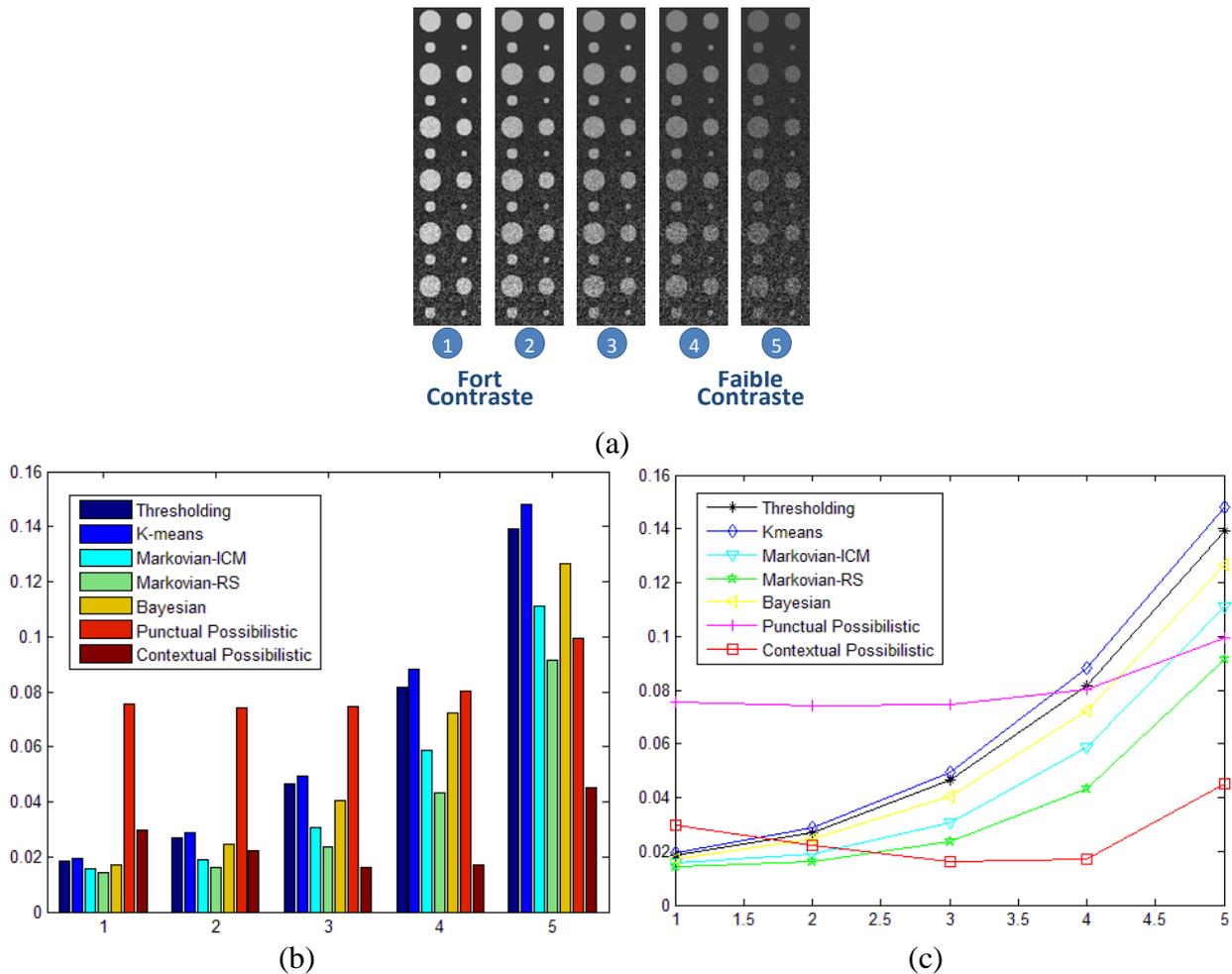


Figure 2.20 Evaluation de la segmentation effectuée sur des blocs de contraste décroissant, entre l'image vérité terrain et les images étiquetées des approches considérées pour la modélisation pixellique :

a- Blocs de contrastes décroissants

b- Distances de Hamming

c- Erreur de reconnaissance

On note que :

- L'approche possibiliste contextuelle réalise de meilleurs scores que l'approche possibiliste ponctuelle, et ce sur tous les niveaux de contraste.
- dans le cas où « tout va bien », c'est-à-dire où les régions sont nettement distinguées grâce à un fort contraste entre l'objet et l'arrière-plan, les 3 approches Bayésiennes présentent les meilleurs scores avec des distances entre les résultats de classification et les images vérité terrain, inférieures à 0.02, alors que les approches possibilistes présentent les distances les plus élevées, de l'ordre de 0.07 pour l'approche ponctuelle et 0.03 pour l'approche contextuelle. Cependant, dans le cas où « tout va mal », c'est-à-dire où les régions sont ambiguës à cause du faible contraste entre l'objet et l'arrière-plan, l'approche possibiliste contextuelle assure les meilleurs résultats avec une distance de l'ordre de 0.04, alors que le rendement des approches Bayésiennes baisse nettement avec des distances de l'ordre de 0.1. On peut en déduire que le modèle possibiliste représente mieux le pixel dans les cas d'ambiguïté.

- Si on observe la dynamique des valeurs de distances calculées pour chacune des approches, on réalise que les approches possibilistes présentent une stabilité remarquable sur leurs valeurs, alors que les autres approches présentent de grands écarts quand l'image se dégrade, tel est l'exemple de l'approche k-means qui donne une distance de l'ordre de 0.02 pour un bon contraste et de l'ordre de 0.15 pour un mauvais contraste.

❖ *Test2 : Evaluation selon bruit croissant*

Les blocs de la 6^{ème} ligne, présentent le plus fort taux de bruit de l'image, alors que les blocs de la 1^{ère} ligne ne présentent aucun bruit (Cf. Figure 2.21-a). Ainsi, en allant sur les lignes de blocs du haut vers le bas, on note une dégradation de la qualité de l'image de point de vue bruit.

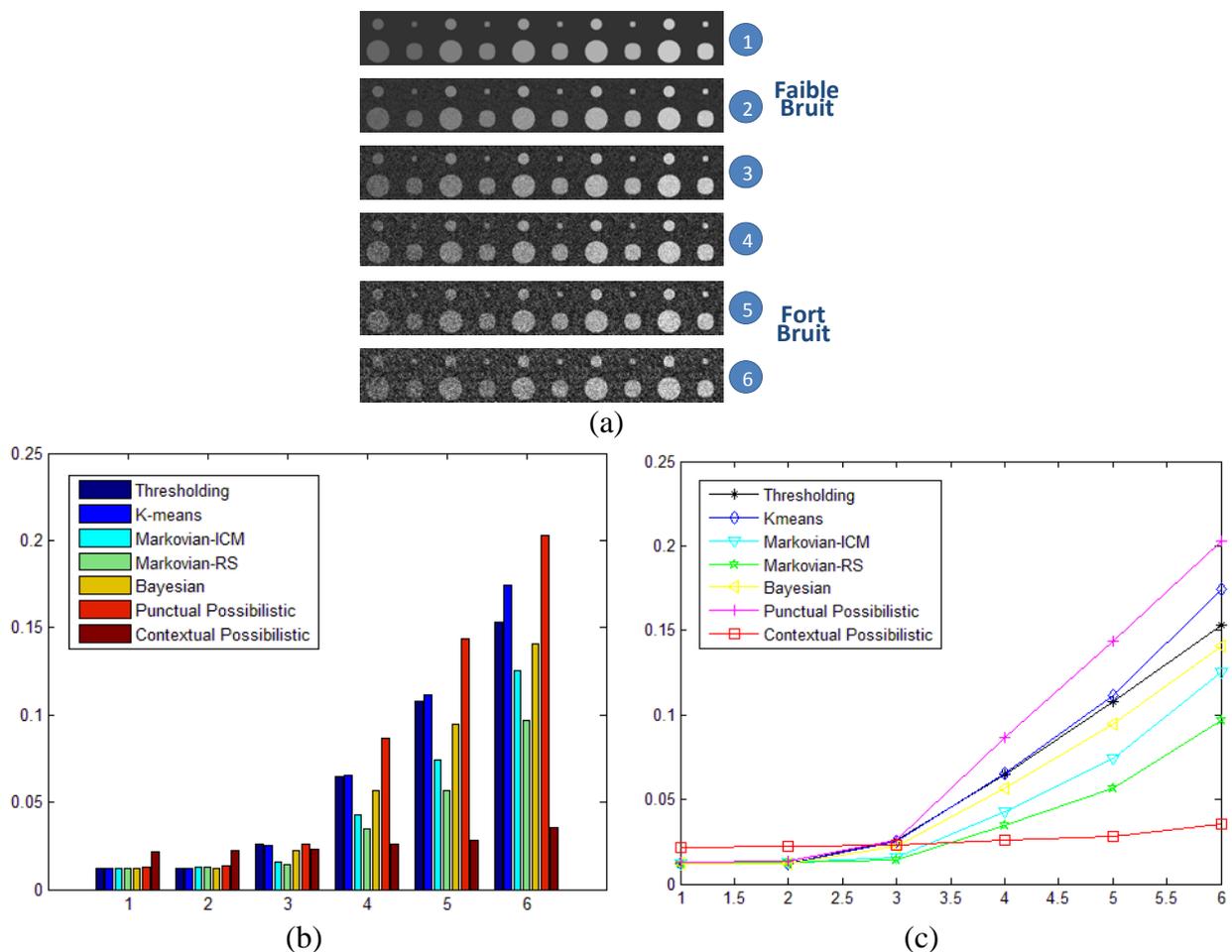


Figure 2.21 Evaluation de la segmentation effectuée sur des blocs de bruit croissant, entre l'image vérité terrain et les images étiquetées des approches considérées pour la modélisation pixelique :

- a- Blocs de bruit croissant
- b- Distances de Hamming
- c- Erreur de reconnaissance

La Figure 2.21-b, illustre les distances de Hamming calculées sur des blocs de bruit croissant (de la 1^{ère} ligne vers la 6^{ème} ligne), entre l'image vérité terrain et les images

étiquetées par les 7 approches considérées pour la modélisation pixellique. La Figure 2.21-c illustre l'erreur de segmentation.

On note que :

- Les résultats de classification sont plus influés par le bruit que par le contraste
- L'approche possibiliste contextuelle garde sa stabilité malgré l'augmentation du taux de bruit. En effet, l'erreur de segmentation pour l'approche possibiliste contextuelle varie légèrement dans l'intervalle [0.02, 0.04], alors que k-means varie dans [0.01, 0.18]

Dans le cas du plus fort taux de bruit, l'approche possibiliste contextuelle réalise le meilleur résultat, avec d'importants écarts avec les autres approches aux mêmes conditions. Cependant l'approche possibiliste ponctuelle réalise le pire résultat, ce qui laisse déduire l'importance de l'information contextuelle dans les mauvaises conditions de définition de l'image.

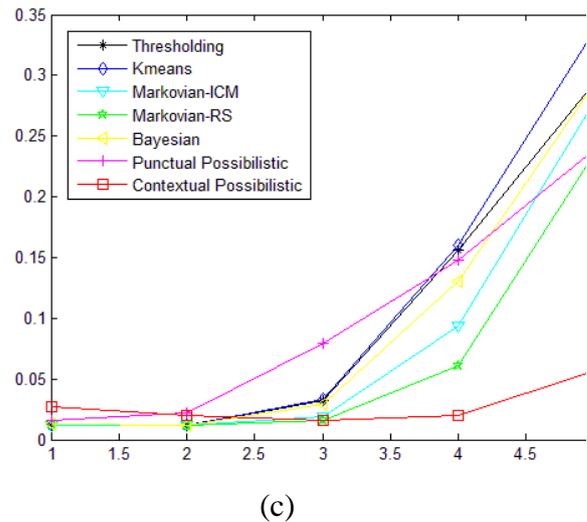
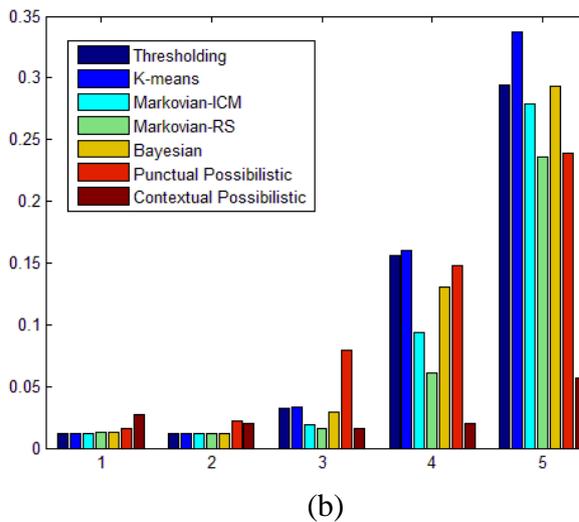
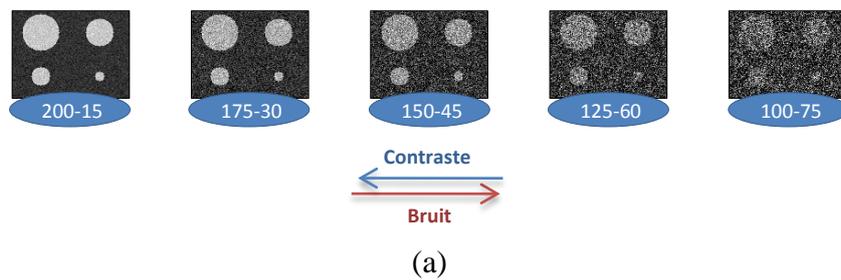


Figure 2.22 Evaluation de la segmentation effectuée sur des blocs de dégradation croissante (contraste décroissant et bruit croissant), entre l'image vérité terrain et les images étiquetées des approches considérées pour la modélisation pixellique :

- a- Blocs de dégradation croissante
- b- Distances de Hamming
- c- Erreur de reconnaissance

❖ Test3 : Evaluation selon dégradation croissante de la qualité (bruit croissant et contraste décroissant)

Pour ce test, la dégradation de l'image est liée à la fois au contraste et au bruit. On réalise l'évaluation par la distance de Hamming sur un balayage diagonal des blocs, allant du bloc (2^{ème} ligne, 5^{ème} colonne) au bloc (6^{ème} ligne, 1^{ère} colonne) (Cf. Figure 2.22-a). Dans ce sens de balayage, la qualité va dans un sens de dégradation croissante.

La Figure 2.22-b, illustre les distances de Hamming calculées sur les blocs ainsi considérés, entre l'image vérité terrain et les images étiquetées par les 7 approches à évaluer pour la modélisation pixellique. La Figure 2.22-c donne l'erreur de segmentation.

On note que par ce test on défie toutes les approches évaluées par une qualité vraiment médiocre de l'image au niveau du bloc (6^{ème} ligne, 1^{ère} colonne). C'est l'ambiguïté totale qui règne et la classe «objet» se trouve fondue dans la classe «arrière-plan». Ce qui explique la nette dégradation des résultats de classification de toutes les approches évaluées (presque le un tiers de l'ensemble des pixels a été mal classifié), excepté l'approche possibiliste contextuelle qui a pu résister. En effet, les observations déjà notées pour les deux tests précédents, concernant la stabilité de la modélisation de la connaissance pixellique par un modèle possibiliste contextuel, et sa nette efficacité par rapport aux autres approches dans les mauvaises conditions de contraste et de bruit, sont de nouveau vérifiées et validées.

c) Tests et évaluation sur des images médicales optiques:

Treize images de lésion de peau (mélanome), dont 3 sont illustrées à titre d'exemple par la Figure 2.23, ont été considérées pour évaluer les sept approches de modélisation pixelliques. Chacune des treize images fait 150×150pixels, soit 22500 pixels à classifier.

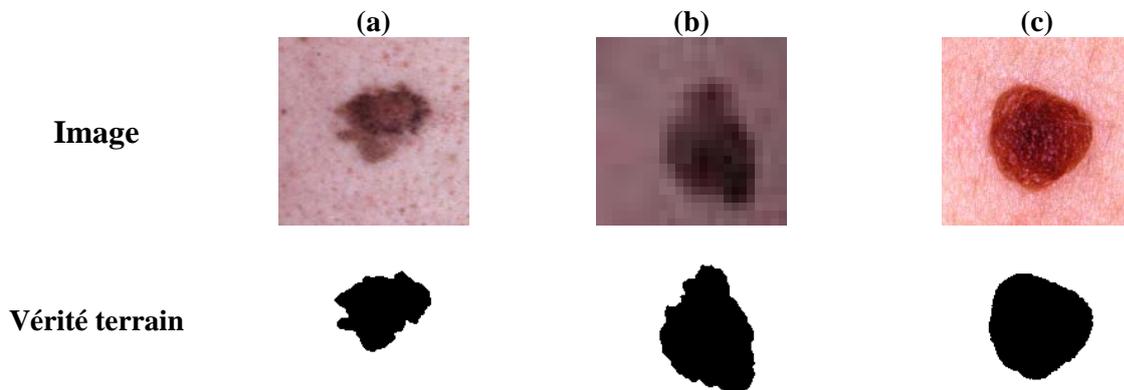


Figure 2.23 Trois images médicales optiques de mélanomes utilisées pour les expérimentations

Précisons, que sur ce genre d'images, une lésion correspond à une région de fort contraste. Cette description, peu précise, est tout à fait caractéristique de la réalité. En effet, les régions d'intérêt peuvent avoir, sur l'image, divers niveaux de contraste selon leur nature, comme la couleur de la peau pour les mélanomes. La grande variabilité des images médicales optiques représente une limitation majeure pour les systèmes d'aide à la détection, à la caractérisation et à la classification des lésions.

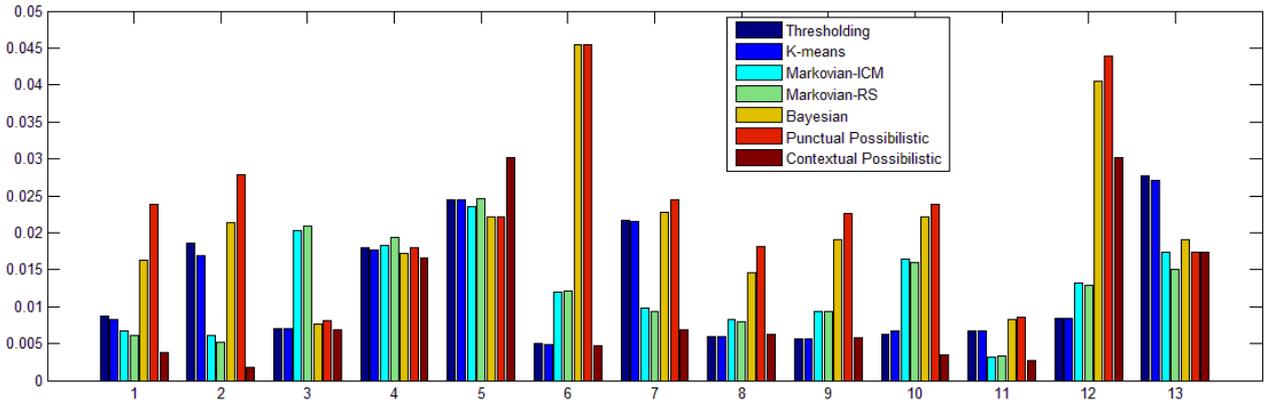


Figure 2.24 Distances de Hamming calculées, entre les 13 images médicales optiques étiquetées et leurs vérités terrain

La Figure 2.24, illustre les distances de Hamming calculées, entre chacune des treize images étiquetées par chacune des 7 approches évaluées pour la modélisation pixellique, et leurs images de vérité terrain correspondantes.

En observant ces résultats, on peut voir que l'approche possibiliste ponctuelle donne des performances très semblables à l'approche Bayésienne classique. C'est complètement prévisible du fait que les deux approches se basent sur des concepts spatiaux semblables et différent juste sur le type du modèle sur lequel se fait la projection des données ponctuelles.

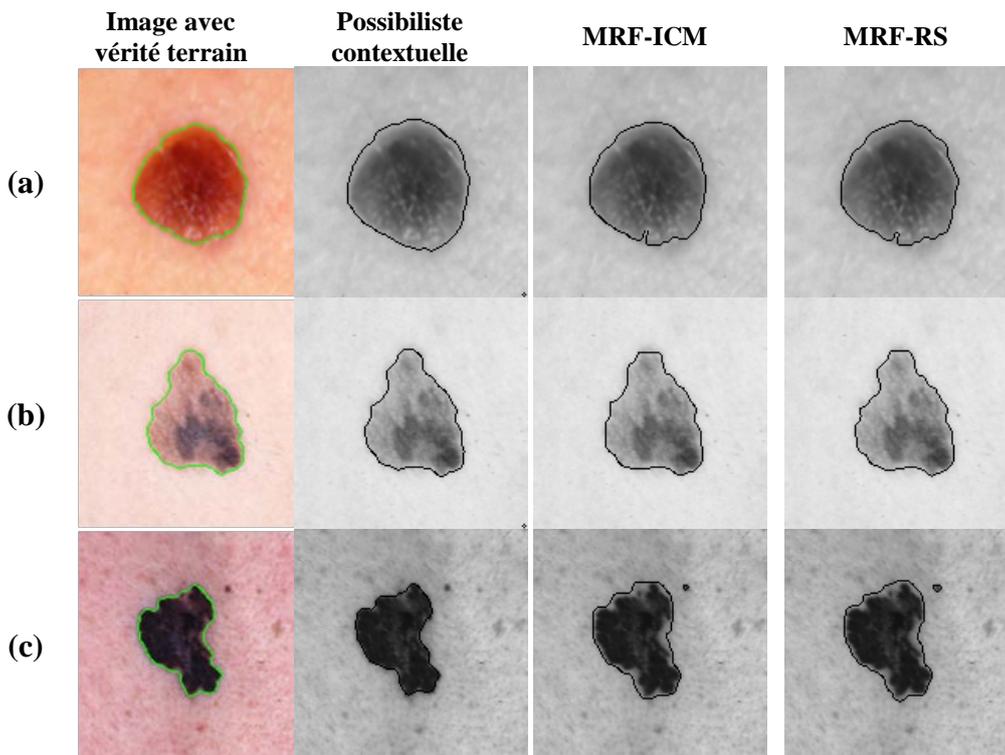


Figure 2.25 Résultats de classification pixellique pour trois lésions : (a) L12, (b) L2 et (c) L3.

L'approche possibiliste contextuelle est comparable aux approches Markoviennes, du fait que toutes les trois prennent en compte l'information contextuelle. Pour la majorité des images, on note une supériorité de la modélisation de la connaissance pixellique par un modèle possibiliste contextuel et sa nette efficacité par rapport aux approches Markoviennes. La lésion L12 est un exemple où les approches Markoviennes ont assuré de meilleurs résultats

de classification pixelliques que l'approche possibiliste contextuelle. Ceci est prouvé par la distance de Hamming (Cf. Figure 2.24). Ce résultat est prévisible, car comme le montre la Figure 2.25-a, cette lésion se montre bien contrastée et ne présente pas trop d'ambiguïté. Mais même dans ce cas, on note que la qualité de la segmentation est meilleure avec l'approche possibiliste proposée (Cf. Figure 2.25a) au niveau du point de réflexion de la lumière.

Par ailleurs, prenons l'exemple de la lésion L2 (Cf. Figure 2.25-b), qui présente un contexte d'ambiguïté qui se manifeste au niveau d'une diversité des niveaux de contraste au sein de la lésion. Certains de ces niveaux de contraste, entraînent une confusion avec la peau. Aussi l'exemple de la lésion L3 (Cf. Figure 2.25-c), qui présente un contexte d'ambiguïté se manifeste au niveau de la forme de la lésion (présence d'encoches). Pour ces deux exemples, l'approche possibiliste contextuelle assure une meilleure qualité de classification pixellique qui reflète une meilleure modélisation des connaissances pixelliques.

Ceci nous incite à évaluer les approches testées par d'autres critères quantitatifs. Nous calculons alors une matrice de confusion pour les 7 approches sur les 13 lésions et nous en déduisons les 8 critères suivants : reconnaissance, erreur, vrais positifs, vrais négatifs, faux positifs, faux négatifs, spécificité et sensibilité. La Figure 2.26 présente pour chacun de ces 8 critères, une moyenne sur les 13 lésions et ce pour chacune des 7 approches.

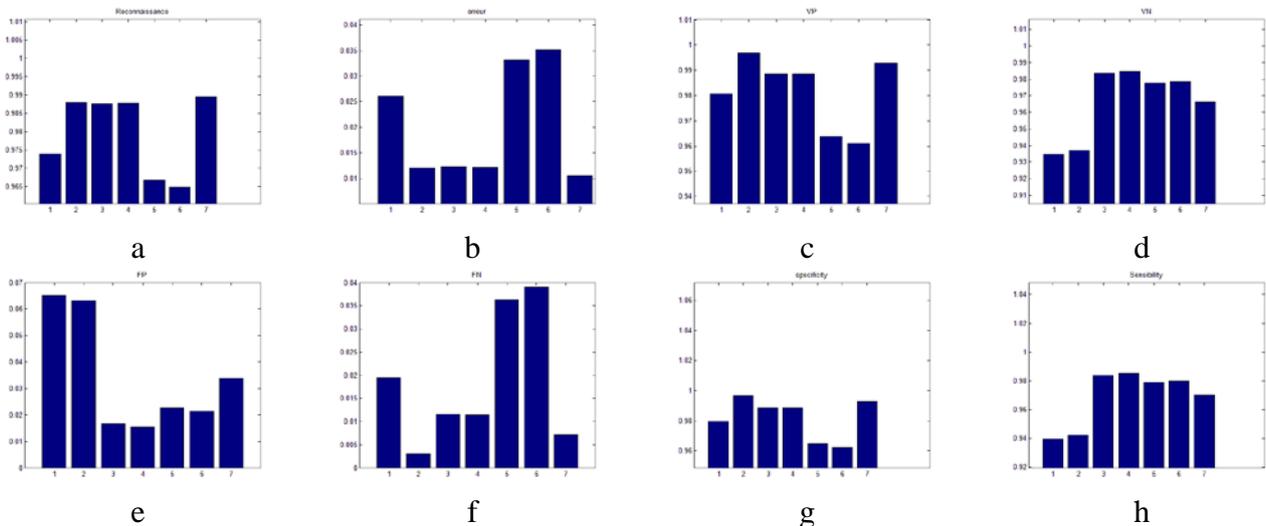


Figure 2.26 Résultats déduits de la matrice de confusion pour les 7 approches sur les 13 lésions, avec 1 : seuillage, 2 : k-means, 3 : Markov(ICM), 4 : Markov(RS), 5 : Bayésien classique, 6 : Possibiliste ponctuelle et 7 : Possibiliste contextuelle.

Les critères sont : a- Reconnaissance, b- Erreur, c- Vrais positifs, d- Vrais négatifs, e- Faux positifs, f- Faux négatifs, g- Spécificité et h- Sensibilité

Les résultats illustrés par la Figure 2.26 montrent de bonnes performances pour l'approche possibiliste contextuelle. Notons surtout qu'elle montre de nouveau son caractère de stabilité. En effet, pour les 8 critères d'évaluation utilisés et cités ci-dessus, l'approche possibiliste contextuelle proposée se montre ou bien la meilleure, sinon parmi les meilleures. Alors que les autres approches excellent pour certains critères mais se dégradent pour d'autres. A titre d'exemple, l'approche k-means réalise le meilleur score pour les faux négatifs, mais par contre un mauvais score pour les faux positifs.

4. 2^{ème} Contribution (en cours de développement): Modélisation contextuelle spatiale locale de deuxième ordre possibiliste d'un pixel; Notion de possibilité des possibilités

a) Description de l'idée

Une modélisation possibiliste du premier ordre s'effectue sur la base des valeurs radiométriques ou les mesures d'attributs qui en dérivent. Il en résulte que chaque pixel de l'image aura une valeur de possibilité d'appartenance à chacune des M classes de l'univers Ω . Il convient à ce niveau d'introduire la notion des cartes possibilistes. En effet, dans de telles procédures de modélisation, nous aurons autant de cartes possibilistes qu'il y en a de classes (M). Chaque carte est de la même taille que l'image initiale. A la position de chaque pixel de l'image, la carte présente sa valeur de possibilité d'appartenance à la classe à laquelle elle correspond.

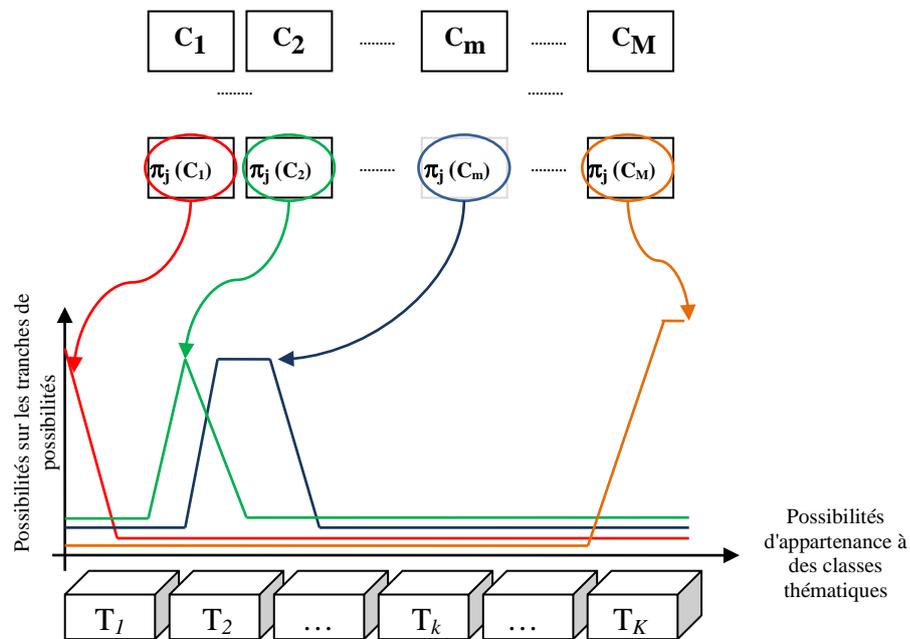


Figure 2.27 Estimation des distributions de possibilités définies sur l'ensemble des K tranches de possibilités d'appartenance aux M classes thématiques et associée au voisinage du pixel j

L'idée que nous proposons dans ce paragraphe consiste à représenter les connaissances pixelliques par un ensemble de M distributions de possibilités, estimées chacune sur le voisinage du pixel, pas dans l'image initiale (ou l'une de ses dérivées par attributs), mais plutôt dans chacune des cartes possibilistes correspondant aux classes thématiques. Les valeurs servant à l'estimation des distributions de possibilités ne sont plus des mesures radiométriques ou d'attributs, mais plutôt des valeurs de possibilités d'appartenance à une classe thématique des pixels voisins au pixel considéré. L'univers de mesures sur lequel nous travaillons dans ce cas n'est plus celui des classes thématiques mais celui des degrés de

possibilités d'appartenance à ces classe thématiques. N'étant pas un ensemble de singletons, nous proposons de le partager en K tranches égales $T_k, k=1..K$ (Cf. Figure 2.27).

En d'autres termes, on considère que le voisinage du pixel j , présente M sources, correspondant aux M classes thématiques, permettant de définir, M distributions de possibilités $\pi_{j,m}$ sur l'ensemble des K tranches de possibilités d'appartenance $\Omega_T = \{T_1, T_2, \dots, T_K\}$. La tranche T_1 correspond à la catégorie "la plus faible possibilité d'appartenance", alors que la tranche T_K correspond à la catégorie "la plus haute possibilité d'appartenance".

L'information spatiale que le voisinage fournit, quant à l'appartenance du pixel central à une classe thématique donnée, est alors une décision qui se formule en se rapportant à la fois aux M distributions $\pi_{j,m}, m=1..M$.

Plusieurs cas pourraient se présenter :

- ✓ **une redondance** : Cas où une seule distribution π_{j,m_0} , présente une forte possibilité d'avoir de fortes possibilités d'appartenir à la classe C_{m_0} . Alors que toutes les autres $\pi_{j,m}$, tel que $m \neq m_0$, présentent de fortes possibilités d'avoir de faibles possibilités d'appartenir aux classe $C_m, m \neq m_0$, Cf. Figure 2.28.
- ✓ **une ambiguïté**: Cas où le maximum de possibilités de plusieurs ou toutes les distributions $\pi_{j,m}, m=1..M$, se situent à la même tranche de possibilités d'appartenance à une classe thématique. A titre d'exemple, il serait, selon le voisinage du pixel, fortement possible que le pixel central ait de fortes possibilités d'appartenir à plusieurs classes à la fois, Cf. Figure 2.29.
- ✓ **un conflit des informations déduites des M distributions de possibilités**: Cas où une ou plusieurs distributions montrent l'ignorance totale comme distribution de possibilités sur les K tranches de possibilités d'appartenance aux classes thématiques, Cf. Figure 2.30. Ceci signifie que nous avons la totale possibilité d'avoir une faible possibilité d'appartenance et en même temps la totale possibilité d'avoir une forte possibilité d'appartenance à la classe thématique concernée (i.e. à la fois ça appartient et ça n'appartient pas = conflit)

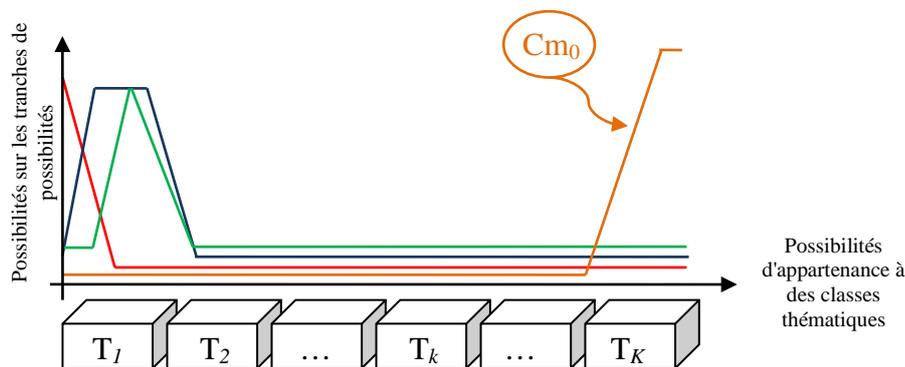


Figure 2.28 Redondance de l'information déduite des M distributions de possibilités définies sur l'ensemble des K tranches de possibilités d'appartenance aux classes thématiques. L'appartenance du pixel candidat est attribuée à la classe C_{m_0} par consensus.

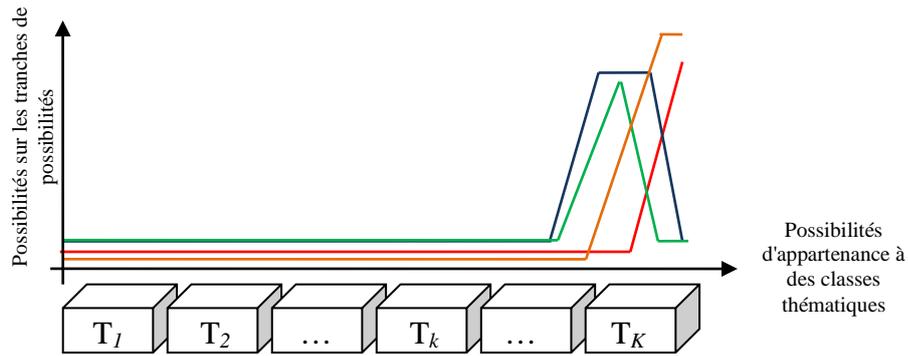


Figure 2.29 Ambiguïté sur l'appartenance à une ou plusieurs classes thématiques.

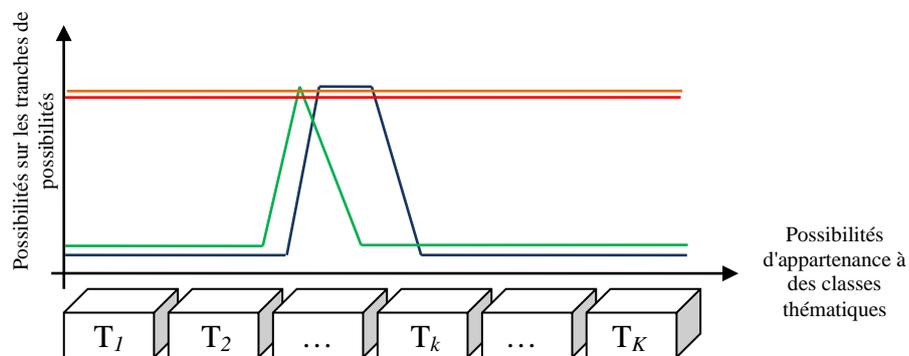


Figure 2.30 Conflit des informations déduites des M distributions de possibilités définies sur l'ensemble des K tranches de possibilités d'appartenance aux classes thématiques.

b) Simulation sur un échantillon

Considérons l'« image test » définie précédemment. La Figure 2.31-a montre la carte possibiliste d'appartenance à la classe "objet", le pixel central mis à l'examen d'appartenance est situé sur l'objet comme le montre la Figure 2.31-b. L'une des 9 cliques de Nagao (un voisinage défini au chapitre 4), portant des voisins du pixel sous examen, est présentée à la Figure 2.31-c. Les valeurs portées par la clique sont des valeurs de possibilité d'appartenance à la classe objet. Ces valeurs sont utilisées pour construire la distribution de possibilités sur K tranches de possibilités (K est prise égale à 5), Cf. Figure 2.31-d. La distribution de possibilités (celle donnée en bleu) est calculée moyennant l'approche proposée d'estimation de distributions de possibilités en se basant sur des intervalles de confiances de sison & glaz [CKA 2014]. La deuxième distribution (celle donnée en rouge) est calculée selon la même approche mais par les mesures de possibilités issues de la carte possibiliste d'appartenance à la classe "fond". La courbe rouge montre qu'on a une forte possibilité d'avoir une faible possibilité (dans la tranche $[0 \ 0.2]$) d'appartenir à la classe "fond". Alors que la courbe bleue montre qu'on a une forte possibilité d'avoir une forte possibilité (dans la tranche $[0.8 \ 1]$) d'appartenir à la classe "objet". Ainsi il est très logique que le voisinage considéré soit résumée par une mesure qui favorise l'appartenance du pixel sous examen à la classe "objet". Notons cependant que la valeur de possibilité d'appartenance

à la classe « objet » est une faible valeur, du fait que c'est représentée par un niveau de gris sombre sur la carte possibiliste de la classe thématique « objet » (Figure 2.31-c)

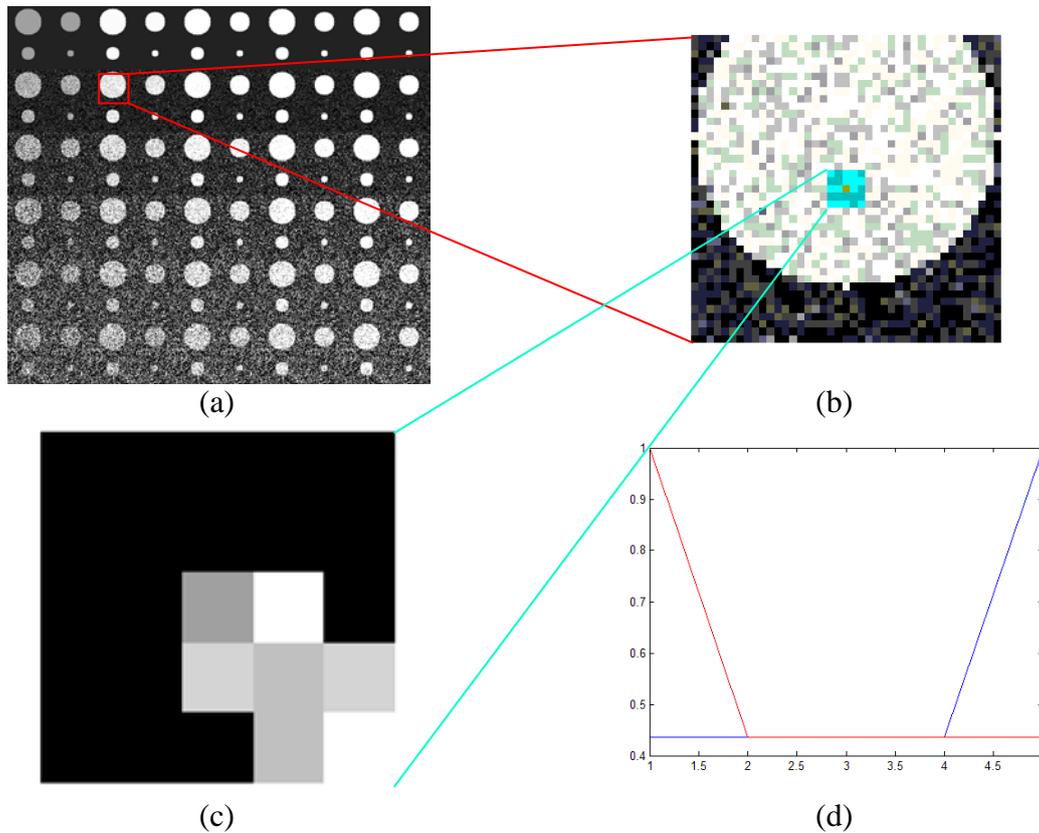


Figure 2.31 Illustration par l'image test

(a) Carte de possibilité de la classe thématique objet; (b) Zoom sur la région contenant le pixel sous examen entouré par ses voisins; (c) une des neuf cliques de Nagao; (d) Distributions de possibilités des possibilités d'appartenance à la classe objet (en bleu) et celle à la classe background (en rouge)

V. Sélection d'attributs en modélisation possibiliste

1. Problématique

La sélection de primitives (attributs, caractéristiques) est devenue de plus en plus importante dans les systèmes d'aide à la décision. Elle a comme but de réduire, outre que la complexité des traitements, diverses limitations, telles que les conflits, les redondances, le sur-apprentissage et bien d'autres, plus détaillées dans la littérature [KJ 1997, BL 1997, GE 2003, LMP 2001, ML 2003, McC 2003, BON 2003, TM 2000, SP 2003, SCY 2009, Amm 2014 et AKS 2015].

Les attributs n'ont pas la même importance vis-à-vis de la modélisation des différentes classes. Certains d'entre eux sont utilisés pour la séparation entre les classes. D'autres sont plus fiables pour la représentation des classes. Un dilemme se pose alors, lors d'un processus de sélection de primitives:

- sélectionner les primitives offrant une meilleure caractérisation des classes, assurant ainsi, une minimisation de la variance intra-classe tout en gardant une cohérence maximale, ou

- favoriser les primitives offrant une meilleure séparation entre les classes, assurant ainsi, une maximisation de la variance inter-classe.

En effet, la minimisation de la variance intra-classe qui assure une meilleure spécification ou caractérisation de la classe, n'a aucune relation avec la maximisation de la variance inter-classes, qui favorise la séparation entre les classes, et inversement. Un compromis entre les deux critères, s'impose dans la majorité des cas. Le processus de sélection d'attribut, se transforme alors lui-même, en un problème de gestion de conflits et prise de décision. La solution dépend le plus souvent de la nature de l'application et de ses objectifs.

Dans cette section nous décrivons le processus de sélection de primitives dans un cadre possibiliste, que nous avons proposé dans [AKS 2015].

L'approche consiste à évaluer la qualité de modélisation possibiliste, d'un attribut dans un univers fermé de classes, afin de juger sa fiabilité. L'évaluation prend en compte deux considérations : la représentation et la séparation des classes. L'appréciation de la qualité de l'attribut envers ces deux considérations, est menée sur deux phases. L'une est basée sur le paradigme de Shapley, l'autre est basée sur la mesure d'incertitude.

2. Cadre de l'étude

On considère pour la suite

- ✓ Ω est un univers fermé de M classes, $\Omega = \{C^1, C^2, \dots, C^m, \dots, C^M\}$
- ✓ Chaque classe C^m regroupe un nombre $Q(m)$ éléments (ou objets, membres...),
$$C^m = \{x_1^m, x_2^m, \dots, x_q^m, \dots, x_{Q(m)}^m\}$$
- ✓ L'univers Ω est observé par K sources $S_k, k=1..K$
- ✓ Chaque source S_k observe l'univers Ω à travers $N(k)$ angles de vue différents, autrement dit avec un nombre $N(k)$ d'attributs $f_{n,k}, n = 1..N(k)$,

Les valeurs de l'attribut $f_{n,k}$ calculées sur les observations des membres d'une classe C^m par une source S_k , servent à construire des modèles pour la classe C^m . Le modèle choisi pour cette étude est la distribution de possibilités obtenue à travers l'une des techniques d'inférence de distribution de possibilités à partir d'une distribution de probabilités. La distribution de probabilités est construite par la méthode d'histogramme.

Un modèle possibiliste $\pi_{n,k}^m$ est ainsi construit :

- pour chaque classe $C^m, m = 1..M$, et
- selon chaque attribut $f_{n,k}, n = 1..N(k)$ et $k = 1..K$.

La Figure 2.32 résume le cadre de cette étude.

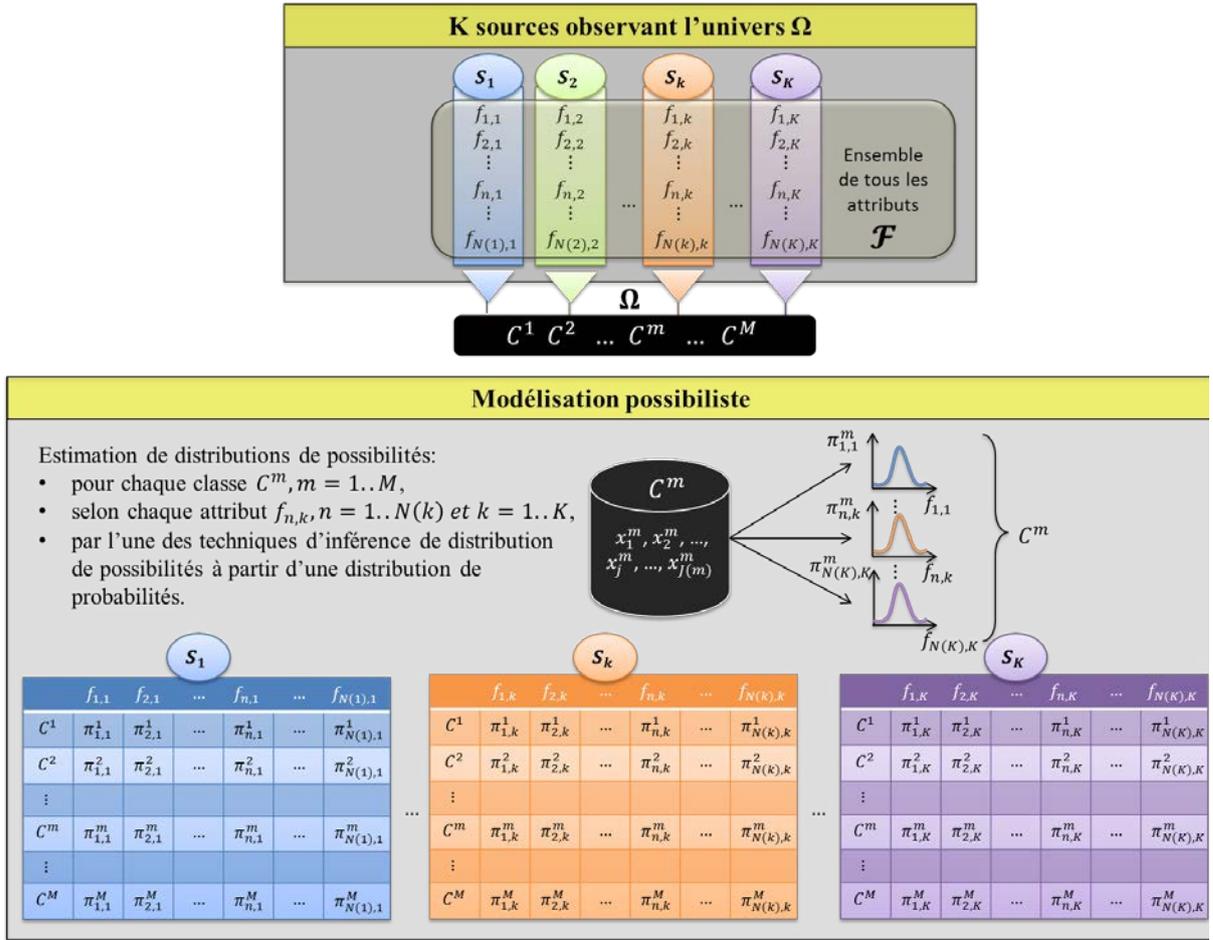


Figure 2.32 Représentation schématique du contexte de l'étude

3. Outils techniques

a) Indice de Shapley

Le formalisme de Shapley est l'un des principaux concepts de solution de la théorie des jeux coopératifs [Chr 2006]. L'objectif le plus considéré pour son utilisation est la répartition des gains dans une coopération. Shapley a proposé une répartition « équitable » des gains de la coalition de n joueurs, en les pondérant par des valeurs dites « indices de Shapley ». Ces indices définissent en quelque sorte la valeur de l'importance de chaque intervenant dans la coopération. Ces valeurs sont basées sur le concept de fonction caractéristique qu'on propose de nommer \mathcal{M} . Soit $\mathcal{M}(L)$ une mesure qui quantifie le gain ou profit maximum de la coalition L . L'indice de Shapley d'un joueur f d'un ensemble de joueurs \mathcal{F} est donné par :

$$v_f = \sum_{i=0}^{N-1} \gamma_i \cdot \sum_{L \subset \mathcal{F} \setminus \{f\} / N_L=i} (\mathcal{M}_{f,L} - \mathcal{M}_L) \quad \text{Eq2.2}$$

N étant le nombre total des joueurs de \mathcal{F} , L est une coalition de \mathcal{F} de cardinal N_L .

$\mathcal{M}_{f,L} = \mathcal{M}(f, L)$ et $\mathcal{M}_L = \mathcal{M}(L)$.

La pondération des coopérants par leurs valeurs de Shapley respectives permet d'assurer un équilibre concurrentiel, dans une coopération. Dans la présente étude, ces valeurs nous

servent à donner une évaluation de la qualité d'un attribut dans la modélisation des classes vis-à-vis du profit exprimé par la fonction caractéristique.

b) Mesure d'incertitude

Lors de la modélisation d'un phénomène, on est souvent en présence d'une information « imprécise, vague voire incomplète ». Cette incertitude est liée à la méconnaissance du phénomène et reflète la limitation des données d'entrée en qualité et en quantité. La mesure d'incertitude permet d'estimer la quantité d'information fournie par un modèle. Elle reflète aussi, entre autre, la qualité informative de la primitive qui a servi à construire le modèle.

La mesure d'incertitude d'un modèle possibiliste, utilisée dans cette étude est définie par :

$$U(\pi) = [1 - \pi(\sigma(B))] \cdot \ln(B) + \sum_{b=2}^B ([\pi(\sigma(b)) - \pi(\sigma(b+1))] \cdot \ln(b)) \quad \text{Eq2.3}$$

avec:

$$\pi(\sigma(B+1)) = 0 \quad \text{Eq2.4}$$

σ est une fonction de permutation arbitraire définie comme:

$$\begin{aligned} \sigma: [1..B] &\rightarrow [1..B] \\ b &\mapsto \sigma(b) \end{aligned} \quad \text{Eq2.5}$$

Tel que :

$$\pi_{n,k}^m(\sigma(b)) \geq \pi_{n,k}^m(\sigma(b+1)), \text{ pour } b = 1..B-1 \quad \text{Eq2.6}$$

4. Description de la contribution

L'indice de Shapley a déjà été utilisé [Gra96], pour caractériser une classe, en utilisant la fonction d'appartenance comme fonction caractéristique. Notre contribution consiste, d'abord, à étendre ce paradigme pour évaluer l'importance de l'attribut dans le processus de discrimination entre les classes, ensuite, à exploiter le concept du degré d'incertitude d'une distribution de possibilités pour traiter le formalisme de Shapley.

L'approche proposée se déroule sur deux phases :

La première phase, illustrée par la Figure 2.33 consiste en une estimation de l'importance d'un attribut par rapport à l'ensemble des autres attributs considérés. Cette estimation de l'importance de l'attribut est basée sur le paradigme de Shapley [Zac 1988]. Les figures Figure 2.34 et Figure 2.35, illustrent les procédures d'estimation de l'importance des attributs dans, respectivement, la séparation et la représentation des classes. Quand le gain à réaliser est la séparation des classes, un seul indice de Shapley est calculé par attribut. Alors que dans le cas où le gain à réaliser est la représentation de classes, M indices de Shapley sont calculés par attribut, autrement dit, un indice de Shapley pour estimer l'importance de l'attribut dans la caractérisation de chacune des M classes. Deux critères sont à vérifier pour la validation du premier test de sélection de l'attribut $f_{n,k}$, se basant sur son indice de Shapley $v_{n,k}$ dans le cas de son évaluation pour la séparation entre les classes et $v_{n,k}^m$ dans le cas de son évaluation pour la représentation de la classe C^m (Cf. Figure 2.33):

Critère 1: $v_{n,k} \geq 1$

Critère 2: $v_{n,k} \geq \delta_1$, tel que $\delta_1 = \frac{\max_{n=1..N(k)}(v_{n,k})}{\sum_{n=1}^{N(k)} v_{n,k}}$

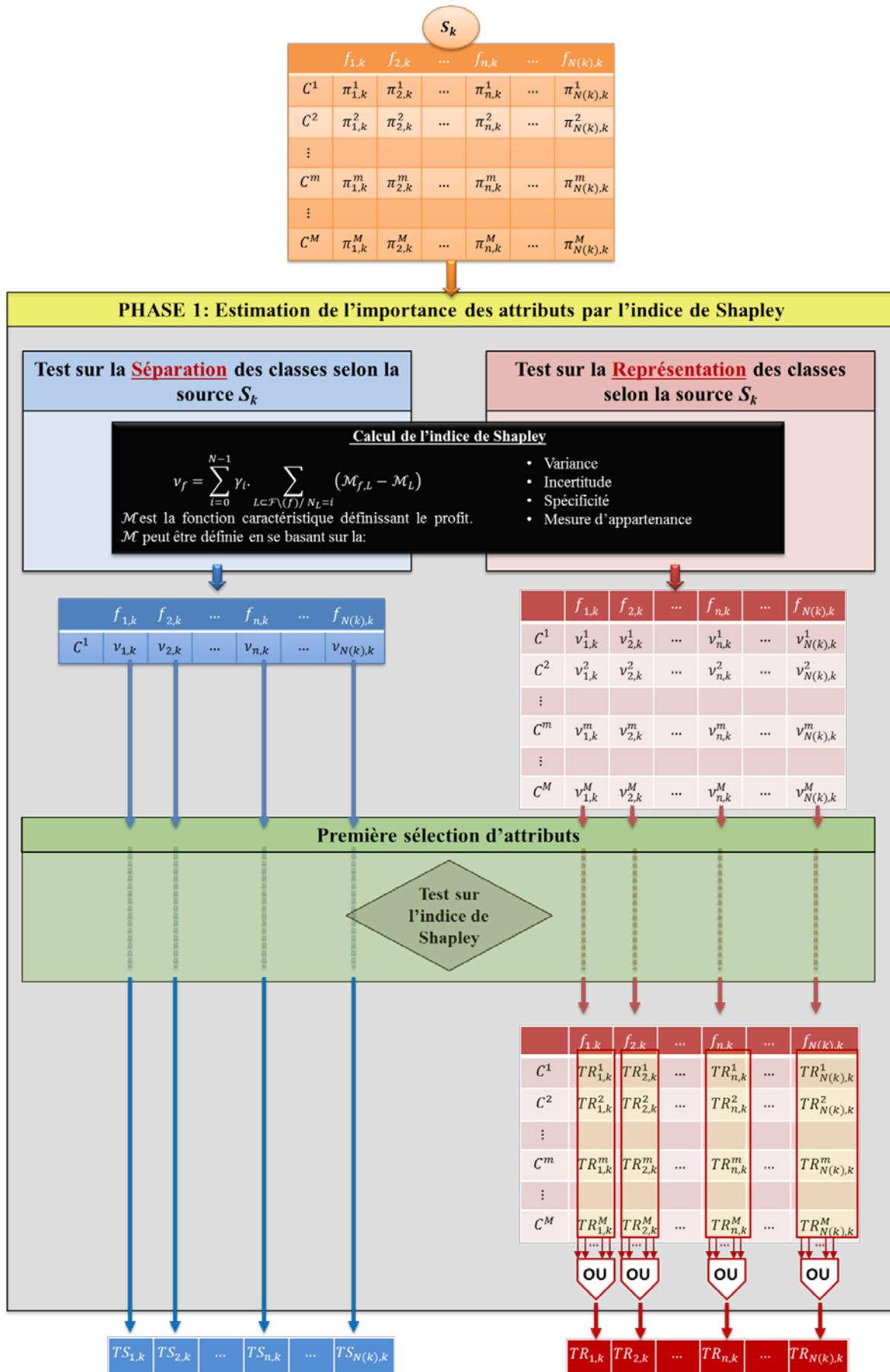


Figure 2.33 Estimation de l'importance des attributs par l'indice de Shapley

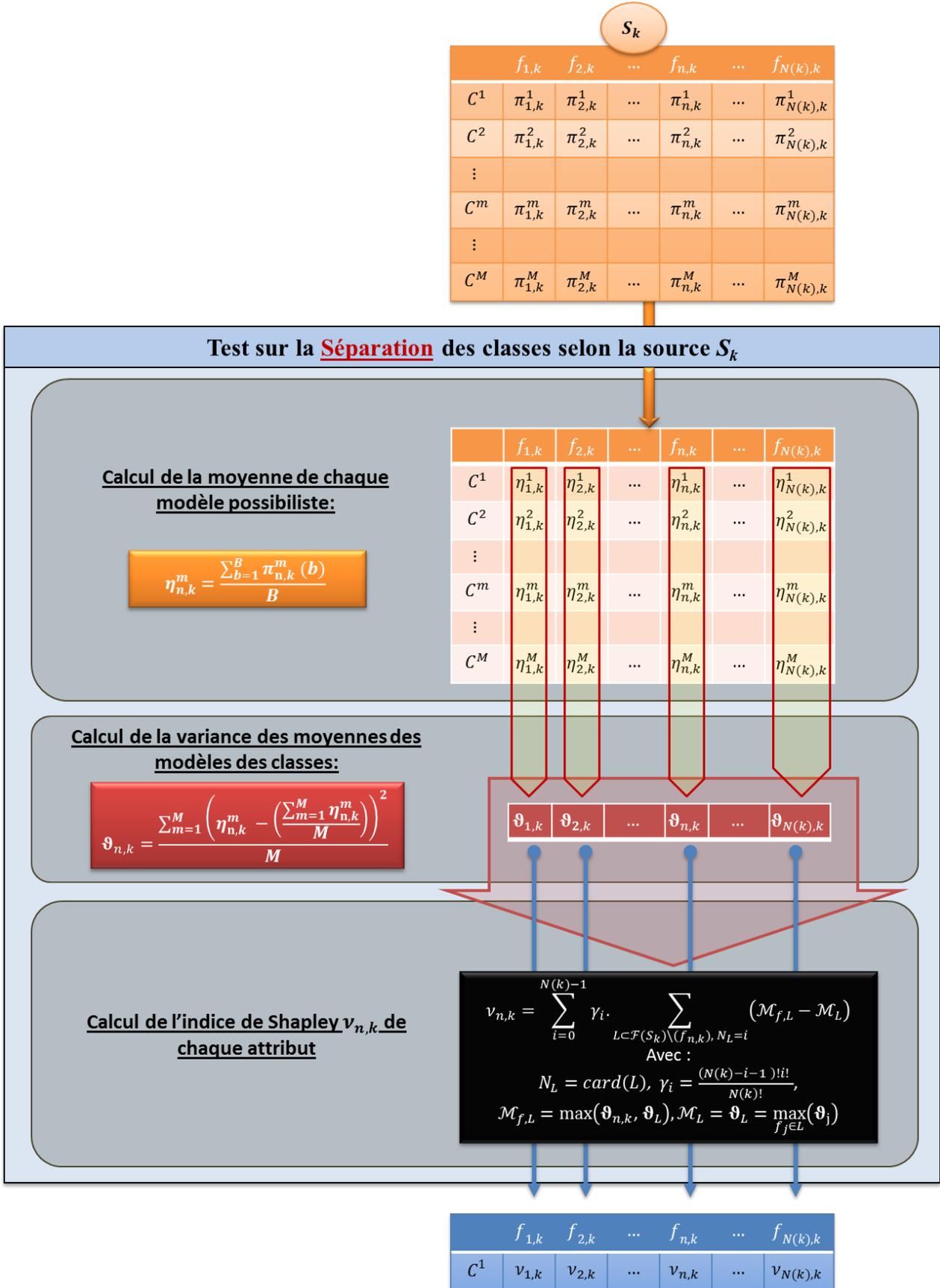


Figure 2.34 Test sur la séparation des classes selon la source S_k

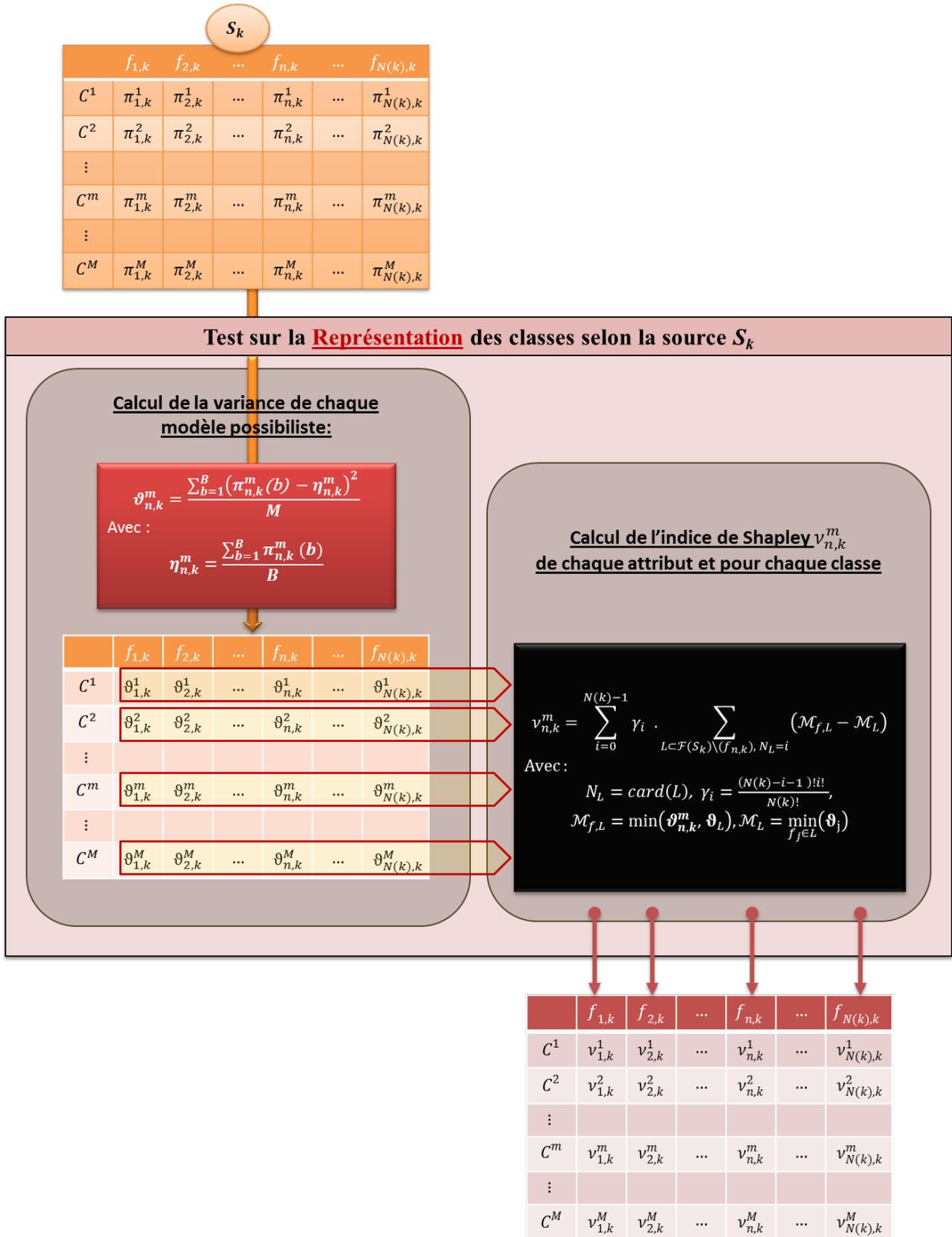


Figure 2.35 Test sur la représentation des classes selon la source S_k

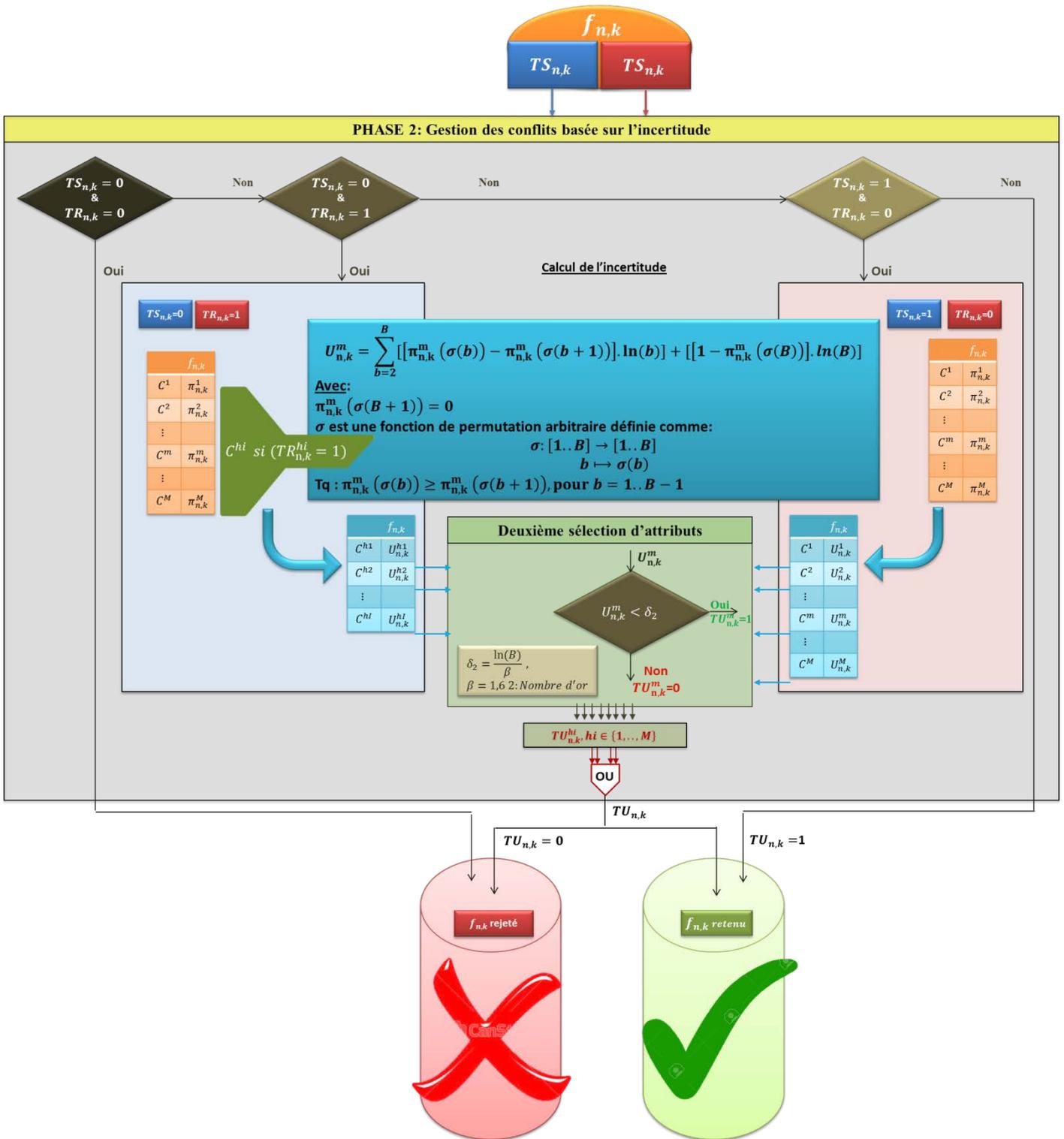


Figure 2.36 Gestion des conflits basée sur l'incertitude

Deux résultats sont issus de ce premier test : un pour la séparation (TS) l'autre pour la représentation (TR) des classes. Un test négatif, $T(S/R)=0$, signifie que l'attribut évalué ne répond pas aux critères suggérés. Un test positif, $T(S/R)=1$, signifie que l'attribut évalué, est bien convenable pour la Séparation/Représentation des classes.

La deuxième phase, illustrée par la Figure 2.36, consiste en une prise de décision vis-à-vis de la rétention ou rejet des attributs en se basant sur les résultats des deux tests de la première phase, à savoir, TS et TR. En effet, les deux tests peuvent convenir sur la même opinion vis-à-

vis d'un attribut, comme ils peuvent diverger d'opinions, de sorte que l'un des tests valide alors que l'autre rejette l'attribut. Ainsi, dans le cas de :

- Une convenance positive : ($TS=1$ et $TR=1$), l'attribut évalué est retenu,
- Une convenance négative : ($TS=0$ et $TR=0$), l'attribut évalué est rejeté, et
- Un désagrément : ($TS=1$ et $TR=0$) ou ($TS=0$ et $TR=1$), un processus de gestion et résolution de conflits de sélection est lancé. Ce processus illustré par la Figure 2.36, consiste à trouver des consensus entre les deux tests en conflit, en se basant sur une estimation des incertitudes engendrées par les modèles fournis par l'attribut en jeu.

5. Expérimentations et discussion

Deux expériences sont menées afin de valider la stratégie proposée, de sélection d'attributs dans un contexte possibiliste.

La première est menée sur des données synthétiques, construites de façon à connaître en avance l'importance de chaque attribut dans la modélisation des classes et ce, vis-à-vis de la représentation et de la séparation. Cette expérience, menée dans un contexte contrôlé, nous sert à vérifier que les résultats prévus, soient réellement atteints. Les données synthétiques sont présentées sous forme de distributions de possibilités, modélisant deux classes. L'univers des deux classes est observé par trois sources. Cinq attributs sont supposés caractériser les observations de chacune des trois sources. L'expérience, détaillée dans [AKS 2015], a montré que l'application de la procédure de l'approche de sélection proposée, a bien mené à la sélection des attributs adéquats et le rejet des attributs sans intérêt à l'application.

Quant à la deuxième expérience, elle consiste à évaluer et comparer, à d'autres approches de sélection d'attributs, l'apport de l'approche proposée dans un processus de classification précédé d'une opération de sélection d'attributs. La tâche de classification est assurée par un classifieur SVM.

Les expérimentations sont menées sur:

- six bases de données de référence (benchmark), à savoir : une base de données de verres [BL 2013, AS 2004, NHR 2011], une base de données des troubles du foie [BL 2013b, ZJ 2004 et GA 1997], la base de données de fleurs d'iris [Bla, TXO 2005, DB 2004], la base de données svmguide3 [GFH 2009, HCL 2003, CL 2011], une base de données de diabète Indiens de Pima [UCI 2008] et une base de données de silhouettes de véhicules [UCI 2008] et
- huit approches de sélection d'attributs : Analyse en composantes principales (Principal Component Analysis : PCA [Smi 2002], Sélection par régression linéaire (Linear Forward Selection : LFS) [GFH 2009], score de Fisher [DHS 2001], la méthode RELIEF [Kon 1994], les trois critères de sélections employés dans notre algorithme, pris chacun à part, à savoir : le critère de représentation (CR), celui de séparation (CS) et celui de l'incertitude (CU) et finalement l'approche proposée.

L'évaluation de l'apport d'une approche de sélection d'attributs, se base sur deux notions : le taux de réduction du nombre d'attributs et le taux de correcte classification. Pour chacune des approches et base de données considérées dans ce travail, le nombre d'attributs sélectionnés (NAS) ainsi que le taux de correcte classification (TCC) réalisés sont donnés dans le Tableau 2.3.

Tableau 2.3 Bases de données

Base de données	Iris		Verres		Troubles du foie		SVMGuide3		Diabètes		Véhicules	
Nombre total des Attributs	4		9		6		21		8		18	
Approche	NAS	TCC	NAS	TCC	NAS	TCC	NAS	TCC	NAS	TCC	NAS	TCC
Sans Sélection	4	96.2	9	95.3	6	67.3	21	83.2	8	78.1	18	82.8
PCA	2	81.3	6	93.9	5	68.4	13	78.3	4	75.6	7	74.7
LFS	2	95.3	7	93.9	3	60.5	4	79.6	4	77.8	10	70.6
Fisher	2	81.3	4	91.5	3	64.8	11	80.4	6	76.4	10	74.2
RELIEF	2	81.3	4	93.4	4	69.3	4	76.2	4	77.2	10	72.8
CS	1	96	2	86.4	2	61.2	16	82.5	2	74.7	3	63.4
RS	2	95.3	5	90.6	4	69.5	17	82.3	2	65.1	7	68
CU	4	96.2	9	95.3	6	67.3	21	83.2	8	78.1	18	82.8
Notre Approche	2	95.3	7	94.4	5	66.3	17	82.3	4	76.3	8	72.7

Nous proposons de quantifier le gain réalisé par une approche de sélection, à travers deux taux :

- Taux de sélection des attributs (TS) :

$$TS = \frac{\text{Nbre}(\text{attributs sélectionnés})}{\text{Nbre}(\text{tous les attributs})} \quad \text{Eq2.7}$$

La valeur de TS est toujours inférieure ou égale à 1. Une importante réduction du nombre d'attributs correspond à une faible valeur de TS .

- Taux de répercussion sur la classification donnée par le taux de classification obtenu après sélection, par rapport à celle obtenue par usage de la totalité des attributs.

$$TRC = \frac{\text{TCC}(\text{après sélection d'attributs})}{\text{TCC}(\text{avec tous les attributs})} \quad \text{Eq2.8}$$

Une valeur de TRC supérieure à 1, signifie que la sélection a amélioré le taux de classification réalisé. Une valeur de TRC inférieure à 1, signifie que la sélection a dégradé la reconnaissance. Plus l'écart $|1-TRC|$ est important, plus l'impact de la sélection sur la classification est important.

La Figure 2.37 illustre la variation de TRC en fonction du TS . Chaque point correspond aux performances (TS, TRC) réalisées par l'approche pour une base de données. Il est bien visible que très peu de couples (Approche/base de données) présentent un TRC supérieur à 1. Ceci peut être dû au fait que les attributs aient pu être présélectionnés. Il est à noter que l'approche proposée montre une stabilité remarquable par rapport aux autres approches, avec des valeurs de TRC très proches de 1. Ceci est visible par l'allure linéaire, horizontale, serrée, rectiligne, proche de l'unité, de sa courbe.

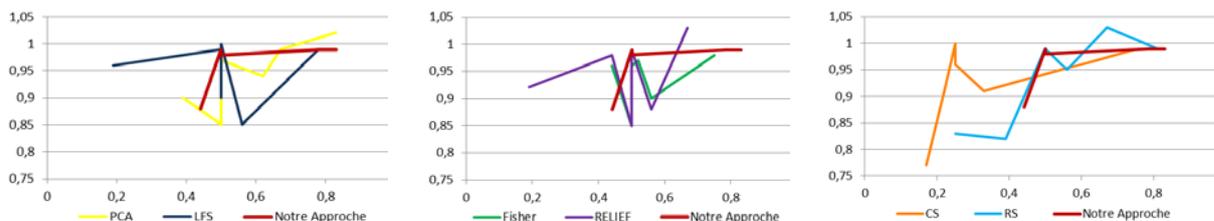


Figure 2.37 Variation du TRC en fonction du TS

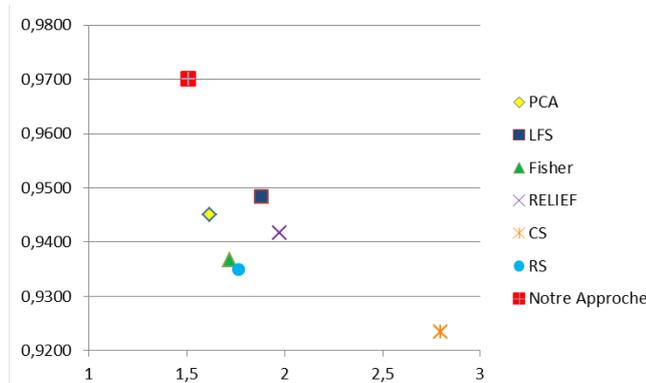
Pour avoir une vision plus globale des performances réalisées par les différentes approches, les moyennes et variances des taux *TS* et *TRC* respectives des différentes approches, obtenues sur les diverses bases de données considérées, sont calculées et présentées par le Tableau 2.4. Les moyennes donnent une estimation globale de la performance de l'approche, alors que la variance renseigne sur la stabilité de cette performance par rapport au type de données testées. Une faible variance signifie que la performance de l'approche est indépendante du type de données testées. Les valeurs marquées désignent les meilleures réalisations des taux *TS* et *TRC*. En effet, il est souhaité qu'une approche offre un bon taux de reconnaissance pour un bas taux de sélection. L'approche proposée montre le plus important *TRC* pour un *TS* modéré. Ceci signifie que la stratégie employée favorise l'atteinte de l'objectif principal de l'application, étant une correcte décision, tout en poussant au mieux la réduction de la complexité. Le critère *CS* par contre, favorise plutôt la réduction de la complexité en dépit d'une diminution de la performance fondamentale.

Tableau 2.4 Moyennes et variances des taux *TS* et *TRC* respectives des différentes approches, obtenues sur les diverses bases de données considérées, les valeurs éclairées sont les meilleures réalisations des performances correspondantes.

Approches	TS		TRC	
	Moyenne	variance	Moyenne	variance
PCA	0,5850	0,0202	0,9450	0,0032
LFS	0,5050	0,0297	0,9483	0,0030
Fisher	0,5450	0,0097	0,9367	0,0022
RELIEF	0,4767	0,0215	0,9417	0,0040
CS	0,3300	0,0392	0,9233	0,0059
RS	0,5300	0,0330	0,9350	0,0066
Notre Approche	0,6433	0,0273	0,9700	0,0016

Afin de mieux voir l'optimisation effectuée par l'approche proposée, nous représentons à la Figure 2.38 pour chaque approche testée, sa moyenne, sur toutes les bases de données, des taux de classification correcte, en fonction de leurs gains de sélection respectifs. Une estimation significative du gain réalisé par chaque approche, en tenant compte des considérations citées précédemment, peut être défini en tant que :

$$Gain = \frac{\text{moyenne}(TRC)}{\text{moyenne}(TS)} \tag{Eq2.9}$$



La Figure 2.38 confirme les déductions obtenues à travers le Tableau 2.4, en effet notre approche de sélection d'attributs, devance toutes les autres approches pour la réalisation du meilleur taux de classification, pour une réduction modérée de la complexité.

VI. Conclusion

Quatre niveaux de représentation des connaissances se placent sur une échelle sémantique d'interprétation, selon la sémantique qu'ils renferment. Le niveau bas correspond aux données brutes, suivi du niveau intermédiaire des primitives, puis vient le niveau sub-sémantique qui regroupe les représentations permettant la modélisation des éléments d'une scène et leur projection dans un autre espace de connaissances. Finalement le niveau symbolique ou sémantique regroupant les décisions. Nous nous sommes intéressés dans ce chapitre à la représentation des connaissances liées aux éléments de la scène au niveau sub-sémantique. Une synthèse des activités a d'abord été donnée, suivie par la description des contributions effectuées pour modéliser des connaissances dans des circonstances particulières d'ambiguïté.

Notre première contribution consiste en une approche d'estimation de distributions de possibilités pour modéliser des données faibles en quantité, en se basant sur les intervalles de confiance simultanés de Sison et Glaz. L'approche consiste en une modification d'une approche de l'état de l'art [MD 2006], en revanche la stratégie d'évaluation est authentique. Cette stratégie conçue et testée sur des données synthétiques et d'autres réelles a permis de souligner l'apport de la modification apportée à l'approche de l'état de l'art adoptées.

Notre deuxième contribution concerne la modélisation possibiliste de connaissances pixelliques. Deux approches de modélisation possibiliste des connaissances pixelliques ont été proposées : une approche ponctuelle et une approche contextuelle. Les deux approches sont évaluées et comparées à d'autres approches de l'état de l'art. L'expérimentation a été menée sur deux types d'images : une image synthétique, de façon à bien contrôler et connaître les modèles de fabrication de l'image et des images optiques médicales qui présentent des contextes réelles d'ambiguïté. La modélisation contextuelle (local based possibility distribution) s'est avérée beaucoup plus intéressante que la modélisation ponctuelle. L'utilisation d'un modèle local d'estimation se basant sur des primitives contextuelles possibilistes, a pu assurer des performances d'une efficacité supérieure aux autres approches, notamment pour des qualités dégradées de l'image.

Dans le même contexte de modélisation des connaissances faiblement définies, se situe notre 3ème Contribution qui est encore en cours de développement. Il s'agit d'une modélisation contextuelle spatiale locale du deuxième ordre possibiliste du pixel. C'est une proposition qui découle de la notion de possibilité des possibilités. L'idée a été décrite, puis une simulation sur un exemple numérique a servi pour donner des interprétations constructives et utiles pour la continuation de ce travail.

Finalement, notre quatrième contribution consiste en une procédure de sélection d'attributs dans un contexte possibiliste, en intégrant la notion de l'incertitude à l'application du paradigme de Shapley. La sélection dépendait de deux critères fondamentaux qui caractérisent l'importance d'un attribut : sa capacité de représenter et celle de séparer des classes.

Chacune de nos contributions a été évaluée et comparées à d'autres approches de l'état de l'art. Les expérimentations ont été menées sur deux types de données : des données

synthétiques, de façon à bien contrôler leurs particularités et des données réelles qui présentent des contextes d'ambiguïtés réelles.

Les résultats d'expérimentation de l'ensemble de ces travaux, ont permis de tirer les conclusions suivantes :

Les approches possibilistes offrent des avantages importants en termes de représentation des connaissances souffrant d'incomplétude en quantité (peu de données disponibles) et en qualité (ambiguïté). Ces avantages permettent à ces approches d'être plus efficaces que les méthodes usuelles de modélisation dans les contextes de connaissances faiblement définies. Une optimisation et une stabilité dans les performances réalisées par les approches possibilistes proposées, ont été spécialement notées.

Dans ce qui suit, nous procédons de la même manière d'étude et d'expérimentation pour présenter nos apports dans le domaine de prise de décision dans un contexte possibiliste.

Chapitre III: **Aide à la décision dans un contexte
possibiliste**

CHAPITRE III: AIDE A LA DECISION DANS UN CONTEXTE POSSIBILISTE	103
I. INTRODUCTION.....	105
1. Base de cas et de connaissances	105
a) Base de cas	105
b) Base de connaissance	106
2. Processus d'association de cas ou de connaissances à des décisions.....	107
II. SYNTHÈSE DES ACTIVITÉS RELATIVES À CE CHAPITRE	108
1. Collaborations et encadrements	108
2. Publications	109
III. MESURE DE SIMILARITÉ	110
1. Notion de similarité dans un contexte possibiliste	111
2. Propriétés d'une mesure de similarité possibiliste	112
3. Etat de l'art sur les mesures de similarité en théorie de possibilités	112
a) Mesure de similarité « Information Closeness »	112
b) Mesure de similarité déduite de l'information de divergence	113
c) Mesure de Sangüesa :	113
d) Distance de Minkowski et ses dérivés	113
e) Mesure de similarité delta δ	114
f) Mesure de similarité « Info Affinity »	114
4. Analyse de l'état de l'art.....	114
a) Simulation comparative sur des exemples numériques	114
b) Critique	114
c) Motivation	115
5. Contribution : Proposition d'une nouvelle mesure de similarité « Info-spécificity »	116
a) Principe	116
b) Définition	116
6. Méthodologie d'évaluation	117
a) Facteurs influant sur l'évaluation	117
b) Données utilisées pour les expérimentations.....	117
c) Approches d'estimation de mesures de similarité à comparer	118
d) Description de la procédure d'évaluation.....	118
7. Résultats et discussions	122
8. Emploi de la mesure InfoSpécificity dans la construction d'un arbre de décision possibiliste.....	125
a) Notions préliminaires concernant un arbre de décision.....	125
b) Contribution et application.....	126
IV. TRAVAUX DANS LE CADRE DE LA NOTION DE FUSION	127
1. Sélection d'opérateurs de fusion dans un cadre possibiliste	127
a) Approche de Weiru Liu [Liu 2007]	128
b) Modification proposée	129
c) Expérimentations et résultats.....	130
2. Analyse multicritère de la décision : MCDA.....	135
a) Objectif	135
b) Contexte	136
c) Approche	136
V. FOUILLE POSSIBILISTE	137
1. Construction d'une carte possibiliste.....	138
2. Prise de décision basée sur les cartes possibilistes	139
3. Application	140
VI. CONCLUSION.....	140

I. Introduction

La prise de décision apparaît dans la littérature, en tant que domaine de recherche actif et à multiples facettes qui se situe aux interfaces de nombreux domaines. Les chercheurs engagés dans ce sujet visent à automatiser le processus de prise de décision, en imitant la manière du cerveau humain. Selon la psychologie cognitive, le cerveau résout le problème de décision selon deux concepts fondamentaux: le premier dépend du traitement des représentations internes de l'information et le second concerne les transformations subies par ces représentations mentales. L'automatisation du processus de prise de décisions défie plusieurs difficultés. L'une des principales, réside à la manière d'exploitation de l'information disponible pour aboutir à une décision. Une forte tendance des travaux de recherche actuels cherche à mettre en coopération des méthodes hybrides permettant d'exploiter, d'une façon conjointe, les deux niveaux de traitement de l'information : le bas niveau (mesures numériques constituant une base de cas) et le haut niveau (le contenu sémantique de l'information traitée constituant une base de connaissances). Notons que l'une des plus grandes difficultés rencontrées, à ce stade, réside dans l'absence d'un formalisme général pour l'intégration de ces deux niveaux informationnels numérique et sémantique. A partir de cette constatation, il semble donc important de rechercher une méthodologie inspirée du raisonnement humain qui permet une telle intégration d'information pour aboutir à une extraction de savoir, à partir des deux bases conjointement, à savoir la base de cas et celle de connaissances. Cette intégration est retrouvée dans les modèles possibilistes à travers leurs façons de représenter l'information par le concept des distributions de possibilités. Dans un contexte possibiliste, la prise de décision dans une application donnée, peut alors se référer à un processus d'association de cas ou de connaissances à l'une des décisions de l'univers des décisions pré-envisagé pour l'application en question. A ce stade, il convient d'introduire ces notions, à savoir, la base de cas, la base de connaissances, et le processus d'association cas/connaissance/décision.

1. Base de cas et de connaissances

Dans un problème de prise de décision, deux types d'informations sont envisageables : les informations exprimées sous la forme de modèle des différentes classes et les informations exprimées sous la forme d'un ensemble d'exemples renseignés (dits : objets, patterns, cas....).

a) Base de cas

Dans le cas où les connaissances sont exprimées sous la forme d'exemples (Cf. Figure 3.1), ce problème peut être formulé de la manière suivante

On dispose de :

- un espace de N objets : $\Theta = \{\theta^1, \theta^2, \dots, \theta^N\}$
- un univers de K classes : $\Omega = \{C^1, C^2, \dots, C^K\}$
- un ensemble de I attributs (features, caractéristique, paramètres, primitives...)
 $A = \{A_1, A_2, \dots, A_I\}$

Chaque objet θ^n est présenté par le vecteur: $\begin{pmatrix} a_1^n \\ a_2^n \\ \vdots \\ a_I^n \end{pmatrix}$ doit appartenir à une unique classe

de l'univers Ω .

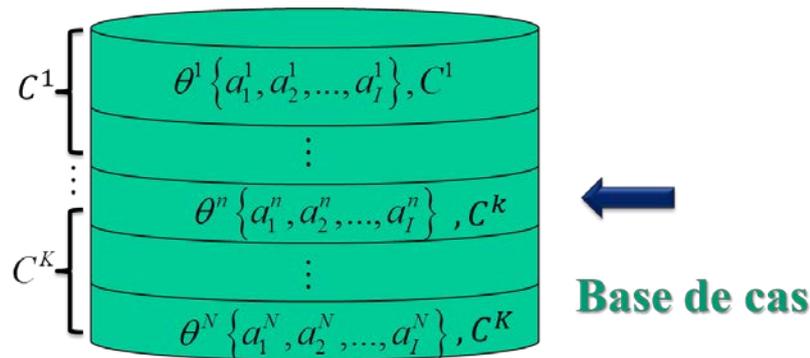


Figure 3.1 Base de cas

b) Base de connaissance

La base de connaissances est supposée encapsuler les connaissances d'experts concernant les différentes prévisions ou éventualités considérées dans une application donnée. Toute éventualité dans cette base est représentée par l'expert en utilisant l'ensemble des modalités possibles de tous les paramètres prédéfinis. Cette représentation se base, donc, sur la description de la relation entre les modalités des différents paramètres et les différentes éventualités. La vision de l'expert, sur l'occurrence d'une modalité pour une éventualité donnée est généralement exprimée de façon probabiliste de la relation « modalité-éventualité » en utilisant des descripteurs ou termes linguistiques ou syntaxique se référant à l'évaluation de l'expert de l'occurrence conditionnelle de la modalité pour l'éventualité donnée. Ce modèle de représentation diffère de l'approche statistique qui calcule la valeur exacte de la fréquence d'occurrence (i.e. relation entre modalité et éventualité) au sein d'une grande base de cas. D'un point de vue classique, la représentation idéale de cette relation consiste à attribuer à chaque couple « modalité-éventualité », une valeur exacte de la probabilité d'occurrence. Néanmoins, ces valeurs sont rarement connues par les spécialistes. Pour cette raison et afin d'exprimer cette connaissance « imprécise/ambiguë » des valeurs probabilistes, les spécialistes utilisent une description qualitative en utilisant des descripteurs linguistiques naturels. Ce mode de description permet aux experts d'exprimer leur incertitude en utilisant des termes linguistiques, plus significatifs que les valeurs numériques utilisées dans la théorie classique des probabilités.

En considérant le cadre incertain, les connaissances liées à des classes sont directement données sous la forme de modèles incertains définies sur l'espace des primitives permettant la description des classes. L'ensemble de ces modèles forme la base de connaissance. La Figure 3.2 illustre une base de connaissances de modèles possibilistes.

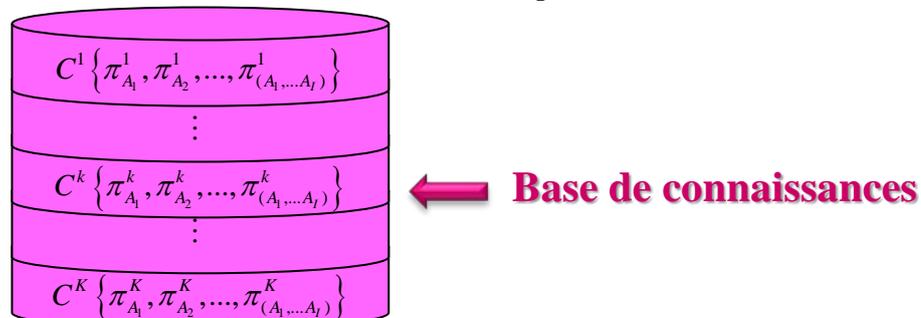


Figure 3.2 Base de connaissances de modèles possibilistes.

2. Processus d'association de cas ou de connaissances à des décisions

On se positionne sous l'hypothèse du monde fermé dans lequel la relation liant chaque cas ou connaissance à sa décision devrait être exhaustive (on est sûr que la décision existe) et exclusive (il n'existe qu'une seule décision à associer à ce cas ou cette connaissance). Cette relation est alors une fonction informative permettant l'attribution d'une étiquette décision ou classe C^k à chaque cas ou connaissance. On parle dans ce sens, de techniques de classification, de reconnaissance, d'identification, d'arbres de décision, d'algorithmes d'apprentissage et de fouilles de données tels que les règles d'associations ou encore les réseaux bayésiens et autres. Il convient ici de distinguer entre la notion de classification thématique et celle sémantique. Pour l'expliquer, nous proposons de prendre le cas de l'image :

La classification thématique du pixel peut précéder la répartition de l'image en régions homogènes (ou objets sémantiques)

En effet, la formation des régions homogènes peut être réalisée à deux niveaux, tel que représenté dans la Figure 3.3:

- Niveau capteur : En cherchant les groupes de pixels homogènes dans l'espace des données brutes en sorti du capteur (i.e. niveau radiométrique, ou niveau de gris). A ce niveau, les régions formées ne portent aucune « étiquette » sémantique ou thématique.
- Niveau sémantique : Ce niveau abouti à des régions étiquetées mais il nécessite de réaliser, d'abord, un processus de classification pixellique. En effet, une fois les données brutes sont classifiées, la formation des régions est réalisée par regroupement des pixels portant la même étiquette de classe.

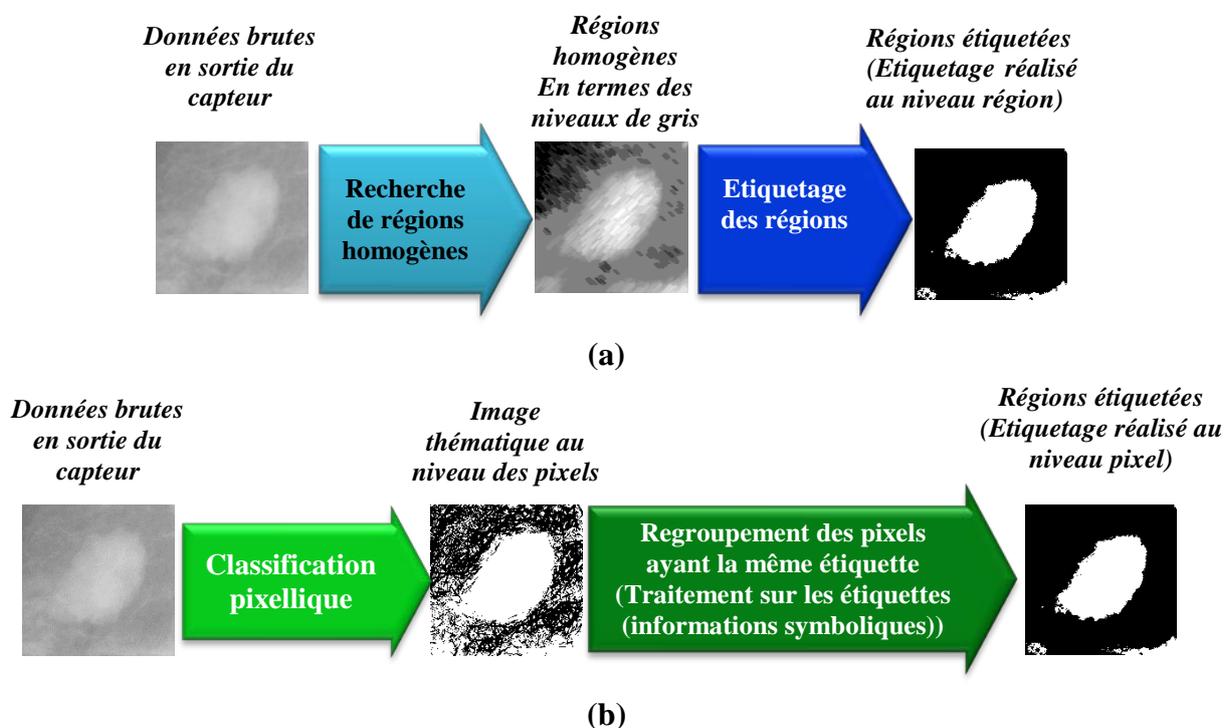


Figure 3.3 Les deux niveaux de formation des régions homogènes :
(a) Niveau capteur ; (b) Niveau sémantique

Les deux niveaux (capteur et sémantique) aboutissent à la formation de régions de deux manières différentes, avec deux approches de raisonnement différentes et selon des sources de connaissance différentes.

En effet, la formation des régions homogènes au niveau capteur passe d'abord par la formation de régions homogènes en niveaux de gris, et accomplit ensuite l'étiquetage des régions. Alors que la formation des régions homogènes au niveau sémantique nécessite d'abord une phase préliminaire de classification pixellique, suivie d'une phase de formation des régions homogènes en se basant, d'une part, sur le critère d'homogénéité, considéré comme celui des pixels portant la même étiquette, et d'autre part sur des critères liés à la sémantique globale de la scène, du type suppression de petites régions....

A ce niveau, il est nécessaire de distinguer entre des « classes thématiques » et des « classes sémantiques ». En effet, une classification « thématique » dans une image, se réfère à une classification des pixels se basant principalement sur une connaissance *a priori* de leurs réponses radiométriques, ou éventuellement de quelques indices de texture ou de voisinage. Alors qu'une classification « sémantique » dans une image, s'applique sur les régions de l'image de façon à aboutir à une répartition de l'image en objets sémantiques. A titre d'exemple on peut définir : Herbe en tant que « Classe thématique », végétation en tant que « Classe Sémantique »

Le défi des techniques de l'intelligence artificielle pour la classification ou d'autres formes de prises de décision, est la recherche des formes d'homogénéité. Ce qui se réfère dans la plupart des cas à une spécification d'une forme de similarité entre les entités assemblées dans des catégories identiques ou différentes. La recherche de la similarité a attiré l'attention de plusieurs chercheurs et elle a marqué son utilisation dans de nombreux domaines. Le choix d'une mesure de similarité adéquate à l'application aux données et à l'attente, présente un critère déterminant dans l'élaboration des modèles de prise de décision.

Nous présentons dans ce chapitre des travaux et des contributions en relation avec la thématique de la prise de décision dans un contexte possibiliste. Notamment des travaux sur les mesures de similarité, les arbres de décision, la fouille possibiliste et la fusion. Une synthèse sur les collaborations et les activités effectuées dans cette thématique est d'abord présentée.

II. Synthèse des activités relatives à ce chapitre

Certaines contributions et travaux effectués dans cette thématique apparaissent dans la majorité des projets et articles cités dans le chapitre précédents, et pratiquement dans le même cadre des collaborations. Ceci est dû au fait que l'étape de prise de décision suit naturellement celle de modélisation. Je cite dans ce qui suit ceux qui ne sont pas encore mentionnés.

1. Collaborations et encadrements

Mes travaux sur la prise de décision dans des contextes possibilistes sont essentiellement réalisés en collaboration avec le Pr. Basel Solaiman¹. Nous travaillons actuellement sur l'élaboration d'un système dynamique incrémental pour une fusion possibiliste multimodale. Dans ce cadre je co-encadre le projet de thèse suivant:

1. Amal Charfi, « Système de reconnaissance possibiliste dynamique appliqué à la biométrie multimodale », thèse en co-tutelle avec iTi de l'IMT-Atlantique, Brest,

¹ Professeur et directeur du département iTi à IMT Atlantique, Bretagne, France.

France. La thèse est sous la direction conjointe de Mr. Nabil Derbel, Professeur de l'Ecole Nationale d'Ingénieurs de Sfax. Ecole doctorale: Ingénierie des systèmes informatiques (ISI), ENIS et Mr. Basel Solaiman, Professeur de l'IMT-Atlantique, Brest, France.

Une collaboration a été entretenue avec Dr. Issam Rebaï², pour élaborer une approche de reconnaissance sémantique d'objets par fusion d'arbre de décision possibiliste une ontologie de l'objet, et ce, dans le cadre d'un encadrement partagé du mastère de recherche suivant:

2. Wafa Damak, "Reconnaissance et classification des objets dans un environnement domestique". Présenté à l'Institut Supérieur d'Informatique et de Multimédia En vue de l'obtention du diplôme de Mastère De Recherche en Informatique et Multimédia (I2M), Soutenu le 18/04/2014.

Une étude approfondie de l'état de l'art sur les mesures de similarité et leurs emplois spécialement pour l'élaboration d'arbres de décisions dans un contexte possibiliste, a été élaboré dans le cadre du projet de mastère de recherche suivant :

3. Amal CHARFI, "Evaluation de mesures de similarité en vue de leur utilisation dans des arbres de décision dans un cadre possibiliste". Présenté à L'Institut Supérieur d'Informatique et de Multimédia En vue de l'obtention du diplôme de Mastère De Recherche en Informatique et Multimédia (I2M), Soutenu le 30/10/2014.

Une proposition d'une nouvelle mesure de similarité ainsi qu'une méthodologie d'évaluation s'en sont suivis. Le travail est en cours de publication.

Des travaux sur les critères d'évaluation de la décision et l'évaluation et la sélection des opérateurs de fusion sont en cours, dans le cadre d'une collaboration avec Pr. Basel Solaiman, Pr. Eloi Bossé et Dr. Sonda Ammar.

2. Publications

Les résultats de certains de ces travaux ont été publiés à travers 4 articles de conférences internationales. D'autres sont en cours de production.

1. S. Ammar Bouhamed, M. Chakroun, **I. Khanfir Kallel**, H. Derbel, "Haralick feature selection for Material rigidity recognition using ultrasound echo", International Conference on Advanced Technologies for Signal & Image Processing, Sousse, 2018.
2. S. Ammar Bouhamed, **I. Khanfir Kallel**, "Merging Operator Selection in Possibility Theory", 14th International Multi-Conference on Systems, Signals and Devices (SSD'17), March 28 - 31, 2017 - Marrakech, Morocco, pp. 772-778.
3. W. Damak, **I. Khanfir** and I. Rebai, "Semantic object recognition by merging decision tree with object ontology". 1st IEEE International Conference on Advanced Technologies for Signal & Image Processing ATSSIP 2014, March 17-19, 2014, Sousse-Tunisia.
4. S. Ammar Bouhamed, **I. Khanfir Kallel**, D. Sellami Masmoudi and B. Solaiman, "Possibilistic modeling of ultrasonic signal for floor state recognition". 1st IEEE International Conference on Advanced Technologies for Signal & Image Processing ATSSIP 2014, March 17-19, 2014, Sousse-Tunisia.

² Enseignant chercheur au département informatique de l'IMT Atlantique, Bretagne, France

Un travail coopératif sur la fusion multi-sources, basée sur la qualité des sous-ensembles à combiner, a été entretenu avec Dr. Sonda Ammar, Pr. Eloi Bossé, Pr. Ronald Yager et Pr. Basel Solaiman. Il s'agit d'une étude comparative entre l'approche de fusion probabiliste proposée par Yager dans [YF 2016] et sa reprise par la théorie des possibilités, qui a montré une supériorité considérable dans les situations où l'on dispose de données faiblement définies et provenant de sources possibilistes. En effet, l'information provenant de sources possibilistes se présente sous forme de témoignages au contraire de celle provenant de sources probabilistes qui permettent d'avoir des statistiques fiables. Ce travail a été récemment soumis à un journal international renommé.

III. Mesure de similarité

Tout système ayant pour but d'analyser ou d'organiser automatiquement un ensemble de données ou de connaissances (système de fouille des données, par exemple) doit utiliser sous une forme ou une autre, une mesure de similarité dont le but est d'établir les ressemblances ou les relations qui existent entre les informations manipulées [Bis 2000]. Quantifier ces ressemblances par une valeur numérique, puis s'assurer que la valeur estimée reflète la réalité physique, n'est pas évident. Ces difficultés sont liées à trois concepts fondamentaux :

✓ *Le concept physique de la notion de similarité*

L'univers est fondé sur la base des différences, ce qui explique qu'une similarité absolue n'existe pas et reste toujours relative mais avec différents ordres de fermeté. La question est alors: « relative par rapport à quoi? »

Considérons un univers d'éléments observé par différentes sources d'information (capteurs ou humains) et décrit par des attributs (primitives, caractéristiques). Les éléments de l'univers sont à affecter à des partitions selon leurs ressemblances, selon la règle « ceux qui se ressemblent s'assemblent ». Leurs degrés de similarité sont relatifs, principalement à deux notions :

- 1- La partition visée de l'univers d'éléments. Exemple (Homme/Chimpanzé; Homme/Homme; Homme/Fils; Homme/lui-même)
- 2- La nature des sources utilisées pour l'«observation» et des attributs utilisés pour la description. En effet, il existe une possibilité d'avoir : des conflits à gérer, des préférences à accorder, etc.

✓ *L'estimation de la similarité*

Selon [BRL 2008], l'estimation de la similarité se base essentiellement sur deux concepts: géométrique et ensembliste. Les mesures de similarité basées sur le point de vue géométrique, appelées aussi mesures de similarité métriques, considèrent les informations à comparer comme étant des points dans un espace métrique de représentation. Ces mesures réalisent alors une transformation de la distance estimée entre les informations à comparer en une mesure de similarité. Quant aux mesures de similarité ensemblistes, elles assimilent une information représentant un objet à un sous-ensemble algébrique des primitives. Ainsi, les mesures de similarité ensemblistes mesurent le nombre de primitives communes ou distinctes entre les informations à comparer. Sur la base de ces deux catégories de mesures de similarité, plusieurs travaux de recherche ont proposé des généralisations par l'intégration des fonctions d'appartenance aux ensembles flous, afin de réaliser la comparaison entre informations

ambiguës représentées comme des ensembles flous [BRL 2008]. Les mesures de similarité ainsi obtenues sont appelées mesures de similarité floues.

Une mesure de similarité entre deux objets ou informations, fournit habituellement une valeur numérique quantifiant le degré de similarité entre elles. La question qui se pose ici « est-ce qu'une proximité métrique quantifiée par une valeur numérique coïncide toujours avec une proximité sémantique ? ». Les performances d'une mesure de similarité dépendent de plusieurs facteurs. On note surtout, la qualité de la représentation (ou modélisation) des connaissances reliées aux entités (objet/classe, objet/objet, classe/classe) à apparier, et encore le concept mathématique de l'approche d'estimation et de calcul de la similarité.

✓ *Les critères d'évaluation de la mesure de similarité*

Pour juger correctement les performances d'une mesure de similarité, il faut se poser dans des conditions de test permettant d'optimiser les influences des autres facteurs sur ses vraies performances. Les facteurs sont reliés principalement aux modèles adoptés pour représenter les entités à apparier, le(s) attribut(s) choisi(s) pour décrire les entités et la nature même des entités.

1. Notion de similarité dans un contexte possibiliste

Le concept de similarité a été vastement étudié dans la littérature, et les mesures proposées sont tellement nombreuses que l'on tombe dans l'embarras du choix dès qu'on a besoin de s'en servir. Cependant, ces mesures n'offrent pas les mêmes performances en dépit de la diversité des circonstances. Dans le cas où les informations à comparer sont de nature possibiliste (i.e. chaque information est représentée par une distribution de possibilités), l'utilisation des mesures de similarité classiques ou communes, risquent de ne pas révéler la ressemblance ou d'exprimer la relation entre ces connaissances. Selon [Als 2014], l'immense majorité des travaux existants, propose de transposer les mesures de similarité floues en substituant les fonctions d'appartenance aux ensembles flous par les distributions de possibilités. Ce point de vue, malgré sa profusion, néglige le fossé sémantique en terme de différence de signification entre une fonction d'appartenance (représentant un ensemble flou) et une distribution de possibilités (encodant la possibilité que chaque singleton soit l'unique qui s'est produit). Notre objectif pour cette activité de recherche est d'étudier la similarité entre des connaissances possibilistes, tout en prenant en considération la particularité de l'encodage possibiliste des connaissances. En effet, nous considérons que la forme d'une distribution de possibilité renseigne sur sa valeur informative. Ainsi une ressemblance entre les formes de deux distributions de possibilités peut conduire à une déduction sur la proximité entre leurs informations respectivement encodées. De cette révélation nous proposons de grouper les mesures de similarités possibilistes en deux catégories : celles qui sont basées sur une évaluation de l'aspect morphologique (forme) des distributions à apparier et celles qui ne le sont pas. La première catégorie regroupe les mesures de similarité qui se basent sur des concepts de quantification de l'information, généralement en se basant sur des mesures d'incertitude. Quant à la seconde catégorie, elle regroupe les mesures de similarité qui sont conçues par des généralisations des mesures métriques.

Pour entamer cette recherche, une large investigation de la littérature a été d'abord effectuée en quête des mesures proposées pour évaluer des similarités entre des distributions de possibilités. Dans ce qui suit, nous présentons d'abord, très brièvement, quelques propriétés de base d'une mesure de similarité possibiliste selon [JBE 2007]. Un bref résumé

de l'analyse critique effectuée dans [Char 2014] sur certaines approches de l'état de l'art est donné. Dans le but de répondre à des lacunes identifiées pour ces approches, nous proposons une nouvelle mesure de similarité qui prend en compte aussi bien des ressemblances métriques que de forme entre deux distributions de possibilités. Une méthodologie pour l'évaluer et la comparer à d'autres approches est enfin proposée puis employée.

2. Propriétés d'une mesure de similarité possibiliste

De nombreuses propriétés d'une mesure de similarité ont été évoquées dans la littérature. Dans un cadre possibiliste, il peut être convenable d'adopter celles qui ont été retenues pour les mesures de similarité floues [BRB 1996, CS 2017, MCB 2004]. Cependant, dans ce qui suit, nous préférons présenter les propriétés qui ont été retenues dans [JBE 2007] comme étant les propriétés de base pour une mesure de similarité possibiliste.

Soient π_k , $k=1..n$ des distributions de possibilités modélisant n entités appartenant à un même univers de discours $\Omega = \{\omega_1, \dots, \omega_n\}$. Notons $s(\pi_1, \pi_2)$ une mesure de similarité possibiliste.

✓ **Propriété 1: Non-négativité**

$$s(\pi_1, \pi_2) \geq 0$$

✓ **Propriété 2: symétrie**

$$s(\pi_1, \pi_2) = s(\pi_2, \pi_1)$$

✓ **Propriété 3: Limite supérieure**

$$\forall \pi_i, s(\pi_i, \pi_i) = 1 \text{ et } \forall \pi_i, \pi_j, s(\pi_i, \pi_j) \leq 1 \text{ (distribution normalisée)}$$

✓ **Propriété 4: Limite inférieure**

Si $\forall \omega_i \in \Omega$,

i) $\pi_1(\omega_i) \in \{0, 1\}$, $\pi_2(\omega_i) \in \{0, 1\}$, et

ii) Si $\pi_2(\omega_i) = 1 - \pi_1(\omega_i)$ alors $s(\pi_1, \pi_2) = 0$

Les distributions de possibilité les plus contradictoires ont le plus faible degré de similitude.

✓ **Propriété 5: La Large inclusion (spécificité)**

$$\text{Si } \forall \omega_i \in \Omega, \pi_1(\omega_i) \leq \pi_2(\omega_i), \pi_2(\omega_i) \leq \pi_3(\omega_i), \\ \text{alors } s(\pi_1, \pi_2) \geq s(\pi_1, \pi_3).$$

✓ **Propriété 6: Permutation**

Soit ρ une permutation d'indice de π_1 et π_2 alors, $\forall \pi_1, \pi_2, s(\pi_1, \pi_2) = s(\pi_{\rho(1)}, \pi_{\rho(2)})$.

La similarité entre deux distributions de possibilités π_1 et π_2 ne change pas si l'ordre des éléments de même indice change.

3. Etat de l'art sur les mesures de similarité en théorie de possibilités

Dans la théorie des possibilités, il existe peu de travaux consacrés à l'estimation de similarité entre des distributions de possibilités. Dans ce qui suit nous présentons brièvement des mesures de similarité, parmi les plus abordées et utilisées dans la littérature [HK 1983, SCC 1998, ASB 2016, SIE 2017, AE 2017a-b, KAS 2018, KAA 2018, Krou 2006, SHC 2009].

a) Mesure de similarité « Information Closeness »

La mesure proposée dans [HK 1983] est donnée par:

$$S_G = 1 - \frac{G(\pi_i, \pi_j) - G_{\min}}{G_{\max}}, \text{ avec :} \quad \text{Eq3.1}$$

$$G(\pi_1, \pi_2) = 2 * U(\pi_1 \vee \pi_2) - U(\pi_1) - U(\pi_2)$$

où \vee est l'opérateur maximum, U est la mesure d'incertitude, $G_{\min} = 0$ et $G_{\max} = 2 * \log_2(|\Omega|) - \log_2(|\Omega| - 1)$.

S_G a bien les valeurs dans l'intervalle $[0,1]$

b) Mesure de similarité déduite de l'information de divergence

La divergence D proposée dans [Krou 2006] est définie comme l'intégrale discrète de Choquet des degrés de la distribution de différence π_d :

$$D(\pi_1 | \pi_2) = \sum_{i=1}^n \pi_d(w_{\sigma(i)}) [\Pi_1(A_{\sigma(i)}) - \Pi_1(A_{\sigma(i+1)})] \quad \text{Eq3.2}$$

Avec :

$\pi_d(\omega) = |\pi_1(\omega_i) - \pi_2(\omega_i)|$ pour chaque $\omega_i \in \Omega$.

σ est une permutation des indices tel que :

- $\pi_d(w_{\sigma(i)}) \leq \dots \leq \pi_d(w_{\sigma(n)})$ et
- $A_{\sigma(i)} = \{w_{\sigma(i)}, \dots, w_{\sigma(n)}\}$, $i = 1..n$ et $A_{\sigma(n+1)} = \emptyset$

La mesure de similarité S_D correspondante à D et ayant des valeurs dans l'intervalle $[0,1]$ est alors :

$$S_D = 1 - D(\pi_i | \pi_j) \quad \text{Eq3.3}$$

Il est clair que cette mesure est non symétrique, ce qui implique qu'elle ne satisfasse pas la propriété 2.

c) Mesure de Sangüesa :

Elle est définie comme la mesure de l'incertitude de la différence entre les distributions [SCC 1998].

$$S(\pi_1, \pi_2) = U(\pi_d) / \log_2(|\Omega|) \quad \text{Eq3.4}$$

d) Distance de Minkowski et ses dérivés

La mesure de distance point par point entre deux distributions est aussi utilisée dans le but d'évaluer la similarité entre des distributions de possibilités.

Soient deux distributions de possibilité π_1 et π_2 sur $\Omega = \{\omega_1, \dots, \omega_n\}$

- ✓ La distance de Minkowski non normalisée est donnée par :

$$D_M(\pi_1, \pi_2) = \sqrt[p]{\sum_{i=1}^n |(\pi_1(\omega_i) - \pi_2(\omega_i))|^p} \quad \text{Eq3.5}$$

- ✓ La similarité basée sur la distance de Manhattan normalisée est donnée par :

$$S_M(\pi_1, \pi_2) = 1 - \frac{\sum_{i=1}^n |\pi_1(\omega_i) - \pi_2(\omega_i)|}{n} \quad \text{Eq3.6}$$

- ✓ La similarité basée sur la distance Euclidienne normalisée est donnée par :

$$S_E(\pi_1, \pi_2) = 1 - \sqrt{\frac{\sum_{i=1}^n (\pi_1(\omega_i) - \pi_2(\omega_i))^2}{n}} \quad \text{Eq3.7}$$

- ✓ La similarité basée sur la distance maximale normalisée est donnée par :

$$S_C(\pi_1, \pi_2) = 1 - \frac{\max_{i=1}^n (\pi_1(\omega_i) - \pi_2(\omega_i))}{n} \quad \text{Eq3.8}$$

e) Mesure de similarité delta δ

La mesure de similarité δ [SHC 2009], entre deux distributions π_1 et π_2 , est définie par :

$$\delta(\pi_1, \pi_2) = \frac{\sum_{i=1}^N \min(\pi_1(x_i), \pi_2(x_i))}{\sum_{i=1}^N \max(\pi_1(x_i), \pi_2(x_i))} \quad \text{Eq3.9}$$

f) Mesure de similarité « Info Affinity »

La mesure de similarité, *Info Affinity* [JBE 2007], est une mesure proposée dans le but de prendre en compte le concept de l'inconsistance. En fait, il s'agit de combiner la mesure de distance et la valeur de l'inconsistance des deux distributions à comparer.

$$InfoAff(\pi_1, \pi_2) = 1 - \frac{d(\pi_1, \pi_2) + Inc(\pi_1, \pi_2)}{2} \quad \text{Eq3.10}$$

Avec $d(\pi_1, \pi_2)$: représente une distance, par exemple la distance de Manhattan entre les deux distributions π_1 et π_2 .

4. Analyse de l'état de l'art

Il est certain qu'une simulation numérique d'un exemple ne peut former un fondement objectif pour analyser le comportement d'une approche ou encore pour comparer plusieurs approches. Cependant, un tel test peut attirer l'attention sur des pistes à explorer.

a) Simulation comparative sur des exemples numériques

Nous visons à travers cette simulation numérique, la vérification du comportement des mesures de similarité citées ci haut, envers différents états de connaissance relatifs au degré de proximité de deux distributions de possibilités. Nous testons, l'égalité, un cas de forte proximité, un cas de divergence informationnelle, un cas de désaccord et finalement un cas de contradiction. Le Tableau 3.1 donne des exemples numériques illustrant ces différents états de connaissances, ainsi que les valeurs de similarités fournies par chacune des approches pour chaque état.

b) Critique

- ✓ *Information Closeness* a bien détecté l'égalité, la forte proximité et le désaccord, mais pas la contradiction. Elle a été induite en erreur par l'information de l'ignorance totale exprimée par la distribution2 dans le cas de la divergence informationnelle. Ces défaillances sont dues au fait que cette mesure est complètement basée sur la notion

d'incertitude donnant la quantité d'information sans se référer à l'écart métrique (point à point) entre les distributions.

Tableau 3.1 Estimation de similarités par différentes approches sur différent cas d'état de connaissance

	<i>Egalité</i>	<i>Forte proximité</i>	<i>Divergence informationnelle</i>	<i>Désaccord</i>	<i>Contradiction</i>
<i>Distribution 1</i>	<1 0 0 0>	<1 0 0 0>	<1 0.5 0.5 0>	<0 1 1 1>	<1 0 1 0>
<i>Distribution 2</i>	<1 0 0 0>	<1 0.3 0.2 0.3>	<1 1 1 1>	<1 0.3 0.2 0.3>	<0 1 0 1>
<i>Information Closeness</i>	1	0.768	0.585	0.231	0.179
<i>Sangüesa</i>	1	0.898	1	0.25	1
<i>Information divergence</i>	1	1	0	0	0
S_M	1	0.8	0.5	0.2	0
S_E	1	0.78	0.5	0.190	0
S_C	1	0.7	0	0	0
<i>Delta</i>	1	0.555	0.333	0.2	0
<i>InfoAffinity</i>	1	0.9	0.75	0.25	0

- ✓ *Sangüesa* est une mesure basée sur l'incertitude de l'écart point par point des distributions. Dans les trois cas : d'égalité, de divergence informationnelle et de contradiction, les écarts entre les distributions, correspondaient à des distributions de totales ignorances, ce qui explique l'obtention des valeurs de l'unité pour ces trois cas.
- ✓ *Information divergence* est une mesure catégorique ne reflétant qu'une réalité relative à des cas extrêmes.
- ✓ Les mesures de similarité basées sur la distance de *Minkowski* donnent des valeurs raisonnables si nous jugeons leurs performances de point de vue écart métrique entre les distributions (càd, point par point). En revanche, si nous examinons ces mesures sur le plan forme (de point de vue quantité d'information), elles ne paraissent pas satisfaisantes. En effet, le cas de la divergence informationnelle est en quelque sorte un piège du fait que la distribution 2 est une totale ignorance donc sur le plan informationnel, sa similarité avec la distribution 1 (qui montre la présence d'une information) devrait être plutôt faible.
- ✓ La mesure *Delta* n'a pas bien exprimé l'état de forte proximité.
- ✓ La mesure *InfoAffinity*, combine l'emploi de la notion de distance point à point de Manhattan et l'emploi de la notion de l'inconsistance qui décrit, d'une certaine façon, la quantité d'information. Les valeurs estimatives fournies sont raisonnables avec une réserve envers le cas de l'appariement avec la distribution d'ignorance totale.

c) Motivation

La majorité des mesures étudiées de l'état de l'art ne sont pas toujours convaincantes envers la fidèle révélation des différents états de connaissance. Leurs défaillances viennent du fait qu'elles se basent exclusivement sur des estimations soit d'écarts métriques point à point, soit de dissimilarité morphologique, se référant à la différence des quantités d'information, estimées par des notions spécifiques (telle que l'incertitude pour *Sangüesa*). Notre objectif est justement de proposer une mesure qui prenne en compte les deux types d'écarts. La mesure *InfoAffinity*, adopte ce principe par l'emploi conjonctif de la distance de Manhattan et

l'inconsistance. Cependant la notion d'inconsistance ne semble pas bien gérer la situation dans les cas où les distributions à apparier disposent de la même hauteur ou présentent le même écart métrique point à point. Un exemple numérique explicatif est détaillé dans [Char 2014] (Exemple 2-3).

5. Contribution : Proposition d'une nouvelle mesure de similarité « Info-spécificity »

a) Principe

La mesure que nous proposons consiste à combiner deux mesures, l'une métrique et l'autre basée sur la quantité d'information (ou sur la forme) avec différentes proportions dépendant des connaissances à apparier. Pour la mesure métrique, la distance de Manhattan nous semble bien convenable. Quant à la mesure basée sur la forme, nous proposons de calculer la différence de la quantité d'information, contenue dans chacune des distributions à apparier. L'une des plus reconnues mesures pour quantifier le contenu informationnel dans une distribution de possibilité, est celle définie par [Yag 1983], à savoir la mesure de spécificité (Sp). Sp est définie pour $\pi(\Omega) \in [0, 1]$, par:

$$Sp(\pi) = \pi(1) - \sum_{j=2}^N p_j \pi(\omega_j) \quad \text{Eq3.11}$$

avec p_j est un poids tel que $p_j \in [0,1]$, il doit satisfaire: $p_j \geq p_i \forall 1 < j < i$ et $\sum_{j=2}^N p_j = 1$.

Nous estimons alors l'écart entre quantités d'information, par la différence des mesures de spécificité de chacune des distributions appariées.

b) Définition

Soit π_1, π_2 deux distributions appartenant au même univers de discours Ω . On définit la mesure « *InfoSpecificity* » comme suit :

$$InfoSpecificity(\pi_1, \pi_2) = \begin{cases} 1 - \frac{\alpha * dist_{Manh}(\pi_1, \pi_2) + \beta * dist_{Sp}(\pi_1, \pi_2)}{\alpha + \beta} & \text{if } dist_{Manh} \neq 0 \\ 0 & \text{if } dist_{Manh} = 0 \end{cases} \quad \text{Eq3.12}$$

Tel que :

- ✓ $dist_{Sp}(\pi_1, \pi_2) = |Sp(\pi_1) - Sp(\pi_2)|$
- ✓ $dist_{Manh}(\pi_1, \pi_2)$ est la distance de Manhattan normalisée et
- ✓ α et β sont deux coefficients tels que $\alpha \in [0, 1]$, $\beta \in [0, 1]$.

Les paramètres α et β permettent à varier l'importance à accorder à chacun des deux types d'écart selon des règles relatives aux distributions à apparier. Ces règles sont en cours de définition afin d'automatiser complètement le choix de ces paramètres. Pour le présent rapport, nous optons pour un choix neutre de ces deux paramètres. Ils sont fixés à 1.

Selon cette définition, la mesure proposée respecte intégralement les propriétés de base exigées d'une mesure de similarité et énoncées ci-haut. Les preuves mathématiques sont détaillées dans [Char 2014].

6. Méthodologie d'évaluation

a) Facteurs influant sur l'évaluation

- ✓ **La qualité des données** : une qualité médiocre des données peut être due à un fort bruit du capteur, une interférence avec d'autres signaux lors de la capture, ou toute autre altération par des artefacts qui rendent les données peu représentatives de l'entité observée.
- ✓ **La nature de l'attribut** : l'attribut peut être la valeur radiométrique d'un capteur, comme il peut être une transformation de celle-ci, mettant en valeur une caractéristique spécifique contenue dans les valeurs mesurées sur une entité (objet, élément, classe, alternative, etc.). Ainsi l'attribut considéré pour la description de l'entité pourrait déceler les spécificités qui la caractérisent faisant ensuite apparaître des similarités entre ses éléments, comme il peut décrire une caractéristique qui ne la rend pas différentes des autres entités et dans ce cas les dissimilarités entre des entités différentes ne seront pas distinguées à travers cet attribut.
- ✓ **L'approche utilisée pour la modélisation** : influe directement sur la qualité du modèle et sa représentativité de l'entité.
- ✓ **La taille de l'échantillon utilisé pour modéliser l'entité** : se référer à l'étude présentée dans le chapitre 2 dans la section III.

b) Données utilisées pour les expérimentations

On a réalisé les différents tests sur 4 bases de données extraites du répertoire « U.C.I » des bases de données destinées pour le domaine de l'apprentissage automatique (Murphy & Aha, 1996).

• **Base1** : Ensemble de données d'authentification de billets de banque contenant 1372 instances de 10 caractéristiques extraites d'images de billets de banque pour décider si le billet est authentique ou falsifié.

• **Base2** : Ensemble de données de classification des blocs de page consacré à l'analyse d'un document. Il contient 5473 instances qui décrivent chacune un bloc d'un document qui sera identifié en fonction des valeurs de ses 10 attributs dans l'une des 5 classes.

• **Base3** : Ensemble de données sur le diabète chez les Indiens Pima, qui vise à prédire le diabète à 768 personnes et sur la base de 8 attributs.

• **Base4** : Ensemble de données Statlog (silhouettes de véhicules) qui cherche à classer une silhouette donnée parmi 946 instances comme l'un des quatre types de véhicules, en utilisant un ensemble de 18 caractéristiques extraites de la silhouette.

Une récapitulation des données relatives à ces 4 bases est donnée au Tableau 3.2

Pour chacune des 4 bases un unique attribut a été considéré pour la modélisation de leurs classes respectives. En considérant l'influence qu'apporte l'attribut sur l'évaluation des mesures de similarité, nous classifions les attributs en trois catégories envers la qualité d'évaluation à engendrer :

- **Attributs non discriminants**: toutes les mesures de similarité donnent des résultats inutiles.

- **Attributs discriminants**: une évaluation basée sur ce type d'attribut conduit à une comparaison significative entre les mesures de similarité.

- **Attributs très discriminants**: permet aux mesures de similarité de différencier facilement les échantillons de chaque classe. De très bons résultats sont obtenus pour toutes les mesures de similarité.

Un jugement a été porté alors sur les attributs de chaque base vis-à-vis de leurs résultats respectifs de classification par le classifieur SVM de l'application Weka. L'attribut sélectionné pour caractériser les classes de chacune des 4 bases, est donné au Tableau 3.2.

Tableau 3.2 Une récapitulation des données relatives aux 4 bases

<i>Base</i>	<i>Base1</i>	<i>Base2</i>	<i>Base3</i>	<i>Base4</i>
<i>Nombre d'instances</i>	1372	5473	768	946
<i>Nombre d'attributs</i>	10	10	8	18
<i>Nombre de classes</i>	2	5	2	4
<i>Attribut utilisé</i>	« Kurtosis of Wavelet »	« Kurtosis about major axis »	« Percentage of black pixels within the block »	« Diabetes pedigree function »

c) Approches d'estimation de mesures de similarité à comparer

Six approches sont soumises à l'évaluation durant cette étude. Le choix des approches a été conduit en se basant sur leurs manières respectives pour l'estimation des écarts entre les distributions. Le Tableau 3.3 résume le choix effectué

Tableau 3.3 Classification des approches évaluées selon leurs manières d'estimation

	<i>Approche métrique point à point</i>	<i>Approche basée sur la forme</i>	<i>Approche combinée</i>
<i>SM</i>	×		
<i>Delta</i>	×		
<i>Information Closeness</i>		×	
<i>Information divergence</i>		×	
<i>InfoAffinity</i>			×
<i>InfoSpecificity</i>			×

d) Description de la procédure d'évaluation

Une mesure de similarité est jugée performante si :

- Elle affecte des valeurs proches à l'unité aux tests d'appariements entre deux distributions construites d'éléments issus d'une même classe, signifiant ainsi qu'elles sont similaires. L'ensemble de ces distributions est dit « Similaires » ou « Génuines » ou « authentiques ». La Figure 3.4, montre la procédure de collecte des valeurs des tests d'appariement de la catégorie « Similaires ».

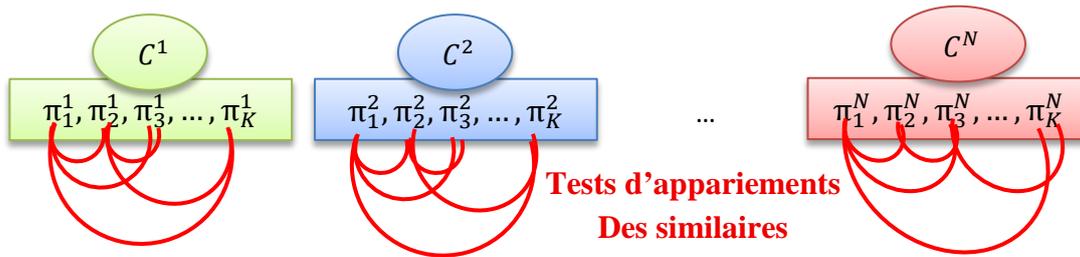


Figure 3.4 Procédure de collecte des valeurs des tests d'appariement de la catégorie « Similaires »

- Elle affecte des valeurs proches de zéro aux tests d'appariement entre deux distributions construites d'éléments issus de classes différentes, signifiant ainsi qu'elles sont distinctes. L'ensemble de ces distributions est dit « Dissimilaires » ou « Faux » ou « Inauthentiques ». La Figure 3.5, montre la procédure de collecte des valeurs des tests d'appariement de la catégorie « Dissimilaires ».

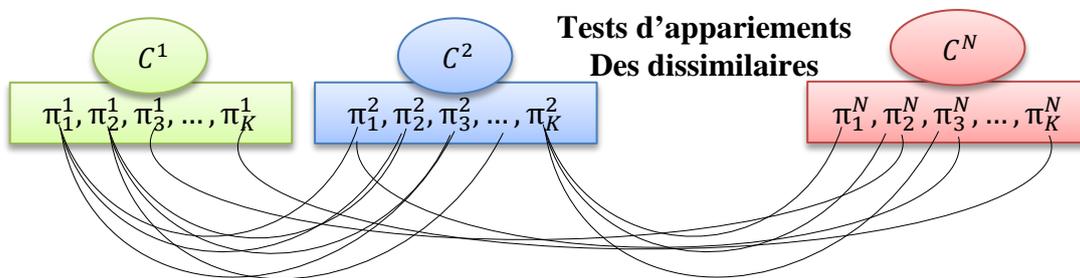


Figure 3.5 Procédure de collecte des valeurs des tests d'appariement de la catégorie « Dissimilaires »

Une représentation des fréquences d'apparition des mesures de similarité dans les deux catégories respectives « Similaires » et « Dissimilaires », sur une échelle de similarité croissante, est donnée par la Figure 3.6.

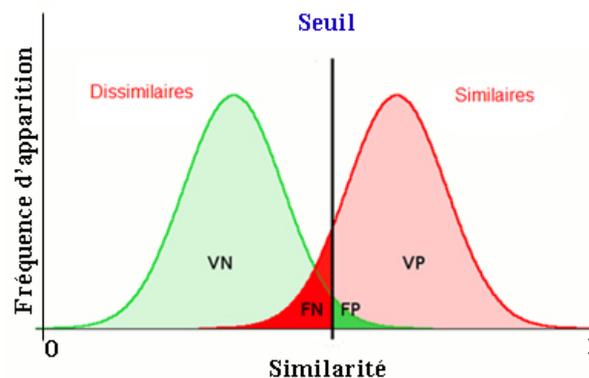


Figure 3.6 Fréquences d'apparition des mesures de similarité dans les deux catégories d'appariement « Similaires » et « Dissimilaires »

Afin de juger le résultat d'un test d'appariement si c'est à considérer du type « Similaire » ou « Dissimilaire », un seuil est à considérer, tel que le montre la Figure 3.6. Des taux de Vrai Positifs (VP), Vrai Négatifs (VN), Faux Positifs (FP) et Faux Négatifs (FN) sont alors à performer en fonction des valeurs du seuil.

Nous utilisons dans ce contexte la courbe ROC (Receiving Operator Characteristics) pour l'évaluation et la comparaison des mesures de similarité. C'est la représentation du taux des vrais positifs (VP) en fonction du taux des faux positifs (FP) tel que le montre la Figure 3.7.

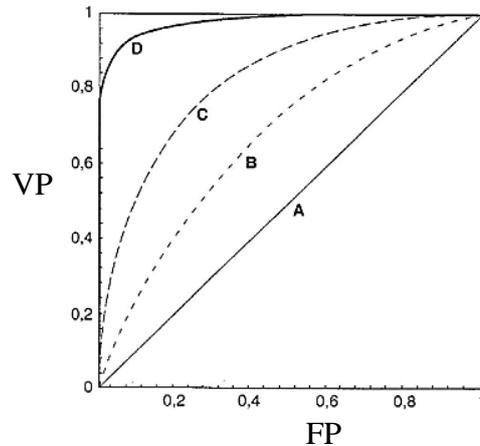


Figure 3.7 Représentations de courbes ROC : (A) Approche sans intérêt, (B) Approche médiocre, (C) Approche meilleure que B, (D) Bonne approche

Le critère d'évaluation adopté pour cette étude est l'aire sous la courbe appelée AUC (Area Under the Curve). La valeur d'AUC varie entre 0.5 et 1. Une mesure de similarité est d'autant plus performante que sa valeur d'AUC est proche de 1. Il est à noter, qu'à partir d'une valeur d'AUC=0.8, l'approche est jugée performante.

Nous proposons la procédure d'évaluation suivante :

Pour chacune des 4 bases de données, faire :

- 1- Construction d'une base de distributions de possibilités. La création de la similarité entre des distributions se fait selon deux manières, engendrant deux tests :

➤ **Test1 :**

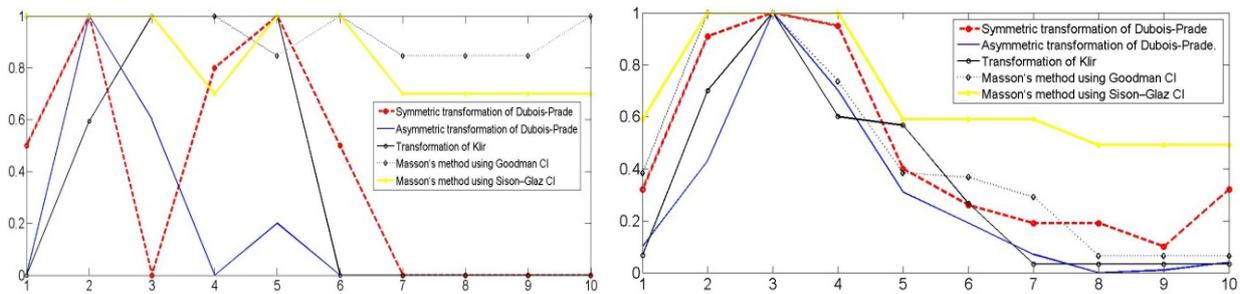
Création de la similarité par la construction de distributions de possibilités à partir d'éléments issus d'une même classe. Pour ce test, 10 tailles d'échantillons (de 10 à 100 à pas de 10) sont considérées afin de varier l'efficacité de la représentativité et tester ainsi le comportement des mesures de similarité envers la dégradation de la représentativité du modèle.

➤ **Test2 :**

Création de la similarité par la construction de distributions de possibilités à partir d'un même échantillon par 5 approches d'estimation de distributions de possibilités (Cf. Figure 3.8), à savoir :

- Transformation symétrique de Dubois et Prade [DP 1988]
- Transformation asymétrique de Dubois et Prade [DPS 1993]
- Transformation de Klir [Kli 1993]
- Technique de Masson et Denoeux (intervalles de confiance de Goodman) [MD 2006]
- Technique de Masson et Denoeux modifiées (intervalles de confiance de Sison & Glaz) [MD 2006]

- 2- Pour chaque test, le scénario donné par le graphe de la Figure 3.4 est ensuite effectué pour chacune des 6 approches de mesure de similarité à comparer.



(a) Très petite taille de l'échantillon

(b) Grande taille de l'échantillon

Figure 3.8 Distributions de possibilités construites pour la modélisation de la même classe de la base de données "Page blocks", par 5 approches d'estimation, à partir d'un échantillon de taille (a) 10 et (b) 100

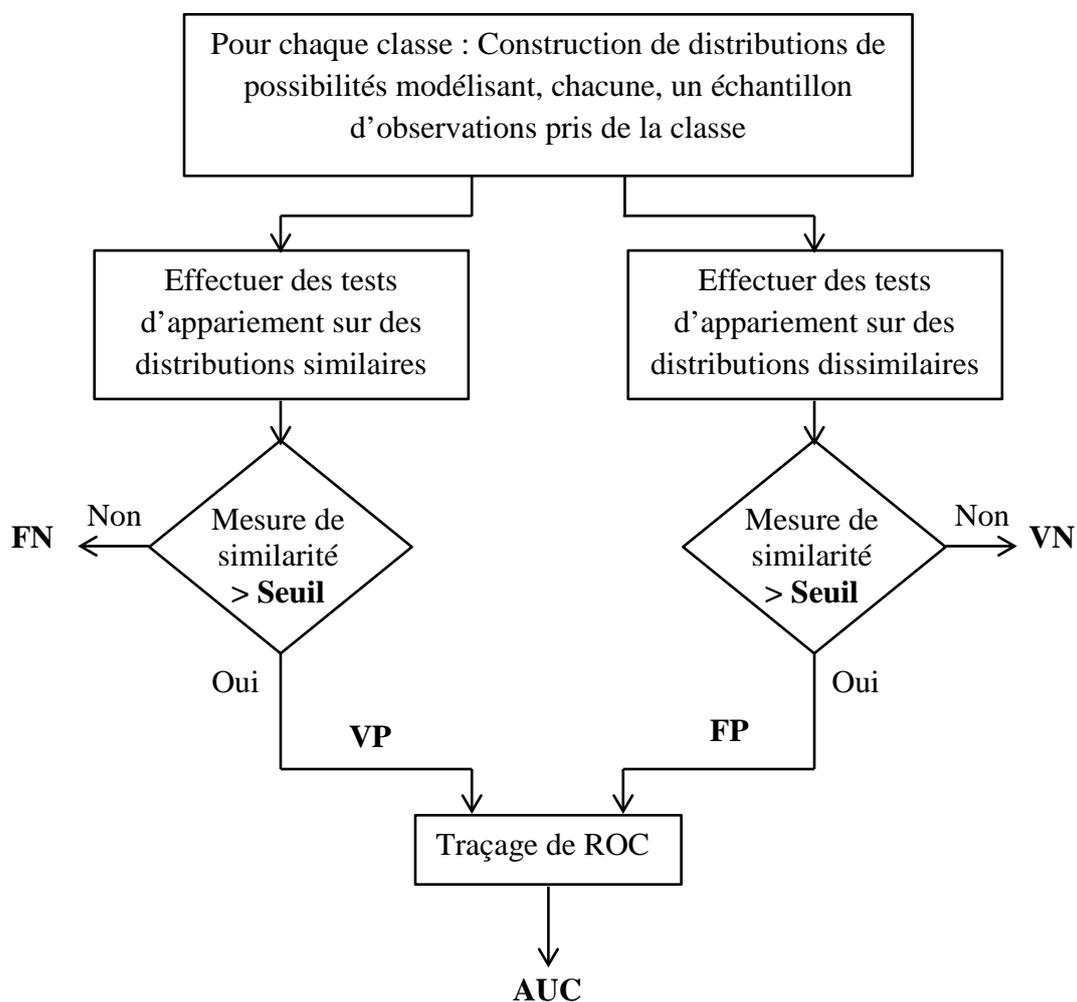


Figure 3.9 Récapitulation de la procédure d'évaluation des techniques de mesure de similarités entre des distributions de possibilités

7. Résultats et discussions

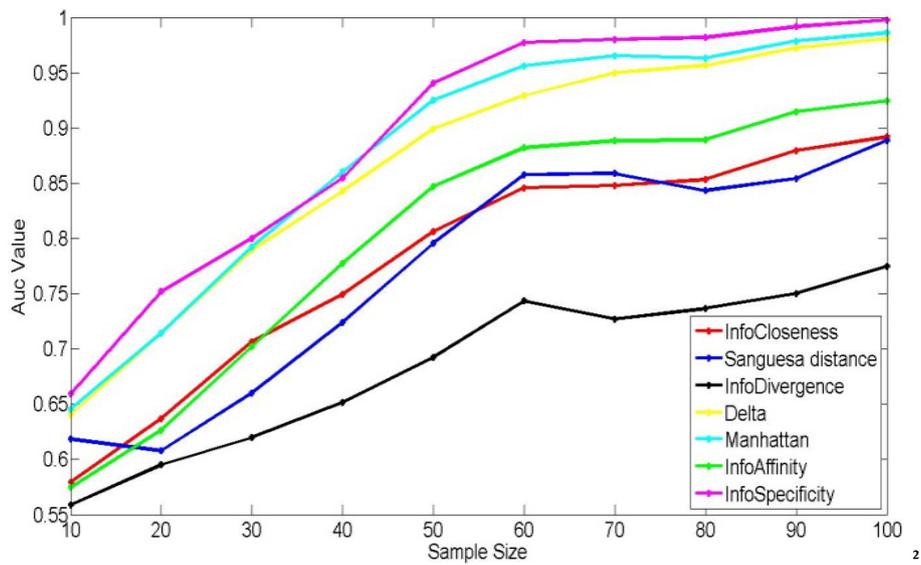


Figure 3.10 Comparaison des mesures de similarité, basée sur les valeurs de l'aire sous la courbe ROC (AUC) pour la base de données « Banknote authentication »

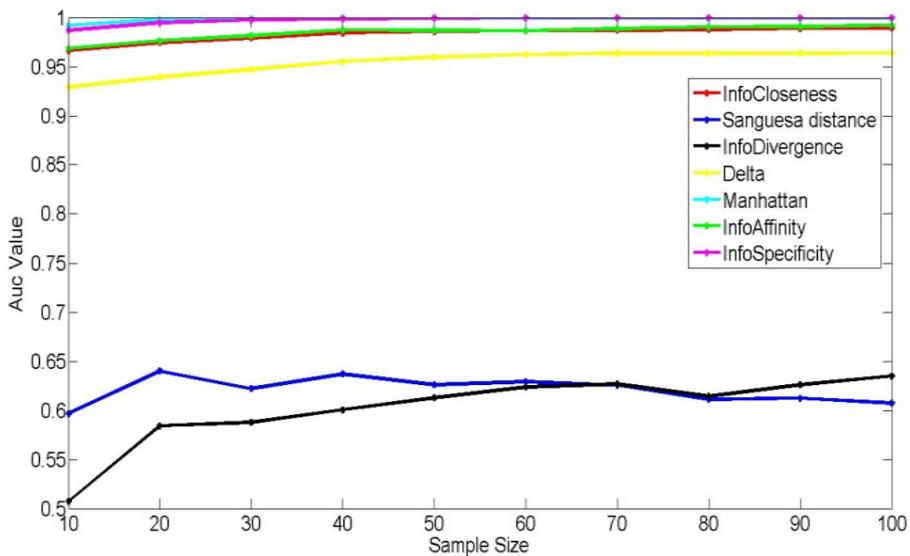


Figure 3.11 Comparaison des mesures de similarité, basée sur les valeurs de l'aire sous la courbe ROC (AUC) pour la base de données « Page Blocks »

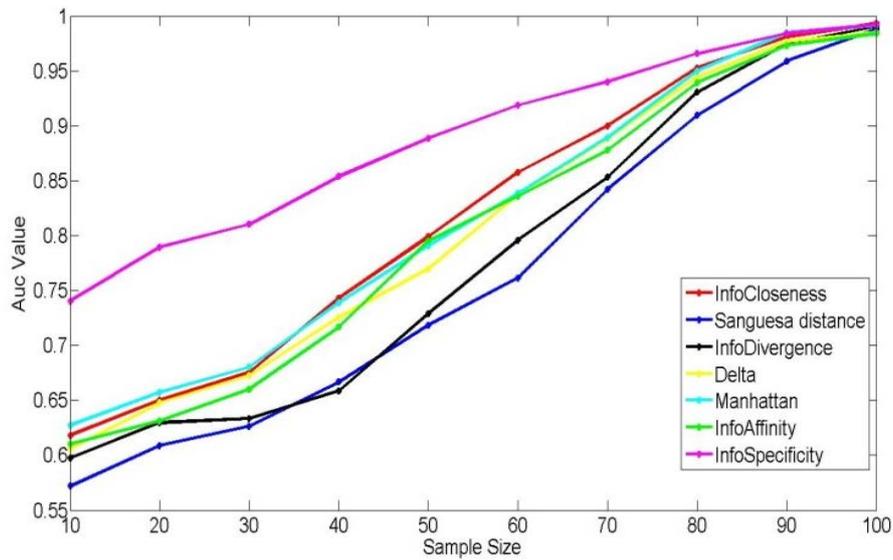


Figure 3.12 Comparaison des mesures de similarité, basée sur les valeurs de l'aire sous la courbe ROC (AUC) pour la base de données « Pima Indians Diabetes »

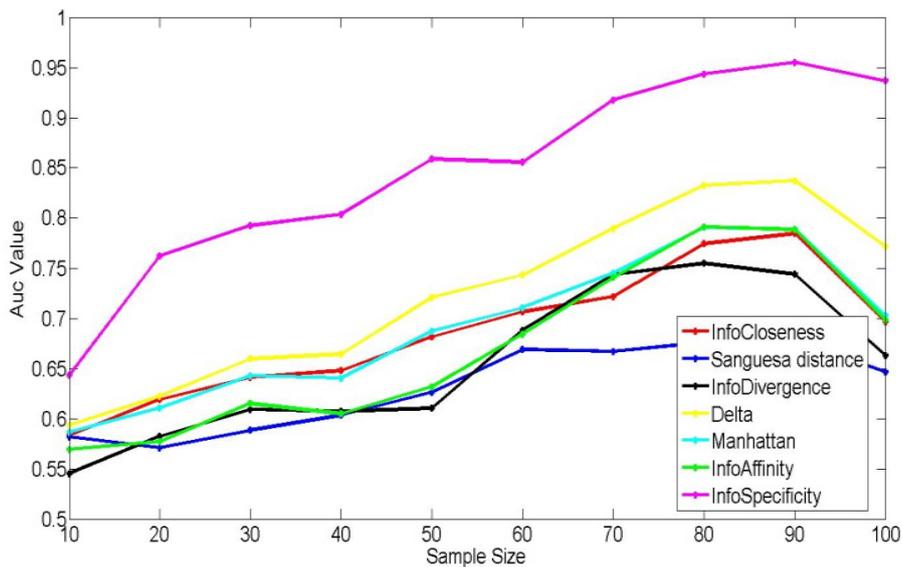


Figure 3.13 Comparaison des mesures de similarité, basée sur les valeurs de l'aire sous la courbe ROC (AUC) pour la base de données « Statlog »

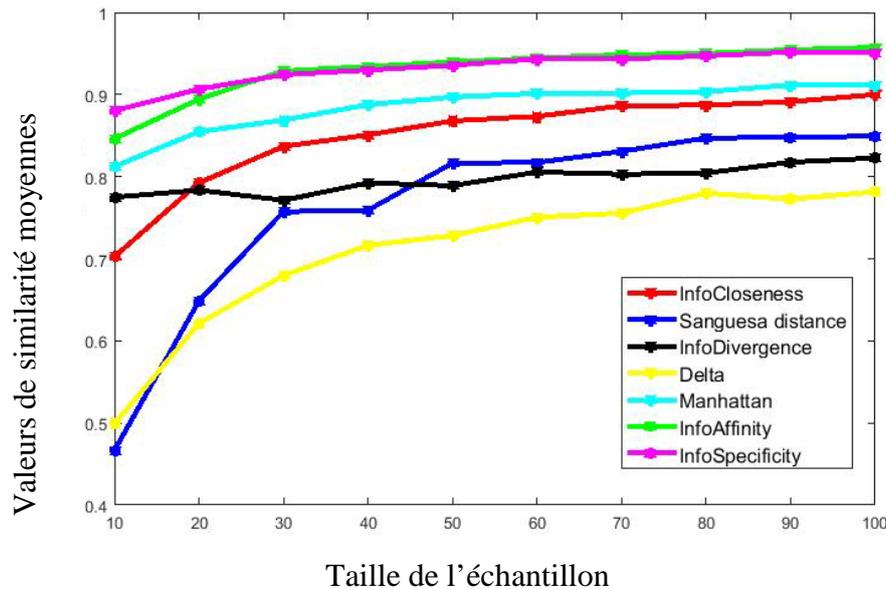


Figure 3.14 Valeurs de similarité moyennes entre distributions construites du même échantillon par les deux approches d'estimation Klir et Asymétrique de Dubois et Prade, en fonction de la taille des échantillons. Test effectué sur la base « Diabetes »

- 1- Les résultats varient d'une base à l'autre, ceci confirme l'influence de la nature des données évaluées. Pour la base « Statlog » (Cf. Figure 3.13), il est clair que la qualité des données ou encore la pertinence de l'attribut envers leur caractérisation, a engendré une difficulté pour la perception des similarités pour la majorité des mesures évaluées. Il est aussi à signaler que cette base fournit le plus petit nombre d'observations disponibles par classe (Cf. Tableau 3.2). Plus la taille de l'échantillon est importante moins nous pouvons effectuer de tests : c'est ce qui explique la chute des courbes pour la taille 100 de l'échantillon (nombre de tests non fiable). Dans ces mêmes conditions, l'approche proposée, a montré de meilleures performances. Spécialement, pour des cas d'incomplétude (cas des échantillons de petites tailles).
- 2- L'efficacité de distinguer entre les distributions semblables et les distributions différentes est aussi dépendante de la qualité de modélisation influée elle-même par la diminution de la taille de l'échantillon employé pour la construction de la distribution. Pratiquement, toutes les approches, excepté « InfoSpecificity », arrivent à déceler les similarités et les différences, à partir d'une taille d'échantillon supérieure à 60 ($AUC \geq 0.8$). Pour l'approche proposée « InfoSpecificity », AUC dépasse 0.8 à partir de la taille 30 des échantillons, (Cf. Figure 3.10, Figure 3.11, Figure 3.12, Figure 3.13). Là encore, un dépassement de performances remarquable envers le défi des petites tailles, est à signaler pour l'approche proposée.
- 3- Les performances de l'approche *InfoSpecificity* sont comparables à celles de la distance de Manhattan. Ce résultat est bien attendu, vu que la première fait intégrer la deuxième par sa définition. En revanche, nous notons l'avantage apporté par la quantification de la différence entre les quantités d'information, assurée par la mesure de spécificité et ce spécialement dans les cas où règne l'ambiguïté soit par la nature des données dans le cas de la base « Statlog » (Cf. Figure 3.13) ou dans les

cas des petites tailles des échantillons (Cf. Figure 3.10, Figure 3.11, Figure 3.12 et Figure 3.13).

- 4- L'évaluation des approches testées envers la perception des similarités entre des modèles construits par différentes approches à partir du même échantillon, a montré que les deux approches basées conjointement sur les deux types d'écart (point à point et quantité d'information), à savoir l'approche « *InfoAffinity* » et celle proposée « *InfoSpecificity* » assurent des performances semblables et meilleures que les autres approches pour les quatre bases testées. La Figure 3.14 illustre un exemple de cette évaluation.

De ces observations, on peut conclure sur la supériorité de l'approche proposée dans l'évaluation du degré de similarité entre deux informations encodées par des distributions de possibilités. On note également ses bonnes performances réalisées dans les conditions de manque de données ou encore de leur faible contenu informatif.

Cette nouvelle mesure a été employée dans plusieurs de nos travaux dont l'un est brièvement décrit dans le paragraphe suivant.

8. Emploi de la mesure InfoSpécificity dans la construction d'un arbre de décision possibiliste

a) Notions préliminaires concernant un arbre de décision

L'arbre de décision est l'une des techniques les plus connues et les plus utilisées dans le domaine de prise de décision et spécialement de fouille des données. Leur principal avantage réside dans la fourniture d'un formalisme puissant pour présenter des classificateurs compréhensibles et souvent faciles à interpréter par des experts et même par des utilisateurs ordinaires [BBE 2004].

➤ ***Les arbres de décision sont composés de trois éléments de base :***

- les nœuds qui correspondent aux caractéristiques (attributs) ;
- les branches ou arcs qui représentent chacune une valeur de la caractéristique testée dans le nœud duquel elles sont issues ; et
- les feuilles qui spécifient la classe résultante (les nœuds terminaux de l'arbre).

➤ ***Trois phases sont essentielles pour un arbre de décision :***

- *Phase de construction ou d'induction*

La construction d'un arbre de décision est une méthode d'apprentissage inductif supervisé : On doit disposer d'une base de cas pour l'apprentissage, ce qui signifie que chaque élément est décrit par un ensemble d'attributs et est affecté à une classe. A partir de cette base, l'apprentissage inductif supervisé consiste à trouver une règle de décision permettant d'associer à tout élément inconnu, décrit par le même ensemble d'attributs, une classe.

- *Phase de classification ou d'inférence*

Cette phase consiste à classer des nouvelles instances non étiquetées en se basant sur l'arbre de décision construit. La classe d'une nouvelle instance non étiquetée est fournie en débutant par la racine de l'arbre et en suivant le chemin correspondant aux valeurs des caractéristiques jusqu'à atteindre une feuille de l'arbre.

- *Elagage de l'arbre*

L'élagage intervient au cours de la construction de l'arbre afin d'enlever les nœuds inutiles dans le but de simplifier l'arbre et la rendre plus compréhensible.

➤ *Trois composants distinguent les algorithmes d'arbres de décision :*

- *La mesure de sélection d'attribut*

La procédure de construction des arbres de décision est basée sur la stratégie du choix d'un attribut à chaque nœud. Une mesure de sélection d'attribut sert à trouver l'attribut le plus informatif à chaque étage de l'arbre. C'est un paramètre critique qui est généralement basé sur l'entropie de l'information.

- *La stratégie de partitionnement*

Elle consiste à diviser l'ensemble d'apprentissage en métaclasse selon l'attribut sélectionné.

- *Les critères d'arrêt :*

Il détermine les conditions d'arrêt du processus du partitionnement.

b) Contribution et application

La notion d'arbres de décision possibilistes n'est pas une nouveauté. Elle est inspirée de l'arbre de décision flou [YS 1995] qui utilise les concepts des ensembles flous et des degrés d'appartenance pour représenter des instances ayant des valeurs imprécises de leurs attributs ou encore d'appartenance aux classes. Dans les modèles d'arbres de décision possibilistes, la même représentation que celle des arbres de décision flous est conservée, sauf que les feuilles sont étiquetées par des valeurs de possibilité au lieu des degrés d'appartenance. La différence entre les différents algorithmes d'arbres de décision possibilistes, tels que les algorithmes NS-PDT [JBE 2008], Sim-PDT [JBB 2008] et Clust-PDT [JBE 2009], réside essentiellement dans le choix de l'ensemble de leurs composants (la mesure de sélection d'attribut, la stratégie de partitionnement et le critère d'arrêt). Une étude de l'état de l'art accompagnée de simulation d'exemples pratiques a été effectuée et détaillée dans [Char 2014].

Notre contribution se situe au niveau de la mesure de sélection d'attributs agissant sur la phase de partitionnement. L'algorithme modifié est celui de Sim-PDT [JBB 2008], étant lui-même une adaptation de l'algorithme C4.5 avec le contexte incertain possibiliste. Nous avons employé notre mesure de similarité *InfoSpecificity* pour la recherche des similarités entre les distributions de possibilités modélisant, à différents niveaux de la construction de l'arbre, les classes et les métaclasses.

Cette modification a d'abord été testée et évaluée dans [Char 2014] pour des bases standards. Les critères d'évaluation qui ont été adoptés sont : la durée de construction (en secondes) à partir d'un ensemble d'apprentissage donné, la taille de l'arbre et le taux de classification.

Dans [Dam 2014, DKR 2014], nous avons proposé une approche de reconnaissance d'objets dans un environnement domestique. Notre contribution consiste à utiliser des fonctionnalités de bas niveau extraites d'images avec des concepts de haut niveau générés à partir d'une ontologie d'objets domestiques pour aboutir à une décision plus riche. L'approche consiste à développer une classification sémantique en fournissant à un utilisateur de canne blanche intelligente, aussi bien la classe de l'obstacle que la scène dans laquelle il se trouve. La classification est effectuée par le modèle de l'arbre de décision possibiliste proposé ci-haut.

Une comparaison avec le classifieur SVM a été effectuée sur la base de 16 classes comportant chacune une vingtaine d'images. 94 attributs ont été utilisés : 30 de forme, 36 de couleur et 28 attributs de texture. La Figure 3.15 illustre les résultats de reconnaissances des deux approches comparées pour chacune des 16 classes. L'arbre de décision possibiliste proposé reconnaît mieux que SVM 10 classes sur 16. L'arbre de décision a également réalisé un meilleur taux global de reconnaissance de 81.87% contre SVM 72.18%.

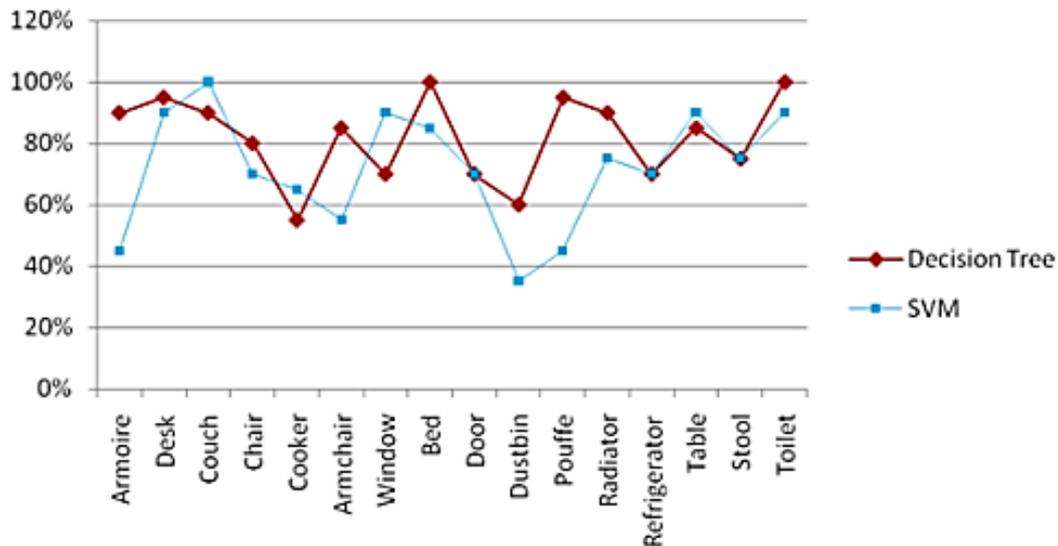


Figure 3.15 Taux de reconnaissance données par l'arbre de décision possibiliste et le classifieur SVM pour les différentes classes

IV. Travaux dans le cadre de la notion de fusion

1. Sélection d'opérateurs de fusion dans un cadre possibiliste

La sélection d'attributs constitue une étape très importante dans le processus de décision. Elle permet d'écartier les attributs corrélés (non discriminatifs) qui risquent de dégrader plutôt que d'enrichir les résultats à l'issue d'un processus de prise de décision. La fusion des attributs sélectionnés, permet d'obtenir une complémentarité descriptive et discriminante plus importante que celle obtenue à partir de chaque attribut à part. La manière de fusionner doit aussi être choisie d'une façon judicieuse afin d'éviter la dégradation dans le processus de prise de décision. Dans le cadre possibiliste, une grande variété d'opérateurs de fusions ont été adoptés et aussi adaptés. Classiquement ils sont classés en trois catégories : opérateurs conjonctifs, disjonctifs ou compromis. Ces modalités fondamentales de fusion n'ont pas le même comportement : la modalité conjonctive généralise l'intersection des événements, la modalité disjonctive généralise l'union des événements et la modalité de compromis utilise une règle de combinaison adaptative pour passer continuellement de la modalité conjonctive à la modalité disjonctive. Décider de l'utilisation d'un opérateur en dépit de l'autre n'est pas une question de préférence ou de hasard. La sélection d'un opérateur de fusion détermine la démarche de la combinaison des sources d'information et doit donc être bien raisonnée. Etablir un processus de sélection de l'opérateur de fusion, suggère une connaissance préalable de la fiabilité des sources et du degré de cohérence entre les informations fournies.

En théorie des possibilités, les informations sont représentées par des distributions de possibilités. Afin que la fusion entre deux distributions puisse avoir un sens, il est nécessaire

que ces modèles soient complémentaires (de façon à éviter les redondances) et cohérents (pour ne pas tomber dans le désaccord). Le degré d'inconsistance Inc du résultat de la fusion de deux distributions de possibilités, a souvent été utilisé pour juger leur cohérence. Cependant, seul, ce critère ne peut pas être suffisant, en particulier lorsque plusieurs paires de distributions de possibilités incertaines ont le même degré d'inconsistance. Pour remédier à cette lacune, il est important d'étudier de manière combinée la complémentarité entre les degrés d'incertitude et d'inconsistance. Dans [Liu 2007], Liu a proposé une idée pour combinaison de ces deux degrés afin de sélectionner l'opérateur de fusion approprié. Son idée n'a cependant pas été ni implémentée ni expérimentée. L'idée de cette combinaison est très intéressante mais il est très important de soigneusement quantifier ces degrés. Cette approche est uniquement basée sur un opérateur conjonctif ou disjonctif. Bien que, dans certains cas, l'utilisation d'un opérateur de compromis, en particulier lorsque les degrés de conflit et d'inconsistance ne sont pas cohérents. De plus, l'utilisation de cette méthode pour fusionner plusieurs sources d'information donne une possibilité de distribution avec des valeurs faibles qui tendent vers zéro. Notre apport dans cette question consiste en une modification apportée à l'approche de Weiru Liu.

a) Approche de Weiru Liu [Liu 2007]

L'approche consiste à évaluer le degré de cohérence entre les distributions de possibilités à fusionner, ainsi que leurs degrés de conflit. Elle considère deux concepts fondamentaux.

Le premier concept, qui consiste à mesurer le degré d'inconsistance d'une distribution, s'allie à la théorie des possibilités. Ce degré renseigne sur le niveau de conflit entre les informations de différentes sources. L'inconsistance d'une distribution π est définie par :

$$Inc(\pi) = 1 - \max_x \{\pi(x)\} \quad \text{Eq3.13}$$

Le deuxième concept est basé sur la théorie des fonctions de croyance. Il est employé pour mesurer la cohérence entre les distributions à fusionner. Ce concept génère des fonctions de masse à partir des distributions de possibilités. Ensuite, ces masses sont transformées en fonctions de probabilité pignistique. La probabilité pignistique, notée $BetP$, permet d'approcher le couple de mesures "croyance" et "plausibilité". Ce processus est réalisé par l'équi-répartition des mesures de masses données aux éléments non singletons sur les singletons qui les composent. La probabilité pignistique pour un élément ω est définie par :

$$BetP_m(\omega) = \sum_{A=2^\Omega, \omega \in A} \frac{m(A)}{|A|} \quad \text{Eq3.14}$$

Où ω est un singleton de l'univers Ω , A est un sous ensemble de l'univers Ω , $|A|$ est le cardinal de A et m est une fonction de masse.

La probabilité pignistique de A un sous ensemble de Ω est définie par :

$$BetP_m(A) = \sum_{\omega \in A} BetP(\omega) \quad \text{Eq3.15}$$

La distance entre les deux fonctions de masses m_1 et m_2 est évaluée par une distance de Tessem [Tes 1993] donnée par :

$$DT_{m_1}^{m_2} = \max_{A \subseteq \Omega} |BetP_{m_1}(A) - BetP_{m_2}(A)| \quad \text{Eq3.16}$$

Le Tableau 3.4 illustre la stratégie de Weiru Liu pour sélectionner l'opérateur de fusion adéquat selon les critères Inc et DT_{m1}^{m2} .

Tableau 3.4 Illustration de la stratégie de Weiru Liu [Liu 2007] pour la sélection des opérateurs de fusion

Valeur		Inc		
		Basse	Moyenne	Grande
DT_{m1}^{m2}	Basse	Produit	Produit	Operator disjunctif
	Moyenne	Produit	Produit	Operator disjunctif
	Grande	Min	Min	Operator disjunctif

b) Modification proposée

Dans [AK 2017], nous avons proposé de faire la sélection de l'opérateur de fusion selon les deux critères, à savoir le conflit et l'incohérence, d'une façon indépendante, puis de traiter les conflits de sélection. Ainsi notre approche se déroule sur trois étapes :

✓ **Etape1 : Choix de la modalité de fusion basée sur le degré de conflit**

Le conflit apparaît naturellement lorsque plusieurs sources d'informations imparfaites sont en jeu. Une interprétation du comportement conflictuel des sources peut se faire selon l'importance de l'inconsistance de la distribution résultante de leur fusion. On distingue 3 cas :

1. Cas où l'inconsistance $Inc(\hat{\pi})$ est faible : Les sources d'information sont consonantes. Nous pouvons de ce fait constater qu'elles sont fiables et que dans ce cas, l'opérateur de fusion doit être « sévère ». Un opérateur conjonctif peut être utilisé, tels que l'opérateur minimum ou l'opérateur produit.
2. Cas où l'inconsistance $Inc(\hat{\pi})$ est partielle : C'est le cas le plus commun, un opérateur de compromis est à choisir, tels que la moyenne harmonique, moyenne géométrique, moyenne arithmétique, moyenne quadratique, etc.
3. Cas où l'inconsistance $Inc(\hat{\pi})$ est importante : Les sources d'information sont dissonantes. On ne peut pas préciser laquelle qui est fiable. Dans ce cas, l'opérateur de fusion doit être « permissif ». Un opérateur disjonctif peut être utilisé, tels que l'opérateur maximum ou l'opérateur somme.

✓ **Etape2 : Choix de la modalité de fusion basée sur la cohérence**

La détermination du niveau de cohérence entre deux distributions de possibilités, peut être effectuée par le calcul de la distance de Tessem entre leurs fonctions de masses respectivement inférées. Plus les fonctions sont proches, plus les sources sont considérées cohérentes. Une interprétation de l'incohérence des sources peut alors se faire sur la base de la valeur prise par cette distance. On distingue 3 cas :

1. Cas où DT_{m1}^{m2} est faible : Les sources d'information sont cohérentes. L'utilisation d'un opérateur conjonctif permet un compromis entre les deux sources.
2. Cas où DT_{m1}^{m2} est partielle : C'est le cas le plus commun, un opérateur de compromis est à favoriser.
3. Cas où DT_{m1}^{m2} est importante : Les sources d'information sont incohérentes. Un opérateur de fusion disjonctif est à adopter afin de permettre toutes les possibilités obtenues par les différentes sources.

Le Tableau 3.5 résume les choix effectués séparément selon les deux critères adoptés.

Tableau 3.5 Choix de modalité de fusion effectués séparément selon les deux critères adoptés (opérateur conjonctif (\wedge), opérateur disjonctif (\vee), opérateur compromis (\bowtie), présentés de la forme : choix(Inc)/choix(DT))

Valeur		Inc		
		Basse	Moyenne	Grande
DT_{m1}^{m2}	Basse	\wedge/\wedge	\bowtie/\wedge	\vee/\wedge
	Moyenne	\wedge/\bowtie	\bowtie/\bowtie	\vee/\bowtie
	Grande	\wedge/\vee	\bowtie/\vee	\vee/\vee

✓ **Etape3 : Sélection de la modalité de fusion selon les deux critères**

Le choix final de la modalité de fusion, en se basant à la fois sur les deux critères, se fait selon deux cas :

1. Cas où les deux critères sont consentants sur la modalité : dans ce cas le choix est évident.
2. Cas où les deux critères divergent sur leurs choix : dans ce cas, favoriser l'un des critères sur l'autre n'est pas une bonne stratégie, alors une modalité de compromis est adoptée.

Finalement la procédure de sélection de la modalité de fusion est illustrée par le Tableau 3.6.

Tableau 3.6 Méthodologie de sélection de la modalité de fusion selon les deux critères de conflit et de cohérence (opérateur conjonctif (\wedge), opérateur disjonctif (\vee), opérateur compromis (\bowtie))

Valeur		Inc		
		Basse	Moyenne	Grande
DT_{m1}^{m2}	Basse	\wedge	\bowtie	\bowtie
	Moyenne	\bowtie	\bowtie	\bowtie
	Grande	\bowtie	\bowtie	\vee

c) Expérimentations et résultats

- 1- La modification apportée à l'approche de Weiru Liu a été comparée à l'approche originale ainsi qu'à trois autres opérateurs de fusion, communément utilisés, à savoir les opérateurs maximum, minimum et produit. Les tests d'évaluation ont été conduits sur deux benchmarks, à savoir « IRIS » [Bla, TXO 2005, DB 2004] et « Diabète » [UCI 2008]. La première base comporte trois classes décrites par quatre attributs, alors que la deuxième base comporte deux classes décrites par huit attributs. Chacune des classes est d'abord représentée par autant de distributions de possibilités qu'il y en a d'attributs considérés pour décrire les observations. Ainsi, pour la base « IRIS » chaque classe dispose de quatre modèles, et pour la base « Diabète », chaque classe dispose de huit modèles. Les approches de fusion à évaluer, sont utilisées afin de n'avoir qu'un seul modèle pour une classe. Les modèles résultants des fusions sont évalués selon trois critères :
 - 1- Un critère psychovisuel : Le modèle possibiliste présente selon les valeurs prises des attributs, des possibilités d'appartenance variant entre 0 et 1. La Figure 3.16 illustre trois zones d'incertitude sur la possibilité d'appartenance à une classe. Un modèle est jugé d'autant plus représentatif de la classe qu'il touche à la fois aux deux zones

extrêmes : celle de forte possibilité d'appartenance et celle de faible possibilité d'appartenance.

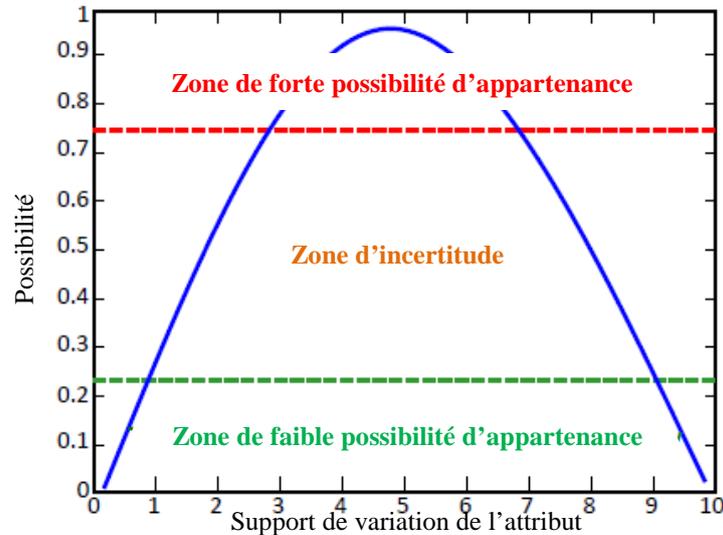


Figure 3.16 Définition de trois zones d'incertitude sur la possibilité d'appartenance à une classe

- 2- Deux critères quantitatifs : (i) une mesure de non spécificité évaluant l'incertitude du modèle après la fusion. En effet, le modèle est d'autant plus informatif que l'incertitude est faible. La mesure d'incertitude U a été utilisée. (ii) une mesure de similarité quantifiant le degré de similitude entre les classes après la fusion. En effet, l'approche de fusion qui permet d'écarter plus les modèles représentatifs des classes (minimiser la similarité), est jugée meilleure. La mesure de similarité « Infocloseness » a été utilisée.

La fusion a été assurée sans passer par une procédure de sélection d'attributs d'abord, ce qui ne rend pas le résultat aussi pertinent que ne le sont les modèles des classes avant fusion. En effet la présence d'attributs non discriminatifs dégrade remarquablement la fiabilité des modèles conçus par fusion, de façon à ne pas favoriser spécialement la séparation des classes. Cette finalité n'étant pas notre objectif actuel, nous ne considérons que les apports relatifs des approches pour les comparer. Les figures Figure 3.17 et Figure 3.18 présentent, respectivement, les distributions de possibilité des classes des bases « IRIS » et « Diabète », obtenues après fusion par les cinq approches considérées. Visuellement, nous pouvons facilement voir que l'approche proposée et l'opérateur « Maximum » sont les seules approches qui ont pu étaler les modèles de classe entre les deux zones de faible et forte possibilités d'appartenance. Les autres approches ne permettent pas une claire interprétation visuelle.

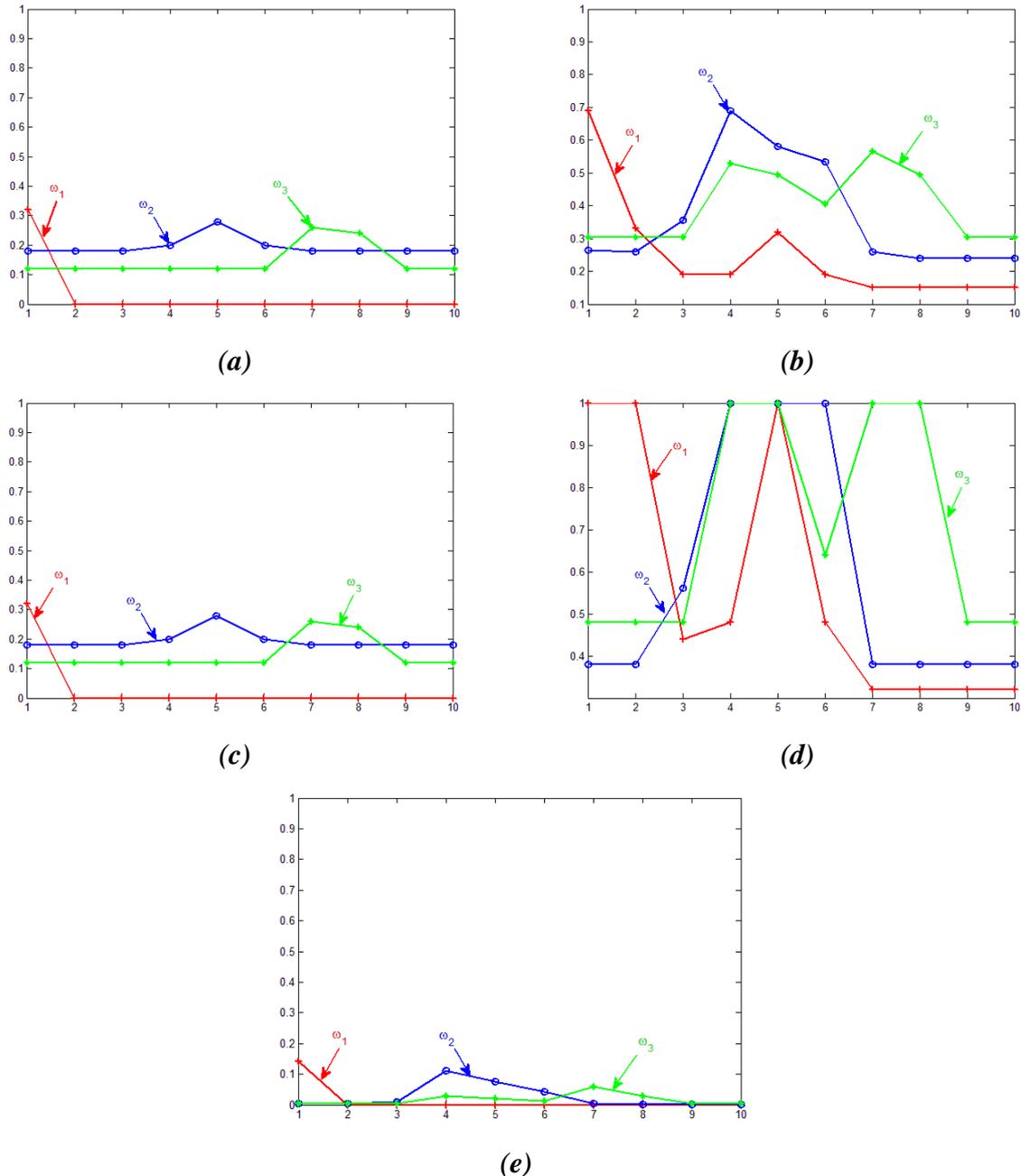


Figure 3.17 Distributions de possibilité des trois classes de la base de données « IRIS », obtenues après fusion par : (a) L'approche originale de Weiru Liu, (b) l'approche proposée, (c) l'opérateur Minimum, (d) l'opérateur Maximum et (e) l'opérateur Produit

La Figure 3.19 donnent les valeurs d'incertitude calculées sur les modèles des différentes classes des deux bases, résultants de fusion. Mis à part l'approche de Weiru Liu qui a remarquablement minimisé l'incertitude pour les deux classes de la base « Diabète » (Cf. Figure 3.19-b) mais au détriment de leur totale confusion (Cf. Figure 3.18-a), il est bien visible que globalement l'incertitude a gardé des valeurs élevées proches du maximum ($\log_2(K=10)$), et ceci est attribué à la non sélection d'attributs pertinents pour la caractérisation avant la fusion. Néanmoins, nous notons que l'approche proposée, bien qu'elle ne réalise pas les meilleures performances pour toutes les classes, globalement elle minimise l'incertitude mieux que les autres approches.

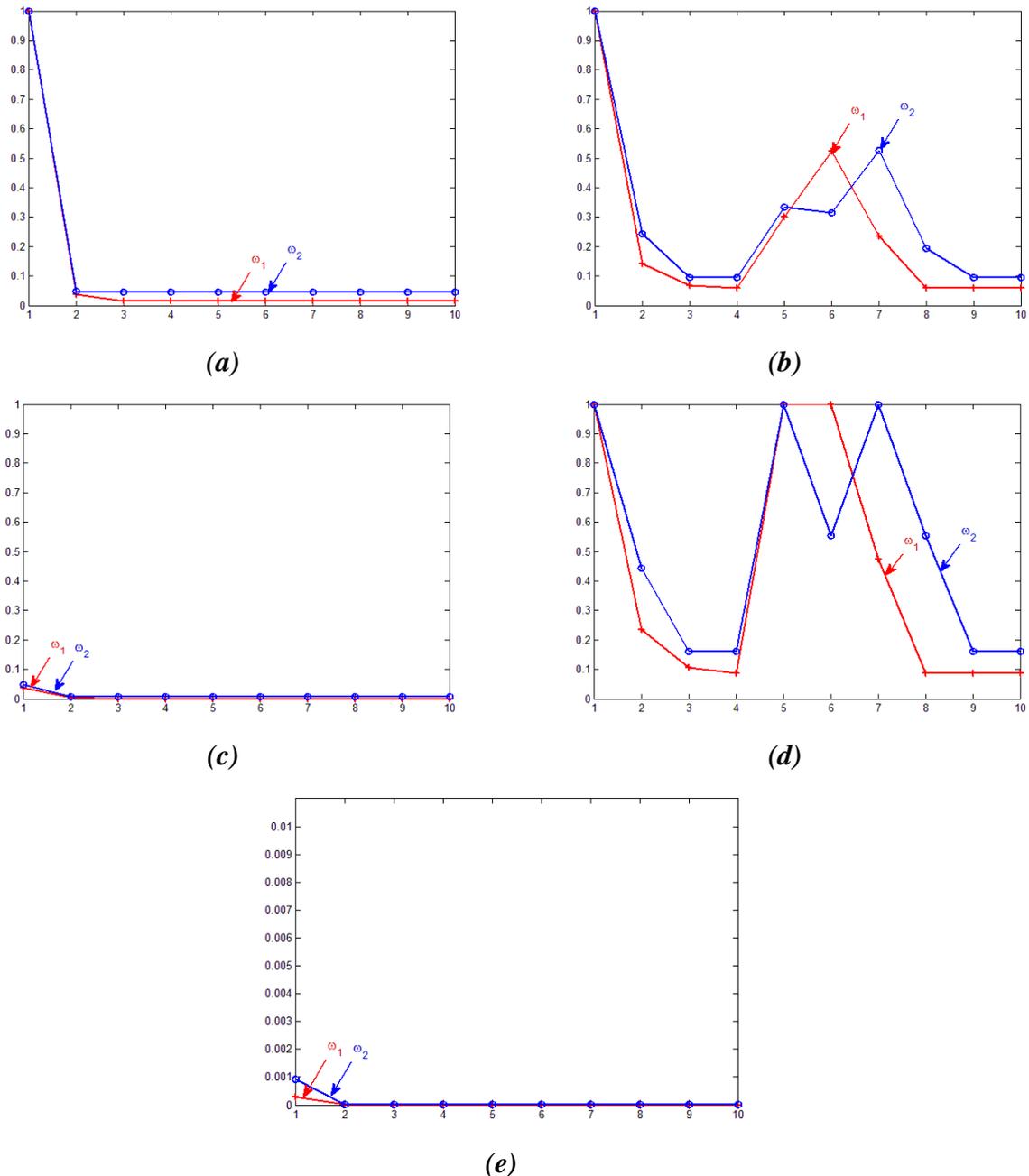
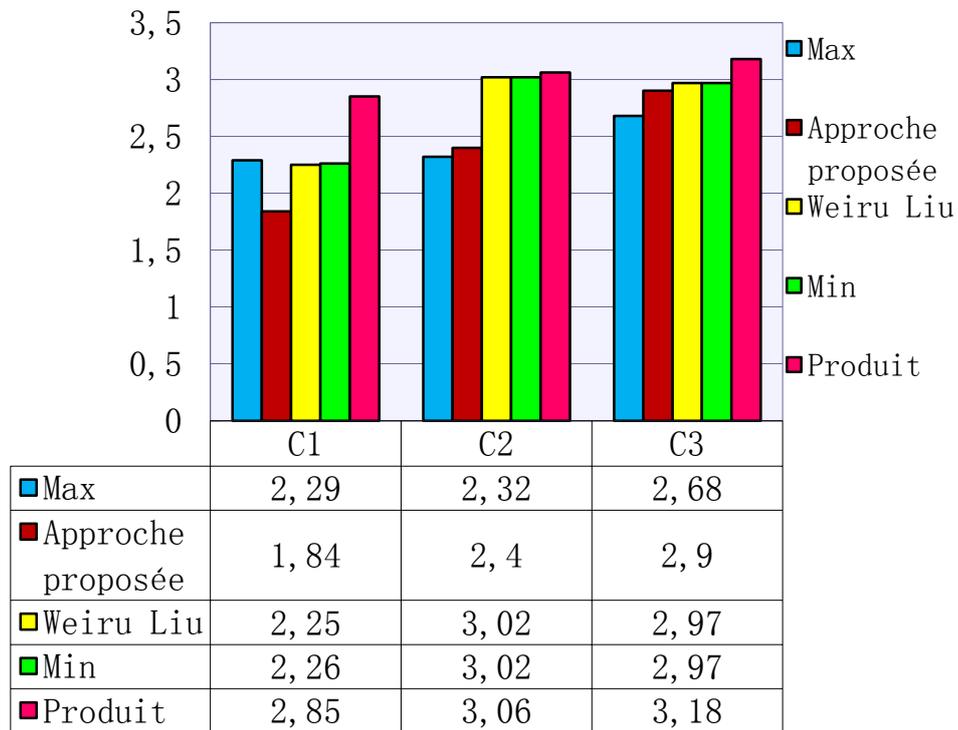


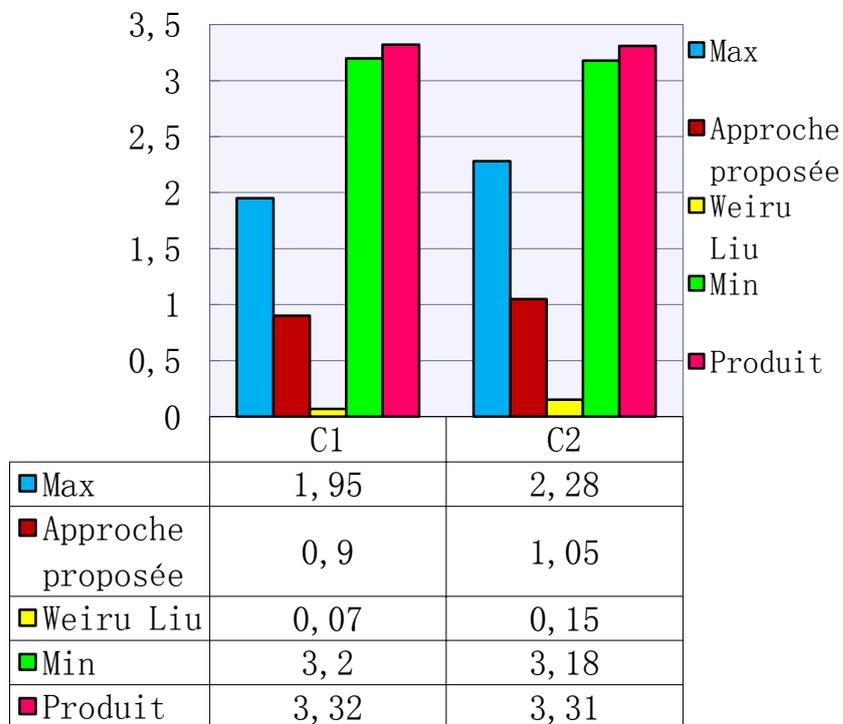
Figure 3.18 Distributions de possibilité des deux classes de la base de données « Diabète », obtenues après fusion par : (a) L'approche originale de Weiru Liu, (b) l'approche proposée, (c) l'opérateur Minimum, (d) l'opérateur Maximum et (e) l'opérateur Produit

La non-sélection des attributs discriminatifs avant la fusion, a fait que les modèles résultants des fusions ne permettent pas une bonne séparation des classes. La Figure 3.20 montre bien qu'un degré élevé de similarité est globalement trouvé entre les classes à séparer. Cependant, nous notons que l'approche proposée ainsi que l'opérateur disjonctif (Maximum) ont mieux favorisé cette séparation. Ceci est dû au fait que ces deux approches ont eu recours à un mode de fusion qui favorise la prise en compte du maximum de l'information disponible. Néanmoins, le recours à ce mode étant exclusif pour l'opérateur « Maximum », ne l'a pas aidé

à minimiser l'incertitude dans la majorité des cas de fusion exécutés (Cf. Figure 3.19), ce qui n'est pas le cas de l'approche proposée.



(a)



(b)

Figure 3.19 Incertitudes calculées sur les modèles possibilistes des différentes classes de : (a) la base « IRIS » et (b) la base « Diabètes »

Nous pouvons conclure que le choix d'un opérateur de fusion selon des règles se rapportant à la nature de l'information à fusionner est une stratégie judicieuse. L'approche proposée est prometteuse et invite à une plus profonde investigation dans l'étude de la relation entre la nature de l'information et celle de l'opérateur de fusion à adopter.

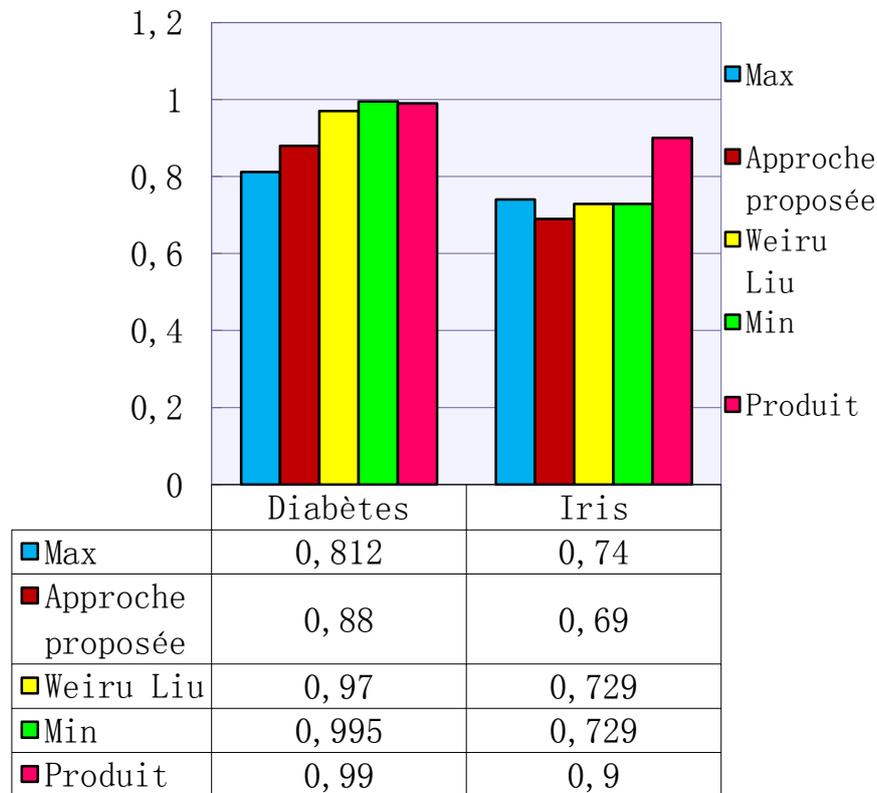


Figure 3.20 Similarités entre les différentes classes, estimées sur les modèles résultants des fusions par les différentes approches

2. Analyse multicritère de la décision : MCDA

Il s'agit d'une activité en cours d'étude et de développement qui consiste en la création d'un système de prise de décision multisources opérant en fonction de la la qualité de l'information de chacun des éléments intervenant (la qualité de l'objet mesuré, la qualité du capteur, la qualité des données radiométriques acquises, la qualité des attributs mesurés sur les données radiométriques). D'où l'appellation du concept : analyse multicritère de la décision (MultiCriteria Decision Analysis : MCDA).

Ce nouveau concept est en cours de développement dans le cadre d'une thèse pour une application de sécurité biométrique, visant à réaliser un système dynamique incrémentale pour une fusion multimodale biométrique basée sur la qualité des différents intervenants dans la procédure de fusion. Nous présentons dans ce qui suit le principe du concept MCDA, et nous soulevons surtout des questions importantes pour le développement de cette idée.

a) Objectif

Nous visons la réalisation d'un système d'identification ou d'authentification présentant les options suivantes :

1. Le système doit s'adapter aux disponibilités des sources d'information
2. Le système affecte un poids d'importance à chacune des sources employées, selon la qualité globale d'information qu'elles fournissent vis-à-vis de la discrimination des classes.

3. Selon la classe à identifier ou authentifier, le système affecte également à chacune des sources employées, un poids d'importance, lié à leurs efficacités vis-à-vis de la discrimination de cette classe, particulièrement, des autres.
4. Si l'univers des classes s'incrémente d'une classe supplémentaire, le système se doit d'ajuster les poids des sources d'une façon incrémentale, sans reprendre dès le début.

b) Contexte

On dispose d'une base d'apprentissage B, comportant N classes.

Chaque classe est observée par K sources pouvant être caractérisées par un ou plusieurs attributs (primitives).

c) Approche

- ✓ Disposant, pour chaque classe C_n , d'un vecteur d'observations X_n , formé par K vecteurs de primitives, issus des K sources observant la classe C_n , la première étape consiste à créer un modèle générique M_n pour chaque classe C_n .

→ M_n caractérise l'information dans C_n selon chaque source (ou dans l'espace de chaque primitive). C'est l'ensemble de K sous modèles de C_n .

$$M_n = \begin{bmatrix} Mod(C_n, S_1) \\ \vdots \\ Mod(C_n, S_K) \end{bmatrix} \quad \text{Eq3.17}$$

- ✓ Pour chacune des K sources, mesurer dans une matrice, des similarités mutuelles entre toutes les classes de l'univers considéré (représentées par leurs sous modèles relatifs à cette source).

→ On dispose à la fin de cette étape, pour chaque source, d'un genre de matrice de corrélation qui caractérise la source. Cette caractérisation peut se rapporter à la source : globalement ou relativement à chacune des classes.

$$A^{S_i} = \begin{matrix} & Mod(C_1, S_i) & \dots & Mod(C_N, S_i) \\ \begin{matrix} Mod(C_1, S_i) \\ \vdots \\ Mod(C_N, S_i) \end{matrix} & \begin{bmatrix} \dots & \dots & \dots \\ \vdots & \ddots & \vdots \\ \dots & \dots & \dots \end{bmatrix} & & \end{matrix} \quad \text{Eq3.18}$$

L'élément de cette matrice étant : $a_{n,m}^{S_i} = Sim(Mod(C_1, S_i), Mod(C_m, S_i))$, vérifiant :

- Si $a_{n,m}^{S_i} = 0 \rightarrow S_i$ confond parfaitement les classes C_n et C_m
- Si $a_{n,m}^{S_i} = 1 \rightarrow S_i$ sépare parfaitement les classes C_n et C_m
- Une caractérisation de la source relative à une classe C_n , peut être :
 - $\tau_{(O_n, S_i)} = \sum_m a_{n,m}^{S_i}$ avec $\tau_{(O_n, S_i)} \in [1, n]$
 - Si $\tau_{(O_n, S_i)} = 1 \rightarrow S_i$ est une source qui met C_n parfaitement à l'écart des autres
 - Si $\tau_{(O_n, S_i)} = n \rightarrow S_i$ est une source qui confond complètement C_n et toutes les autres classes

- Une caractérisation globale de la source peut être formulée par l'ensemble des $\tau_{(O_n, S_i)}$, $n=1 \dots N$
- ✓ A ce niveau plusieurs questions se relèvent :
 - Quelles informations peut-on extraire globalement des K matrices « Similarités/Source » : A^{S_i} pour $i=1..K$?
 - Quelles informations peut-on extraire individuellement de chacune des K matrices « Similarités/Source » : A^{S_i} pour $i=1..K$?
 - Quelle mesure de similarité utiliser pour calculer les $a_{n,m}^{S_i}$?
 - Comment évaluer la qualité de la similarité ? faut-il le faire par calcul des taux d'erreurs de classification et se baser sur des matrices de confusion ? ou se baser sur les taux de reconnaissance ?
 - Comment introduire une nouvelle classe C_j , sans refaire tous les autres calculs ?
- ✓ L'étape suivante consiste à extraire un poids d'importance pour chaque source.

L'un des scénarios peut être le calcul des poids $\alpha_{(O_n, S_i)} = f\left(\frac{1}{\tau_{(O_n, S_i)}}\right)$

 - $\alpha_{(O_n, S_i)}=1 \rightarrow S_i$ est complètement prise pour la classification de l'objet O_n
 - $\alpha_{(O_n, S_i)}=0 \rightarrow S_i$ est complètement écartée pour la classification de l'objet O_n
- ✓ Quantifier la qualité de l'information et des données de chaque vecteur d'observation $x_i^n \rightarrow \beta_{(O_n, S_i)}$
 - A ce niveau une recherche de corrélations possibles entre des $\tau_{(O_n, S_i)}$ et des paramètres qualitatifs de x_i^n , peuvent servir pour renseigner sur la source sans aller vers les similarités mutuelles
 - Une appréciation globale de la source S_i peut aussi être déduite
- ✓ Construire un nouveau modèle pour chaque classe C_n , tenant compte des $\alpha_{(O_n, S_i)}$ et des $\beta_{(O_n, S_i)}$

V. Fouille possibiliste

Généralement, l'objectif de la fouille de données est d'aider à la prise de décision en fournissant des modèles compréhensibles et interprétables. Elle est basée sur l'extraction de connaissances à partir des données disponibles. Notre objectif pour cette activité de recherche est l'application des concepts de la théorie des possibilités à la fouille de données de façon à extraire de nouvelles connaissances à partir d'informations issues de plusieurs sources de données et de gérer l'approximation dans les prises de décision. Le mécanisme de prise de décision proposé pour cette fin, est basé essentiellement sur la construction de résumés de connaissances pertinentes extraites des différentes sources d'information. Ces résumés sont présentés sous forme de « cartes possibilistes », à raison d'une carte pour chaque classe de l'univers des discours. La construction des cartes se fait selon une procédure itérative, son critère d'arrêt est lié à la stabilité de la carte. L'ensemble des cartes offrent une vision à la fois claire et concise de l'ensemble des informations disponibles, facilitant l'interprétation et la prise de décision.

1. Construction d'une carte possibiliste

Considérons le même cadre d'étude présenté à la section V-2 du deuxième chapitre du présent rapport. Une distribution de possibilités $\pi_{n,k}^m$ est ainsi construite pour chaque classe C^m , $m = 1..M$, de l'univers des discours Ω et selon $N(k)$ attributs $f_{n,k}$, $n = 1..N(k)$, sélectionnés pour décrire les observations de la $k^{\text{ème}}$ source d'informations S_k , $k=1..K$, tel qu'illustré par la Figure 2.32.

✓ Initiation de la carte possibiliste

Un seul modèle possibiliste π_k^m selon chaque source S_k et conditionnel à la classe C^m , est construit par fusion des $\pi_{n,k}^m$, $n = 1..N(k)$. L'approche de fusion proposée à la section IV-1 du présent chapitre, peut être employée. Une version initiale de la carte possibiliste modélisant C^m , est un regroupement des connaissances pouvant être extraites par les différentes combinaisons possibles, des modèles possibilistes π_k^m relatifs à toutes les sources. Ainsi une carte comporte $(2^K - 1)$ distributions de possibilités, construites par fusions des distributions initiales selon les K sources d'information, selon l'équation :

$$\hat{\pi}_{Cb_i}^m)_{t=0} = \bigwedge_{Cb_j \subset Cb_i} (\hat{\pi}_{Cb_j}^m), (i, j) \in \{1..(2^K - 1)\} \quad \text{Eq3.19}$$

Où Cb_i est l'une des $(2^K - 1)$ combinaisons possibles des sources d'informations et \wedge est un opérateur de fusion. Notons que les combinaisons sont de tailles variables, allant de 1 jusqu'à K .

La Figure 3.21 illustre l'exemple de la carte possibiliste initiale conditionnelle à une classe C^m observée par 4 sources d'information ($K=4$). Les modèles résultant de fusion, sont disposés en étages selon la taille des combinaisons.

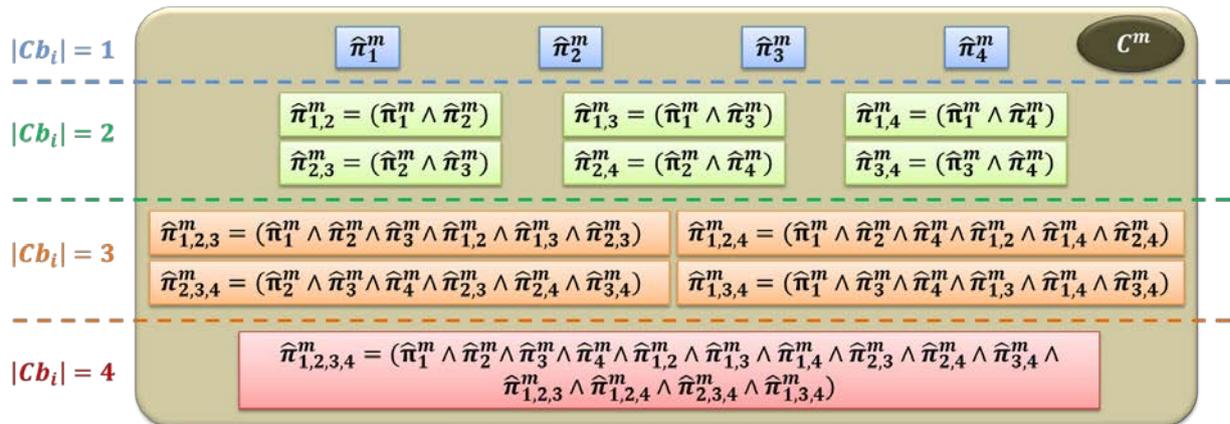


Figure 3.21 Initiation de la carte possibiliste conditionnelle à la classe C^m observée par 4 sources d'information

✓ Calcul des poids pour le raffinement des fusions

Les $(2^K - 1)$ distributions de possibilités de la carte possibiliste sont des modèles représentant la classe C^m , cependant ils n'ont pas tous les mêmes performances pour la distinguer et la séparer des autres classes. Cette considération est prise en compte pour raffiner la carte possibiliste de C^m . L'approche de raffinement consiste à affecter les

différentes combinaisons de sources, de poids correspondant à leurs importances relatives dans la séparation de C^m des autres classes. Le poids à affecter à chaque combinaison est proportionnel à la distance estimée entre les modèles respectifs de C^m et de la réunion de toutes les classes de l'univers Ω , autres que C^m , correspondants à cette même combinaison. Cette distance est calculée comme suit :

$$D_{Cb_i}^m = D \left(\hat{\pi}_{Cb_i}^m, \hat{\pi}_{Cb_i}^{\{1..M\} \setminus \{m\}} \right) \quad \text{Eq3.20}$$

Le poids $\alpha_{Cb_i}^m$ est proportionnel à $D_{Cb_i}^m$ de façon à ne pas écraser la distribution de possibilités résultante de la fusion entre les modèles correspondants aux combinaisons faisant intervenir Cb_i .

✓ **Processus itératif de raffinement**

Dans chacune des itérations de ce processus, les modèles de la carte possibiliste de C^m , correspondant aux différentes combinaisons possibles des sources d'information disponibles, sont mis à jour selon l'équation suivante :

$$\hat{\pi}_{Cb_i}^m)_t = \bigwedge_{Cb_j \subset Cb_i} \left[\alpha_{Cb_j}^m)__{t-1} \times \hat{\pi}_{Cb_j}^m)__{t-1} \right], (i, j) \in \{1..(2^K - 1)\} \quad \text{Eq3.21}$$

t étant le numéro de l'itération.

Après la mise à jour de la carte possibiliste, les poids $\alpha_{Cb_i}^m)_t$ sont recalculés en fonction des distances $D_{Cb_i}^m)_t$, conformément à l'équation 3.20.

✓ **Critère d'arrêt**

Les itérations s'arrêtent quand la carte possibiliste de C^m devient stable. En effet, à la fin de chaque itération un écart entre les poids est évalué. Le poids, relatif à une combinaison, présentant un écart par rapport à l'itération précédente, inférieur à un seuil T fixé expérimentalement, est jugé stable. Les itérations s'arrêtent quand tous les poids sont jugés stables.

2. Prise de décision basée sur les cartes possibilistes

Les modèles possibilistes constituant la carte possibiliste de la classe C^m , sont des représentations raffinées de la classe C^m mais d'efficacités variables. Nous employons les poids $\alpha_{Cb_i}^m)__{t_{fin}}$, générés à la dernière itération, comme des facteurs de confiance à affecter aux degrés de possibilités, qu'ils fournissent respectivement. Ainsi, étant x , un état de connaissance sur le monde réel, le degré de possibilité $\pi(C^m)$ que C^m soit l'unique occurrence, est une moyenne des degrés de possibilités respectifs des modèles de la carte possibiliste de C^m , affectés de leurs coefficients de confiance respectifs. $\pi(C^m)$ est donné par :

$$\pi(C^m) = \sum_{i=1}^{2^K-1} \alpha_{Cb_i}^m)__{t_{fin}} \times \hat{\pi}_{Cb_i}^m)__{t_{fin}}(C^m) \quad \text{Eq3.22}$$

$$\text{Avec } \hat{\pi}_{Cb_i}^m)__{t_{fin}}(C^m) = \hat{\pi}_{Cb_i}^m)__{t_{fin}}(x) \quad \text{Eq3.23}$$

Où x exprime l'état de connaissance sur le monde réel.

Le critère de décision employé ensuite est celui de la possibilité maximale. Ainsi la classe gagnante C_G est tel que :

$$\pi(C_G) = \max_{m=1..M} \pi(C^m) \quad \text{Eq3.24}$$

3. Application

Nous avons utilisé l'approche de fouille possibiliste basée sur l'utilisation des cartes possibilistes dans [Amm 2014] pour classifier l'état du sol en trois classes, à savoir : surface plane, escalier ascendant et escalier descendant, en se basant sur le traitement et l'analyse des signaux ultrasonores. Un seul capteur ultrason a été employé, nous avons alors exploité des transformations du signal temporel fourni par ce capteur comme sources d'information. Trois sources ont été ainsi considérées : le signal temporel, son spectrogramme et son périodogramme. L'approche de fouille possibiliste proposée a permis un taux de classification de 90% contre un taux de 75% donné par une approche basée sur des méthodes conventionnelles d'analyse du signal que nous avons proposée dans [AKS 2014].

VI. Conclusion

La prise de décision fait intervenir plusieurs mécanismes et notions, qui se doivent d'interagir et de s'adapter à la nature des informations disponibles d'une part et de l'objectif de la décision d'autre part. Dans le contexte possibiliste adopté pour notre recherche, les notions de la similarité et la fusion sont fondamentales. Ce chapitre a été consacré à la description de certaines de nos activités reliées à ces deux notions. Nous avons également présenté une approche s'apparentant à la fouille de données tout en restant dans le contexte possibiliste.

Nous commençons par présenter le concept de la similarité, ainsi que différentes notions marquant son caractère relatif. Les catégories des mesures de similarité apparues dans la littérature, identifient les mesures de similarité possibilistes en tant qu'une généralisation des mesures de similarité floues par la substitution de la fonction d'appartenance floue par une distribution des possibilités. Conscients du fossé sémantique entre ces deux dernières notions, nous avons recherché dans la littérature des mesures de similarité possibilistes qui tiennent compte des particularités du concept d'une distribution de possibilités. Une répartition des mesures de similarité possibilistes en deux catégories se basant sur la notion de forme est proposée. Cette référence à la forme vient du fait que l'apparence morphologique de la distribution renseigne sur la quantité d'information qu'elle renferme. Une nouvelle mesure de similarité « *Infospecificity* » basée sur le concept de non spécificité, combiné à une métrique basée sur l'écart d'amplitudes, est proposée dans la catégorie des mesures basées sur la forme. Nous avons ensuite proposé une méthodologie d'évaluation des mesures de similarité, à travers deux scénarios : le premier met à l'épreuve la mesure face à la détection des similitudes entre des membres de la même classe et des différences entre des membres de classes différentes ; tandis que le deuxième scénario met la mesure à l'épreuve quant à la détection des similitudes entre des distributions de possibilités créées pour un même échantillon mais par différentes techniques. Cette évaluation a montré que les approches combinant la prise en compte de la quantité d'information avec l'évaluation de l'écart point à point, sont mieux adaptées au contexte possibiliste. Spécialement la mesure « *Infospecificity* » que nous avons proposée a donné des résultats prometteurs, ce qui nous a incités à l'utiliser dans plusieurs de nos travaux dont l'un a été brièvement décrit dans ce chapitre. Il s'agit de l'utilisation d'un arbre de décision possibiliste employant cette mesure pour la reconnaissance d'objets dans un environnement domestique.

Deux travaux dans le cadre de la notion de fusion ont été ensuite présentés : le premier se rapporte à une approche de sélection d'opérateurs de fusion. En effet, nous parlons de fusion quand nous disposons de multiples sources d'information. Ces sources n'ayant pas toutes la même fiabilité, leurs exploitations ne doit pas être de la même façon. De plus la fusion se doit d'améliorer l'interprétation des données acquises en vue d'une représentation plus riche de la scène observée, il est alors crucial de choisir l'opérateur de fusion le plus approprié aux informations fournies. La contribution que nous avons proposée dans ce sens est une modification apportée à l'approche de Weiru Liu qui consiste à sélectionner l'opérateur de fusion selon des règles se rapportant à l'évaluation du conflit entre les sources d'une part et de leur cohérence d'autre part. L'évaluation de cette méthode nous a montré qu'elle souffre de diverses lacunes. Pour répondre à ces lacunes, nous avons réétudié les choix proposés par cette méthode. Le raffinement proposé est validé expérimentalement sur deux benchmarks, à savoir « IRIS » et « Diabète ».

Le deuxième travail décrit dans ce chapitre, dans le cadre de l'étude de la fusion dans un contexte possibiliste, consiste en une analyse multicritère de la décision. Le système proposé, fait intervenir dans la procédure de prise de décision, des facteurs de qualité concernant l'objet observé, le capteur, les données radiométriques acquises et les attributs mesurés sur les données radiométriques. Cette approche est en cours de développement.

La dernière section du chapitre présente une approche de fouille de données par l'extraction de connaissances possibilistes sous forme de distributions de possibilités correspondant à des combinaisons raffinées des sources d'information. Le processus de raffinement est itératif et l'ensemble des modèles résultant constitue une carte possibiliste comportant un résumé concis et pertinent pour la représentation d'une classe. Étant donné un état de connaissance sur le monde réel, la décision est formulée en se basant sur l'ensemble des cartes possibilistes relatives aux classes et des facteurs de confiance générés par le processus itératif de la construction des cartes.

La majorité des approches proposées et décrites dans ce chapitre ainsi que dans le chapitre précédent, ont été validées sur des acquisitions réelles, souvent capturées par des dispositifs que nous avons réalisés. Nous décrivons dans le chapitre suivant, nos activités de construction de scènes, ainsi que quelques approches conçues pour leurs interprétations.

Chapitre IV: **Interprétation de scène**

CHAPITRE IV: INTERPRETATION DE SCENE	142
I. INTRODUCTION	144
II. SYNTHÈSE DES ACTIVITÉS RELATIVES À CE CHAPITRE	146
a) Articles de Revue	147
b) Conférences internationales à comité de lecture	147
III. CONSTRUCTION DE SCÈNES	148
1. <i>Scène Biométrique</i>	148
a) Conception du protocole	148
b) Conception des deux irisopes	148
2. <i>Scène d'environnement intérieur</i>	150
a) Description.....	150
b) Schéma fonctionnel	152
IV. DIFFUSION DE CONNAISSANCES POSSIBILISTES ITERATIVES, APPLICATION À L'INTERPRETATION DE SCENE MEDICALE	153
1. <i>Problématique</i>	153
2. <i>Principe</i>	155
a) Système de prétraitement :	156
b) Système de décision :	158
c) Système de diffusion possibiliste :	158
d) Système d'évaluation et de mise-à-jour :	158
3. <i>Diffusion classe/classe</i>	159
4. <i>Diffusion croisée</i>	161
5. <i>Contribution : Diffusion croisée avec une représentation possibiliste des connaissances dans une clique</i>	165
a) Description du processus.....	165
b) Simulation sur un échantillon	167
6. <i>Résultats et discussions</i>	170
a) Estimation initiale des cartes possibilistes.....	170
b) Diffusion itérative Classe/Classe des connaissances possibilistes	171
c) Diffusion itérative Croisée des connaissances possibilistes	175
d) Effet de l'initialisation des connaissances possibilistes	177
e) Utilisation de la diffusion croisée pour la segmentation d'images mammographiques :	180
f) Effet de la modélisation possibiliste du contenu informationnel de la clique sur la diffusion croisée ...	181
V. FOUILLE POSSIBILISTE D'IDENTITÉ, INTERPRETATION DE SCENE BIOMETRIQUE : NOTION DU CARNET D'IDENTITÉ	185
VI. CONCLUSION.....	186

I. Introduction

La notion de scène est floue dans l'état de l'art. L'étude de plusieurs travaux concernant la définition et la reconnaissance de scène a permis dans [Bla 2016] la mise en évidence des notions combinées de : composition, contexte, éléments, lieu et action. Le terme scène selon [Als 2014], désigne un sous-ensemble d'un environnement au sein duquel sont considérés de multiples objets ayant un sens sémantique et des relations spatiales reliant ces objets. D'une façon usuelle, une scène est capturée par un ou plusieurs systèmes d'imagerie, projetant le monde réel en 3D sur un ou plusieurs supports 2D constituant ainsi une ou plusieurs images. Pour cette raison l'interprétation de scène est généralement rendue à une interprétation d'une ou de plusieurs images représentant cette scène. Le terme "interprétation de scène" désigne les mécanismes de raisonnement et d'analyse visant à extraire la sémantique contenue dans ces images [Hud 2005]. L'interprétation de scène est une opération cognitive effectuée habituellement par l'humain pour donner un sens à ses observations. Réalisée par l'humain, cette opération est intuitive se basant sur ses connaissances a priori d'une part et ses capacités cognitives d'observation et d'analyse d'autre part. Automatiser l'interprétation des données acquises d'une scène, fait appel à des outils et approches permettant l'identification du contenu sémantique de ces données collectées d'une ou de plusieurs sources de connaissances. Cinq niveaux d'interprétation ont été identifiés dans [Als 2014] en se rapportant à l'image comme support de mesures physiques approximant la scène: le niveau sous-pixelique, le niveau pixelique, le niveau région, le niveau objet et le niveau scène globale. Le résultat d'une interprétation dépend de plusieurs facteurs qui peuvent être résumés en l'objectif recherché, les connaissances a priori liées à la scène ou au capteur ou à l'image, la modélisation des connaissances et les mécanismes adoptés pour réaliser l'interprétation.

Ainsi définie, la tâche d'interprétation d'une scène n'est pas évidente mais le recours au développement d'outils et de stratégies appropriés est toujours de pointe. C'est dans ce cadre que se situent les travaux présentés dans ce chapitre. Deux principaux volets sont évoqués : Le premier concerne nos contributions au développement d'outils d'interprétation de scène affecté d'ambiguïtés . Quant au deuxième , il s'intéresse à la synthèse des scènes à interpréter.

Deux outils d'interprétation de scène sont présentés dans ce chapitre. Le premier est basé sur l'extraction de connaissance par la théorie possibiliste des données radiométriques, puis sa diffusion sur toute la scène afin de réaliser une interprétation. Trois approches sont proposées puis appliquées sur une scène médicale. Le deuxième outil concerne l'interprétation d'une scène biométrique. L'idée présentée dans ce chapitre est encore en cours de développement. Elle consiste à créer une sorte de carnet d'identité pour une personne à partir de ses éléments biométriques.

Par ailleurs, j'ai orienté l'application de mes travaux de recherche, que ce soient ceux portant sur l'interprétation ou ceux portant sur la modélisation et la prise de décision, vers diverses types de scènes, je précise : scène médicale, scène biométrique, scène d'environnement intérieur et, en cours, scène satellitaire. Ce choix de diversité est bien raisonné dans le sens à prouver que les approches proposées ne sont pas exclusivement conçues pour une application spécifique. L'idée n'est guère de prédire que les approches proposées sont universelles, mais plutôt il s'agit de lier leurs utilisations au contexte dans

lequel l'information, à traiter par l'application, est disponible. Par ailleurs, le choix spécifique de ces scènes a été, pour certaines, induit par les collaborations académiques, (je cite le domaine médical et le domaine satellitaire). C'est l'accès aux données offert par ces collaborations qui a orienté le choix. Ce n'était pas le cas pour les deux autres types de scènes, à savoir biométrie et environnement d'intérieur. Le choix dans ces deux cas a été orienté par la tendance internationale, tant politique, économique et sociale que de recherche scientifique, vers l'axe de la sécurité contre la menace terroriste (pour l'intérêt porté à la biométrie) et l'axe de la gérontechnologie et l'assistance aux personnes en déficience physiques (pour l'intérêt porté à l'exploration d'un environnement d'intérieur). Pour ces deux dernières scènes, les données ne nous ont pas été fournies.

Pour la biométrie pourtant les bases de données disponibles en ligne sont abondantes, mais elles ne répondent pas à notre besoin pour l'évaluation et le développement de nos approches. Notons à titre d'indication, qu'une base biométrique multimodale présentant des acquisitions de multiples éléments biométriques appartenant à la même personne, n'existe pas sur le web à nos dernières connaissances. Afin de créer notre propre base, notre équipe a participé à un projet PRF¹ intitulé « Plateforme Biométrique Multimodale de Prévention contre le Terrorisme ». Notre contribution dans la formation de la scène biométrique a été par la conception d'un iriscope opérant dans la lumière visible et un autre opérant en proche infrarouge.

Pour l'exploration d'un environnement d'intérieur, l'intérêt a été d'abord orienté vers la réalisation d'un outil d'aide à la navigation pour les non-voyants. L'évolution de nos travaux nous a orienté à élargir le champ de cette application de façon à concevoir un outil d'aide à la navigation qui peut être utilisé comme gadget par des personnes en déficit, comme il peut équiper un robot de service. Dans ce cadre un prototype de recherche a d'abord été conçu dans les laboratoires de l'Institut Supérieur de Biotechnologie de Sfax sous mon encadrement de projets de masters professionnels et de recherche. Nous avons ensuite eu recours aux services d'une boîte de développement électronique pour avoir un prototype plus professionnel qui nous a servi de système d'acquisition multicapteurs de la scène. Les données qu'offre ce système sont de 3 types : de profondeur (ultrasoniques), optiques (caméra mono visible) et de position (gyroscope et accéléromètre). Actuellement, nos travaux de recherche dans cette application sont encore poussés vers l'exploration d'autres capteurs de profondeur et l'étude, tant de la qualité de leurs données que l'évaluation de leur contenu informationnel. Cette étude est menée en vue d'améliorer les performances de notre outil d'aide à la navigation.

Dans ce qui suit, nous commençons par présenter les collaborations effectuées au cours de nos activités portant sur les approches académiques pour l'interprétation de scène puis sur la construction des scènes. Ensuite, les activités entretenues pour la construction d'une scène biométrique, et d'une scène d'environnement d'intérieur sont présentées. Les trois approches proposées de diffusion de connaissances possibilistes dans une scène imagée en vue de sa segmentation, ainsi que le principe de formation d'un carnet d'identité biométrique pour l'identification, sont par la suite détaillés.

¹ PROJETS DE RECHERCHE FEDEREE (PRF) 2017-2020, « Lutte contre le Terrorisme »

II. Synthèse des activités relatives à ce chapitre

1. Collaborations et encadrements

Dans le cadre de mes travaux sur l'interprétation de scène j'ai participé avec Pr. Solaiman dans l'encadrement des travaux de thèse portant sur l'interprétation d'une scène médicale par la diffusion de connaissances possibilistes pour la segmentation :

1. Wael EZZIDIN, « Segmentation itérative d'images par propagation de connaissances dans le domaine possibiliste : Application à la détection de tumeurs en imagerie », thèse sous la direction de Mr. Basel Solaiman, Professeur de l'IMT-Atlantique, Brest, France. Cette thèse a été soutenue le 14/12/2012.

Mon orientation vers une diversité de scènes, pour les applications de mes travaux de recherche, m'a permise de m'ouvrir sur l'environnement professionnel et de collaborer avec de nombreuses parties tant du milieu académique que professionnel.

✓ En biométrie :

Dans le cadre du projet PRF : Plateforme Biométrique Multimodale de Prévention contre le Terrorisme PBMLT, j'ai collaboré avec le laboratoire GEOGLOB de la FSS, le Centre de biotechnologie de Sfax, le laboratoire REGIME de l'ENIS, le laboratoire LATIS de l'ENISO pour le développement des algorithmes nécessaires à la conception de la plateforme multimodale ainsi que pour la conception de la base de données et l'organisation et mise en œuvre de compagnes d'acquisition des éléments biométriques.

Une collaboration a été aussi établie avec des boîtes professionnelles : la société PROCAN pour le développement de nos algorithmes dans l'environnement cloud adapté aux notions des Big Data et la boîte NOVELTI pour la réalisation des irisopes selon nos modèles proposés et conçus.

Des activités de recherche sont en cours de développement avec Pr. Saber Masmoudi, professeur du Centre de biotechnologie de Sfax, concernant la recherche de corrélations entre les informations contenues dans la texture et la couleur de l'iris et celles contenues dans l'ADN

✓ Pour l'aide à la navigation :

Dans le cadre académique, une coopération est entretenue avec Pr. Olfa Kanoun², portant sur l'étude des capteurs de profondeur et l'évaluation qualitative et quantitative du contenu informationnel des données acquise par certains de ces capteurs. Je co-encadre avec Pr. Kanoun une thèse en co-tutelle dans cette thématique :

2. Mouna Medhyoub, « Intégration de méthodes de détection et reconnaissance d'obstacles en temps réel par des capteurs de profondeur pour des systèmes embarqués », thèse en co-tutelle avec Chair for Measurement and Sensor Technology, Technische Universität Chemnitz, Chemnitz, Germany. La thèse est

² Chair for Measurement and Sensor Technology Technische Universität Chemnitz, Chemnitz, Germany

sous la direction conjointe de Mr Nabil Derbel, Professeur de l'Ecole Nationale d'Ingénieurs de Sfax et de Mme. Olfa Kanoun, Professeur de Technische Universität Chemnitz, Chemnitz, Germany.

Dans l'objectif de concevoir et réaliser l'outil d'exploration multicapteur de l'environnement qui répond exactement à nos besoins et nous fournit les données réelles pour nos expérimentations, j'ai partagé la tâche sur des projets de mastères professionnels que j'ai encadré à l'ISBS :

3. Marwa Mhaddheb et Jilani Rinchi, " Acquisition et synchronisation de scène pour une canne blanche électronique. "
4. Sabra Chouaibi et Nahed Missawi, " Conception et réalisation d'un dispositif de communication d'une canne blanche électronique par oreillette Bluetooth. "
5. Imene Abbadi et Khaoula Boukouba, " Synchronisation spatiale et temporelle d'un signal de profil sur une scène par usage d'un accéléromètre. "
6. Khouloud Ben Amor et Chourouk Ben Amor, " Etude d'acceptabilité d'une canne blanche électronique "
7. Zemzemi Nasia, Abir : « Conception et réalisation d'un localisateur d'objets »
8. Imene Dziri et Afef Youssfi, "Etude et conception du dispositif d'acquisition optique pour une canne blanche électronique"
9. Bassem Chrif et Mona Barbar, "Conception et réalisation du dispositif d'acquisition ultrason et du feedback sonore d'une canne blanche électronique".

Nos échanges coopératives avec la société NOVELTI, a permis la réalisation d'un prototype professionnel la base du prototype de recherche conçu au cours de ces mastères professionnelles.

✓ ***Dans le cadre médical :***

Une coopération a été entretenue avec le Service d'Imagerie Médicale. CHU Hédi Chaker., à travers Dr. Hela Fourati Mseddi, qui nous a fourni des vérités terrain de lésions naeviques, ainsi que pour des images scintigraphiques.

2. Publications

Les résultats de ces travaux ont été publiés à travers 1 article de Revue et 4 articles de conférences internationales.

a) Articles de Revue

1. I. Khanfir Kallel, S. Almouahed, E. Bosse and B. Solaiman, "An Iterative Possibilistic Knowledge Diffusion Approach for Blind Medical Image Segmentation", Pattern Recognition, 2018.

b) Conférences internationales à comité de lecture

2. M. Medhioub, S. Ammar Bouhamed, I. Khanfir Kallel, N. Derbel, B. Solaiman and O. Kanoun, "An Electronic Embedded System for Stair Recognition", Sensorica 2017: Series of Lectures of the University of Applied Sciences Ruhr West, June, 8th - 9th
3. S. Ammar, I. Khanfir, and D. Sellami. "A new Approach of hybrid decision tree based stair recognition". International Conference on Control, Engineering &

Information Technology (CEIT'14) Proceedings - Copyright IPCO-2014, pp. 135-139.

4. Sonda Ammar Bouhamed, Imen Khanfir Kallel, Dorra Sellami Masmoudi. « Stair Case Detection And Recognition Using Ultrasonic Signal”, International Conference on Telecommunications and Signal Processing (TSP'13) Jul 02- 04, 2013, Rome, Italy.
5. Sonda Ammar Bouhamed, Jihen Frikha Eleuch, Imen Khanfir Kallel, Dorra Sellami Masmoudi. “New Electronic Cane For Visually Impaired People For Obstacle Detection And Recognition”, IEEE International Conference on Vehicular Electronics and Safety (ICEVES'12), Jul 24- 27, 2012, Istanbul, Turkey.

III. Construction de scènes

1. Scène Biométrique

La collecte et l'annotation de données biométriques d'un échantillon de 500 individus volontaires dans une base prototype constitue la première activité du projet PBMLT. Elle s'étend sur les trois premières années et fait contribuer les cinq équipes de recherche impliquées dans le projet : GEOGLOB (FSS), LPCMC (CBS), CEM-Lab (ENIS), LATIS (ENISo) et REGIM-Lab (ENIS). Notre équipe a assuré la conception et la mise en œuvre de la procédure de collecte des images d'iris d'une part, ainsi que la conception, la mise en œuvre du protocole et l'organisation des campagnes de collectes.

a) Conception du protocole

1. Mise au point d'un scénario de prises (comment se poser, que dire, que faire ...)
2. Mise au point d'un protocole de tests
3. Premiers tests, suivis de mise à jour (maîtrise et validation du matériel)
4. Simulation réelle suivie de mise à jour (prise des données biométriques des membres de l'équipe)
5. Préparation d'une fiche de consentement de collecte des données biométriques suivant les lois Tunisiennes, datant depuis 2004 et appuyées par la Constitution de la République tunisienne en date du 27 janvier 2014. (Commençant par l'Article 24)
6. Planification des dates et des lieux

b) Conception des deux irisopes

Selon les principales propositions du comité de biométrie de l'INCITS (International Committee for Information Technology Standards), les conditions suivantes doivent être vérifiées dans l'acquisition de l'iris :

- ✓ La résolution du système d'acquisition doit être d'au moins 15px/mm dans la zone de l'iris. Facteur de compression JPEG < 6:1.
- ✓ La densité des niveaux de gris doit être supérieure à 256. Ceci est équivalent à dire que le codage de l'information doit s'effectuer sur un minimum de 1 octet par pixel.
- ✓ L'illumination de l'environnement de l'acquisition doit régner dans la gamme de longueurs d'ondes proche infrarouge, entre 700 à 900 nm. L'axe optique d'éclairage doit être supérieur à 5 degrés.

- ✓ Contraste : un minimum de 90 niveaux de gris entre l'iris et la sclérotique et un minimum de 50 niveaux de gris entre l'iris et la pupille sont exigés afin de bien distinguer entre les différentes régions de l'œil et permettre ainsi une bonne localisation de l'iris.
- ✓ Zone de l'iris visible : au moins 70% de l'iris ne doit pas être caché par une réflexion spéculaire ou encore des cils ou des paupières.
- ✓ L'iris doit apparaître avec un diamètre compris entre 160px et 300px et doit être centrée dans l'image. La distance entre les contours de l'iris et les bords de l'image doit être supérieure à 70px.
- ✓ La tête de l'opérateur lors de l'acquisition doit être verticale. La rotation est tolérée entre $\pm 10\%$.

Un capteur classique de résolution convenable, avec une option de prise macro pour les plans rapprochés, pourrait permettre de vérifier ces conditions. Cependant d'autres difficultés sont à défier. En effet, l'image d'iris acquise est souvent perturbée par des reflets du fait que l'iris est localisé derrière la cornée, qui est un miroir hautement réfléchissant. L'iris est aussi couvert, parfois en de grandes zones, par des paupières et des cils dont la détection est difficile (surtout les cils). Le flou peut aussi affecter l'image d'iris acquise dans des conditions non contrôlées. De ce fait les capteurs classiques, même avec de grandes performances de résolution et d'options, ne peuvent pas, isolés, présenter une solution pour l'acquisition de l'iris en vue d'identification. Il faut considérer un système imposant un environnement contrôlé, sur le plan luminosité, réflexions et positions.

Les stations satisfaisant nos suggestions, dites irisopes, ne sont pas disponibles sur le marché local et sont très chères sur le marché étranger. Notre équipe s'est alors engagée à confectionner deux irisopes, l'un opérant sur tout le spectre visible (Cf. Figure 4.1) et l'autre sur uniquement le proche infrarouge (Cf. Figure 4.2). Ce choix s'est effectué suite à une longue étude aussi minutieuse que délicate, pour produire ces deux dispositifs. Outre que les principales propositions de l'INCITS qui ont été parfaitement respectées, nos dispositifs sont pratiques, confortables pour les participants, non nuisibles à leurs yeux et encore faciles d'utilisation. L'utilisation du proche infrarouge est choisie suite à la constatation de deux avantages majeurs par rapport à la lumière visible. Premièrement, la lumière est invisible, l'utilisateur ne sera pas aussi gêné qu'en lumière visible par une puissante illumination. Deuxièmement, le proche infrarouge possède un pouvoir de pénétration de la cornée qui est largement plus grand que celui de la lumière visible et il est ainsi possible de récupérer une richesse de texture supérieure à celle obtenue en lumière visible pour les iris sombres. Cependant, le visible demeure plus intéressant pour les yeux clairs.

Une application embarquée a été également conçue pour permettre l'annotation des images et leur stockage dans une base de données préalablement définie.

L'usinage de ces deux stations, a été effectué par la boîte NOVELTI.



Figure 4.1 Iriscope utilisé pour l'acquisition de l'iris dans le spectre visible



Figure 4.2 Iriscope utilisé pour l'acquisition de l'iris dans le PIR

2. Scène d'environnement intérieur

a) Description

Le système d'aide à la navigation conçu est un dispositif électronique multimédia destiné à l'assistance des personnes en déficience visuelle comme les aveugles ou les personnes âgées, afin de les aider à interagir avec leur environnement et à mieux surmonter leurs incapacités.

Le système d'aide à la navigation conçu vient d'une part défier des limites notées sur les produits qui existent déjà sur le marché et répondre d'autre part à des besoins et des désirs recueillis auprès d'utilisateurs potentiels suite à une étude de besoin réalisée à cet effet.

- 1- Le dispositif est d'une dimension et poids très convenables pour être porté soit en pendentif ou posé sur une canne blanche ordinaire. En effet, les réclamations de look des futurs utilisateurs sont partagées en deux avis, pratiquement selon les âges : les plus jeunes souhaitent avoir une assistance dans l'ombre, de façon à ressembler au maximum aux gens sains. De ce fait, le système d'aide à la navigation conçu leur est offert sous forme d'un boîtier qui se place facilement sur une ceinture et qui apparaît comme un téléphone portable ordinaire. Alors que pour les moins jeunes la canne blanche ne représente pas uniquement un outil d'exploration et un moyen pour éviter des obstacles, mais encore un symbole incitant les autres personnes à céder le passage ou plus encore à les aider. Ils souhaitent alors garder cette apparence symbolique. De ce fait, le système d'aide à la navigation conçu leur est offert sous forme d'un boîtier qui se fixe facilement sur leur canne blanche.
- 2- le système d'aide à la navigation conçu est équipé de :
 - ✓ Deux capteurs ultrason (US) à positions ajustables selon la taille du porteur. L'un des deux capteurs est dirigé vers le haut afin de capter les obstacles en hauteur, l'autre est dirigé vers le bas afin de détecter l'état du sol (plat, escalier ou marche descendant, escalier ou marche ascendant) ;
 - ✓ Une caméra couleur en liaison encastrement avec deux autres capteurs ultrason, l'ensemble étant destiné à acquérir la scène présente devant l'utilisateur de façon à la reconstruire en profondeur et permettre au système d'accomplir la reconnaissance des différents objets existant dans la scène ;
 - ✓ Trois servomoteurs servant à maintenir les positions des différents capteurs afin de préserver la stabilité de leurs acquisitions ;

- ✓ Une oreillette Bluetooth, pour la communication des informations à l'utilisateur ;
- ✓ Un accéléromètre et un gyroscope pour compenser les variations des positions des capteurs lors de la manipulation du dispositif ;
- ✓ Vibreur afin d'avertir l'utilisateur juste avant la réception d'un message vocal
- ✓ Deux cartes de commande à microprocesseur ;
- ✓ Une batterie rechargeable.

La conception du modèle que nous proposons est détaillée dans la Figure 4.3.

- 3- Les dispositifs existants sur le marché assurent convenablement la tâche d'avertissement de la présence d'obstacles. Ceci est un besoin primordial pour les déficients visuels. Le système d'aide à la navigation conçu remplit aussi parfaitement ce rôle, mais prend aussi en considération le fait que les utilisateurs suggèrent également d'être informés de quel obstacle s'agit-il, et à quoi ressemble la scène dans laquelle ils se retrouvent. En effet, grâce à l'emploi de la caméra couleur accompagnée par les capteurs US, le système d'aide à la navigation conçu assure :
 - une détection robuste des obstacles
 - une détermination de leurs distances par rapport à l'utilisateur
 - une reconnaissance de ces obstacles (formes, couleurs, consistance et dimensions)
 - une reconnaissance précise de l'état du sol (sol plat, marche ascendante, marche descendante)
- 4- Lors du déplacement de l'utilisateur, l'acquisition de données est perturbée sous l'effet de la démarche et des balayages aléatoires. L'équipement du système d'aide à la navigation conçu par un accéléromètre et un gyroscope, permet de détecter ces perturbations. Leur compensation est ensuite assurée par une mise à jour continue et temps réel des positions des capteurs par des servomoteurs. Ainsi le système d'aide à la navigation conçu se distingue par un système d'acquisition stable permettant la collecte de données de qualité supérieure aux autres dispositifs.
- 5- Les dispositifs existants sur le marché consistent à transmettre un message à l'utilisateur par voie tactile, vibratoire ou bips sonores. De ce fait, un apprentissage d'utilisation est indispensable, ce qui n'est pas à la portée cognitive ou encore matérielle de tous les utilisateurs. Le système d'aide à la navigation conçu présente le retour d'information sous forme d'un retour informationnel commun, facile à comprendre beaucoup plus accepté par les utilisateurs. Il s'agit d'un retour vocal, sous une langue choisie par l'utilisateur dès la réception de l'appareil, précédé par une légère vibration (pour avertir l'utilisateur de la réception du message vocal). De ce fait, une formation n'est plus nécessaire.
- 6- La détection et la reconnaissance des obstacles ont une grande importance dans un système d'aide aux déplacements. Les systèmes de détection d'obstacles existants sur le marché permettent la détection des objets en relief par rapport au sol mais ne sont pas aptes à la détection des baisses du niveau de sol, tels que les escaliers descendants ou marches descendantes ou même des fossés. Alors, l'incapacité de reconnaître ces entraves, peut entraîner un vrai danger pour la sécurité des personnes à déficience visuelle. Dans notre invention et contrairement aux autres dispositifs, un intérêt

particulier est orienté vers les entraves qui présentent une sérieuse menace à la sécurité des non-voyants. Nous citons spécialement le danger posé lors du changement de l'état du sol. Le système de détection proposé vise à la reconnaissance de tel changement. Une telle fonctionnalité est importante dans un système d'aide destiné aux personnes à déficience visuelle.

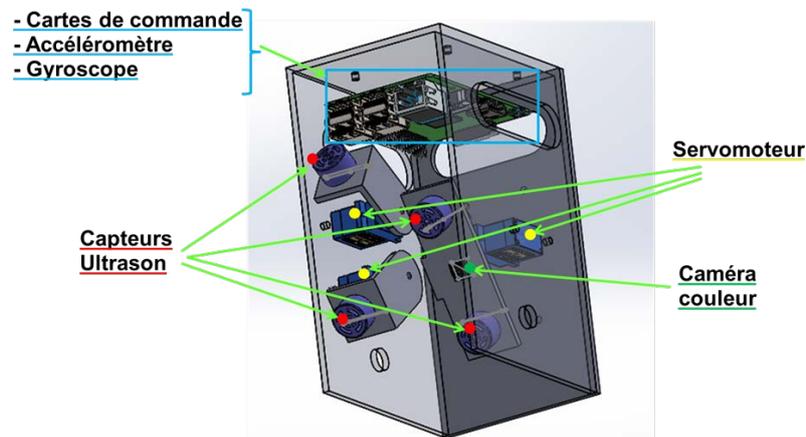


Figure 4.3 Vue en 3d du dispositif du le système d'aide à la navigation conçu
b) Schéma fonctionnel

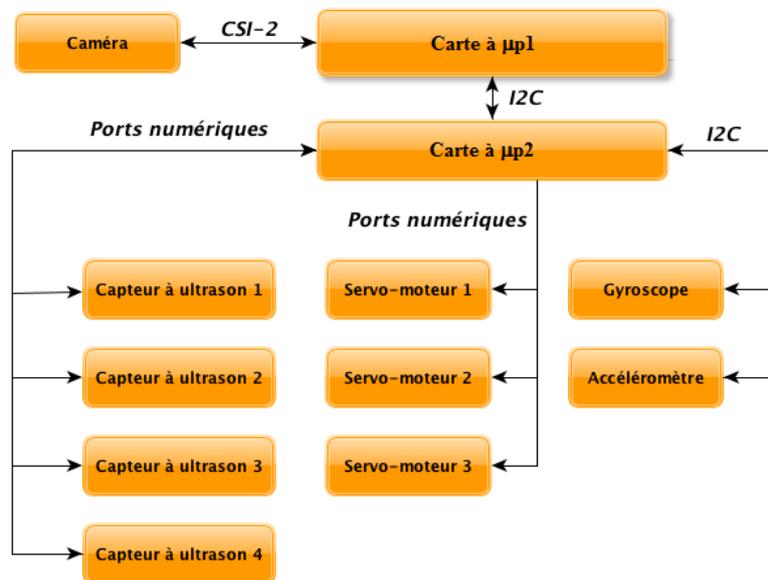


Figure 4.4 Schéma fonctionnel du système d'aide à la navigation conçu

IV. Diffusion de connaissances possibilistes itératives, application à l'interprétation de scène médicale

1. Problématique

La segmentation constitue une étape importante pour l'extraction des informations qualitatives et quantitatives de l'image, et par conséquent, de la scène réelle analysée. Elle fournit une description de haut niveau de l'image analysée : l'étiquette de chaque région détectée, la forme de chaque région (taille, orientation, ...), la distribution spatiale (connexions, voisinages, ...) des différentes régions, etc.

La majorité des méthodes utilisées en segmentation d'image se base uniquement sur l'analyse des niveaux de gris observés. La question que l'on peut soulever est la suivante : est-ce que le niveau de gris étant proportionnel à la valeur mesurée par le capteur, est le meilleur choix pour réaliser la segmentation de l'image?

Pour répondre à cette question, on propose d'analyser l'histogramme des niveaux de gris de deux classes thématiques contenues dans une image (Cf. Figure 4.5).

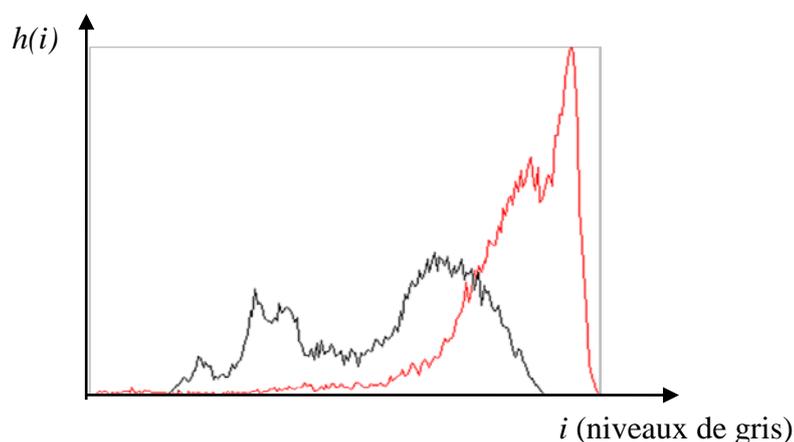


Figure 4.5 Un histogramme bimodal montrant un exemple de chevauchement entre deux classes thématiques

On note des zones de chevauchement importantes. L'effet de cette imperfection, de nature statistique, peut être réduit en analysant les relations spatiales entre les pixels. Ceci revient à considérer la distribution texturale des niveaux de gris, qui paraît avoir un niveau sémantique plus élevé que celui de l'information fournie par les niveaux de gris uniquement. Ainsi on aura affaire à plusieurs caractéristiques de natures différentes. Il semble donc qu'une « bonne méthode de segmentation » doit prendre en considération ces caractéristiques tenant compte de leurs facteurs d'incertitude et d'ambiguïté. Notons qu'une telle incertitude et une telle ambiguïté peuvent être traitées en utilisant les méthodes de segmentation à base de processus itératif capable de réduire l'ambiguïté de l'information (liée au contenu de l'image) via l'intégration de connaissances spatiales. De plus, lors de la classification des pixels de l'image, l'incertitude sera diminuée parce que l'information manipulée devient plus fiable grâce à la diffusion spatiale des connaissances via un processus itératif. Pour cette raison, les

méthodes de segmentation d'images proposées dans la littérature, peuvent être présentées en deux familles selon leur nature itérative.

- les méthodes non itératives de segmentation d'images fournissent une image segmentée en réalisant un seul passage d'analyse du contenu informationnel de l'image. Ainsi, la segmentation n'admet pas de résultats intermédiaires pour raffiner ceux qui les succèdent et n'offrent aucun moyen afin d'évaluer et de modifier d'une manière dynamique le résultat de la segmentation dans l'objectif d'augmenter la fiabilité des résultats obtenus. Parmi ces méthodes, on cite celles basées sur la notion de classification de pixels, dites aussi par seuillage d'histogramme [DSW 2017, GW 2002, MKK 2009], les méthodes de segmentation bayésienne [ARB 2007, LWL 2001, GG 1984] et celles basées sur l'analyse texturale et morphologique [CSM 2006, LH 2007, Tang 1998].
- Les méthodes itératives offrent l'avantage de répéter le processus de segmentation afin d'améliorer, au fur et à mesure, sa qualité. leur principe consiste à exploiter l'information du voisinage des pixels, appelée le contexte spatial pour augmentation le pouvoir de discrimination des différents pixels, améliorant ainsi la fiabilité de leur classification. Une analyse de l'état de l'art des méthodes de segmentation itérative nous permet de les classer en plusieurs catégories :
 - ✓ Les méthodes basées sur l'estimation itérative des paramètres de segmentation [GG 1984, YC 1997, KKB 2004];
 - ✓ Les méthodes de segmentation à base de contours actifs [ZXZ 2018, KWT 1988, KMH 2005, Ros 2002, Lec 2005];
 - ✓ Les méthodes de segmentation markoviennes [GG 1984];
 - ✓ Les méthodes de segmentation par découpage itératif d'image (découpage Top-down) [KBW 2009, RMJ 2002];
 - ✓ Les méthodes de croissance de régions (Bottom-Up) [KC 2002, BGE 1996, Til 1989, Til 1988]; et
 - ✓ Les méthodes hybrides de segmentation d'image [KC 2002, WH 1999, IMT 2007, KBC 1994, HS 2010].

L'analyse des principes et des performances de différentes méthodes de segmentation (non itératives et itératives) permet de déduire les remarques suivantes :

- ✓ Pour les images à la fois de faible contraste et de nature fortement texturée, toutes les méthodes de segmentation proposées dépendent très fortement des valeurs choisies des paramètres (seuils, valeurs moyennes, variances ...) et des modèles exploités par ces méthodes (densité de probabilité, fonction d'appartenance ...). Par conséquent, une faible erreur d'estimation de ces paramètres ou de ces modèles peut aboutir à des résultats de segmentation de qualités médiocres en termes de taux d'erreur au niveau pixel. Cette mauvaise estimation risque également de faire rater la détection des petites régions;
- ✓ Les méthodes de segmentation itératives aboutissent, d'une façon incontestable, à des résultats de meilleures qualités que les méthodes non itératives ;

- ✓ Les méthodes de segmentation par croissance de régions se basant sur les mesures liées aux niveaux de gris (y compris des mesures texturales) où des mesures probabilistes, apportent une bonne identification initiale des régions d'intérêt mais souffrent de l'inconvénient majeur de localisation des contours de ces régions.
- ✓ L'approche de segmentation par les contours actifs présente de bons résultats au niveau de la localisation des contours des régions d'intérêt à condition que l'initialisation de ces contours ne soit pas trop éloignée des contours finaux. Néanmoins, la nature texturée des images fait souvent aboutir à de multiples faux contours à l'intérieur des régions détectées.
- ✓ Les approches hybrides capables de combiner les informations issues de plusieurs méthodes semblent être les plus prometteuses en segmentation d'images.

En guise de conclusion de cet état de l'art des méthodes de segmentation d'images, on peut souligner que les lacunes constatées sur les méthodes les plus prometteuses, basées sur les approches hybrides et sur la croissance de régions, ont pour origine la propagation spatiale des imperfections telles que les incertitudes, les imprécisions et les ambiguïtés, qui sont dues à la nature des images et à la forte dépendances des méthodes appliquées par rapport à l'initialisation des paramètres et des modèles exploités par ces méthodes.

Dans ce qui suit, nous proposons une approche de segmentation itérative d'images destinée à tirer profit de divers avantages des méthodes que nous venons d'analyser. Cette approche est basée sur la propagation des connaissances spatiales, dans un espace différent de celui des niveaux de gris ou des mesures probabilistes. Il s'agit de l'espace des connaissances possibilistes.

2. Principe

Les considérations précédentes nous conduisent à proposer une approche de segmentation d'image permettant de mieux interpréter le contenu informationnel de l'image. Cette approche, fortement inspirée des méthodes de croissance de région, consiste, dans un premier temps, à transposer les différentes sources de connaissances disponibles en un espace possibiliste de représentation des connaissances, et d'opérer, dans un second temps, un processus de diffusion des connaissances possibilistes (d'une façon similaire à l'opération de croissance de régions) sur la base d'informations contextuelles.

L'architecture conceptuelle de l'approche de segmentation que nous proposons est schématisée sur la Figure 4.6.

Cette architecture est composée de quatre systèmes :

1. Système de prétraitement ;
2. Système de décision ;
3. Système de diffusion possibiliste ; et
4. Système d'évaluation et de mise-à-jour.

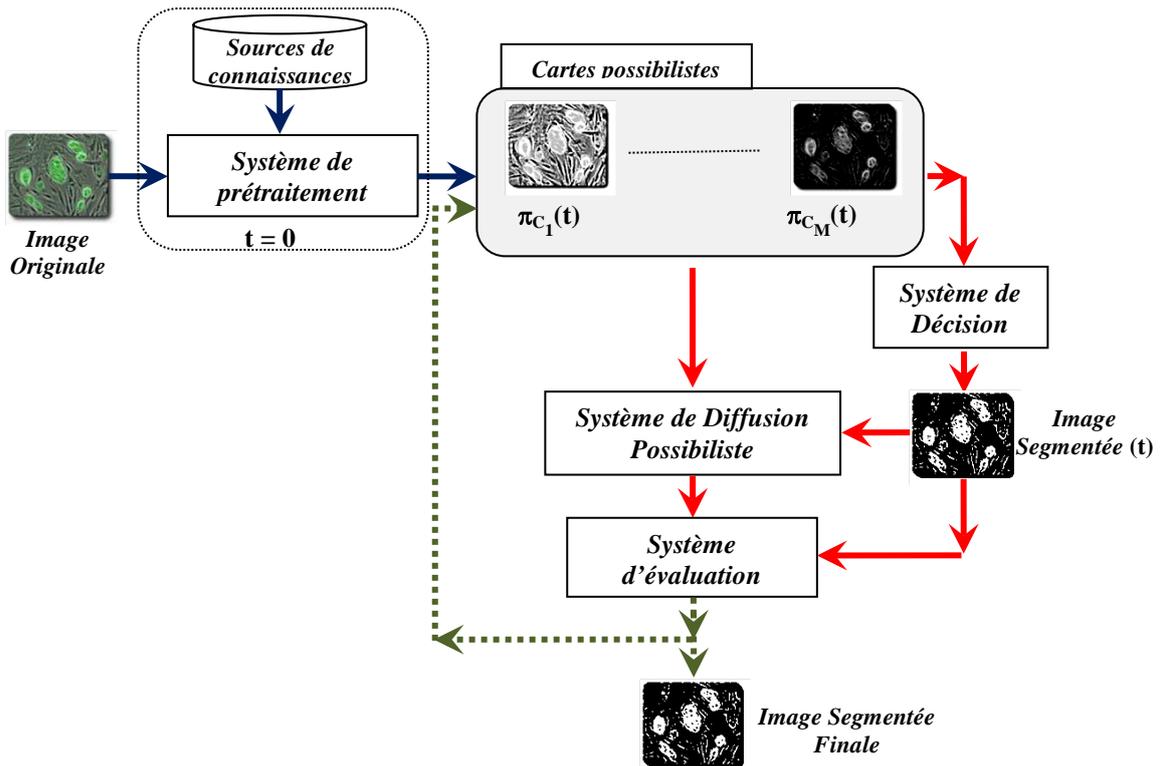


Figure 4.6 Architecture conceptuelle de l'approche itérative de segmentation

Afin de détailler ces systèmes nous supposons que :

- les images à l'entrée du système de segmentation sont de taille $K \times K$;
- chaque pixel (i,j) , i et $j \in \{0, 1, \dots, K-1\}$ est observé sur N niveaux de gris ($x(i,j) \in \{0, 1, \dots, N-1\}$) ;
- les images comportent M classes thématiques $\Omega = \{C_1, C_2, \dots, C_M\}$; et
- nous opérons sous l'hypothèse du monde fermé (i.e. Les M classes thématiques sont les seules classes rencontrées dans les images considérées).

a) **Système de prétraitement :**

Le système de prétraitement a pour objectif d'appliquer les différentes méthodes de prétraitement sur l'image originale (filtrage, rehaussement, etc.), d'extraire l'ensemble des primitives jugées pertinentes et d'estimer, en sortie, pour chaque pixel de position (i,j) une distribution « initiale » de possibilités définie sur l'ensemble des classes thématiques $\Omega = \{C_1, C_2, \dots, C_M\}$:

$$\begin{aligned} \pi_{t=0,(i,j)}: \quad \Omega &\longrightarrow [0, 1] \\ C_m &\longrightarrow \pi_{t=0,(i,j)}(C_m) \end{aligned} \quad \text{Eq4.1}$$

Cette estimation initiale des distributions de possibilités est réalisée en exploitant des sources de connaissance *a priori* liées à la forme des distributions de possibilités ou à une base d'échantillons représentant les différentes classes thématiques considérées.

Pour l'ensemble des positions (i,j) , i et $j \in \{0, 1, \dots, K-1\}$, ces distributions de possibilités peuvent être représentées, Figure 4.7, sous la forme de M images possibilistes, que l'on appelle *cartes possibilistes initiales* :

$$\pi_{t=0}(C_m) = \{ \pi_{t=0,(i,j)}(C_m) \}_{i \text{ et } j \in \{0, 1, \dots, K-1\}}, \quad m = 1, 2, \dots, M \quad \text{Eq4.2}$$

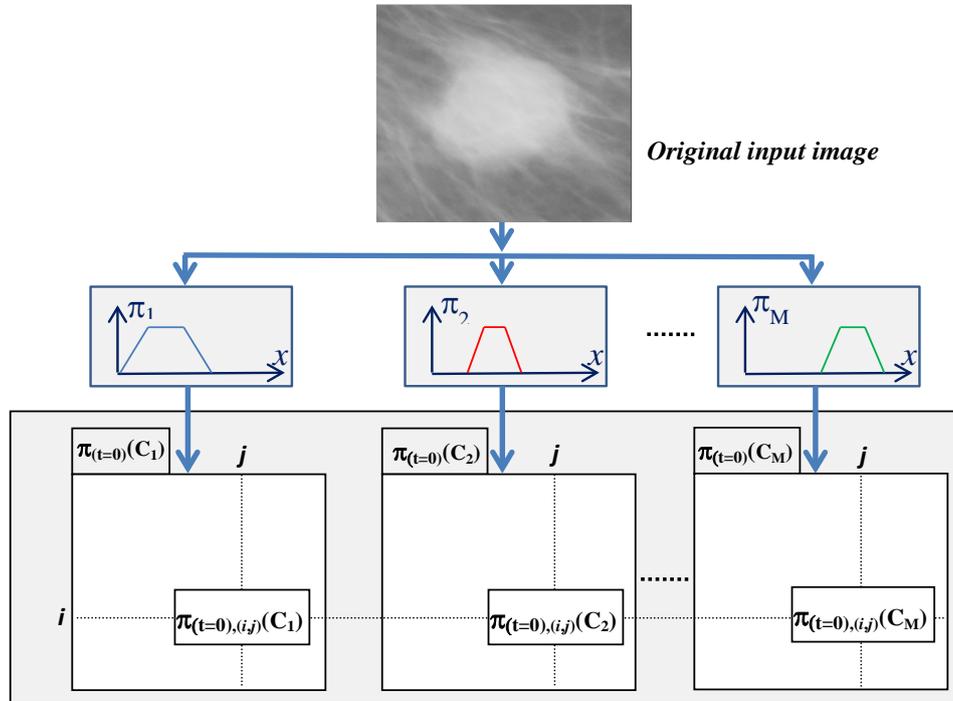


Figure 4.7 Les M cartes possibilistes initiales en sortie du système de prétraitement

Dans ce sens, notre système d'estimation des M cartes possibilistes initiales procède en trois phases: celle de génération de sources de connaissances, celle d'estimation de distributions de possibilités selon chacune des sources de connaissances et enfin celle de fusion de ces distributions afin d'en avoir qu'une seule carte possibiliste.

Dans la première phase, R filtres sont appliqués séparément. Si l'application de ces filtres offre une atténuation du bruit et une meilleure homogénéité dans l'image permettant ainsi l'amélioration de la fiabilité et la robustesse des résultats de l'estimation, leurs résultats respectifs présentent carrément des sources de connaissances. En effet chacun de ces filtres fait apparaître, selon sa manière d'opération, des informations non aperçues dans l'image à son état brute. Les R filtrages fournissent R images différentes de l'image originale, contenant des informations plus pertinentes, différentes les unes des autres tout en étant complémentaires. Ces filtres jouent ainsi le rôle d'observateurs dans notre système. Chaque pixel appartenant à une classe donnée est donc représenté par un vecteur de R observations.

La deuxième phase de notre système exploite les données issues de l'étape de prétraitement afin d'estimer les différentes fonctions d'appartenances liées aux classes contenues dans l'image. L'estimation est du type indirect guidé par les connaissances d'expert. Chaque classe de l'image est représentée par R fonctions d'appartenance correspondant aux R sources de connaissances. Grâce au postulat possibiliste de zadeh [Zad 1978], un passage à l'espace possibiliste est effectué. Les fonctions d'appartenances sont ainsi considérées comme des distributions de possibilités. Nous avons nommé les images associées à ces distributions : cartes possibilistes.

Une fusion des R cartes possibilistes issues des R sources est ensuite effectuée, afin d'obtenir pour chaque classe une seule carte possibiliste.

b) Système de décision :

Le système de décision a pour objectif d'analyser les M cartes possibilistes correspondant aux M classes, présentées à son entrée, dans l'objectif de produire, en sortie, une carte thématique « particulière ». La particularité de cette carte (appelée *Image segmentée*(t), $t = 0, 1, \dots, t_{\text{final}}$) réside dans le fait que chaque pixel peut se voir attribuer une décision de classification à une classe appartenant à Ω , ou, un label indiquant que l'état des connaissances possibilistes disponibles ne permet pas d'attribuer une décision à ce pixel. Il s'agit d'une classe de rejet de classification temporaire.

c) Système de diffusion possibiliste :

Le système de diffusion possibiliste, constitue le véritable cœur de l'approche de segmentation proposée. Le rôle de ce sous-système est d'analyser les connaissances possibilistes disponibles pour chaque pixel de l'image afin d'apporter des mises à jour permettant de réduire l'incertitude rencontrée pour la prise de décision de classification. C'est ce système qui est supposé réaliser l'opération de croissance de régions dans l'espace de représentation possibiliste des connaissances au lieu de l'espace « habituel » des niveaux de gris. Notons que cette mise-à-jour peut être réalisée en exploitant des connaissances potentielles issues de l'image segmentée de l'itération précédente. Cette exploitation peut se limiter à apporter une contrainte (indiquant les seuls pixels pour lesquels la mise à jour des connaissances possibilistes doit être réalisée), mais peut aussi bien s'étendre à l'exploitation des connaissances sémantiques qui y sont contenues.

d) Système d'évaluation et de mise-à-jour :

Le système d'évaluation et de mise-à-jour des connaissances a une double mission :

- Au niveau pixellique, ce système est en charge d'évaluer l'incertitude décisionnelle liée à la distribution des possibilités produite par le système de diffusion possibiliste. Cette évaluation a pour objectif d'enrichir l'ensemble de pixels constituant les germes de référence des différentes classes. Dans le cas où un pixel devient un « germe de référence » d'une classe donnée, alors, cette décision est considérée comme irréversible et sa distribution des possibilités sera modifiée en conséquence.
- Au niveau global, ce système est en charge d'évaluer la nécessité de poursuivre le processus itératif de diffusion possibiliste des connaissances. Dans le cas où la décision est d'arrêter ce processus, ce système produira en sortie une image segmentée dans laquelle chaque pixel est associé à un label de classe (appartenant à Ω), ou, un label de « rejet » indiquant que malgré les différentes itérations de diffusion des connaissances, le système n'est pas parvenu à prendre une décision. L'ensemble des pixels labélisés « rejet » doit être analysé et les raisons de rejet doivent être identifiées (hypothèse du monde fermé, pixels comportant des mélanges de classes, connaissances initiales non valides, etc.)

L'approche qu'on propose permet d'établir une sorte d'analogie avec le raisonnement humain lors de l'analyse perceptuelle d'une image. En effet, l'interprétation humaine du contenu d'une image (en termes de sa segmentation en régions homogènes) consiste à « repérer » des zones de référence de chaque classe et d'enrichir, ensuite, ces zones d'une façon incrémentale en analysant leur voisinage spatial afin de raffiner les possibilités d'appartenance aux différentes classes. Cet enrichissement est assimilé au processus de

propagation par diffusion de connaissances possibilistes visant à raffiner les valeurs de possibilité des pixels encore labélisés comme étant de la classe rejet.

Le système de segmentation proposé offre les avantages suivants :

- L'utilisation du formalisme de la théorie des possibilités permet de décrire les informations disponibles ainsi que les différentes formes d'imperfection (incertitude, ambiguïté, etc.) liées à ces informations dans un même cadre. La raison de cette orientation réside dans la flexibilité des outils associés à la représentation des connaissances possibilistes.
- Le principe de propagation de connaissances au voisinage spatial des pixels permet l'ajustement des distributions de possibilités afin de diminuer l'imperfection inhérente à la connaissance dont nous disposons sur les classes des pixels.
- L'utilisation de la théorie des possibilités permet de développer des mécanismes possibilistes de décision d'une façon analogue au raisonnement d'un observateur humain.

Nous proposons dans ce qui suit d'étudier deux familles de méthodes de diffusion des connaissances possibilistes qui n'exploitent pas d'autres sources de connaissances que les seules cartes possibilistes. En d'autres termes, ni les connaissances que l'on peut obtenir par la classification des cartes possibilistes, ni les germes de références des classes ne sont mis à contribution dans le processus itératif de modification des cartes possibilistes.

3. Diffusion classe/classe

La première stratégie que nous présentons pour la diffusion des connaissances possibilistes est appelée la diffusion possibiliste Classe/Classe. Elle consiste à réaliser la mise-à-jour des M cartes possibilistes $\pi_t(C_m) = \{\pi_{t,(i,j)}(C_m)\}_{i,j \in \{0, 1, \dots, K-1\}}$, $m = 1, 2, \dots, M$ d'une façon indépendante les unes des autres :

$$\pi_{t+1}(C_m) = g(\pi_t(C_m)), m = 1, 2, \dots, M \quad \text{Eq4.3}$$

Ce qui est schématisé dans la Figure 4.8.

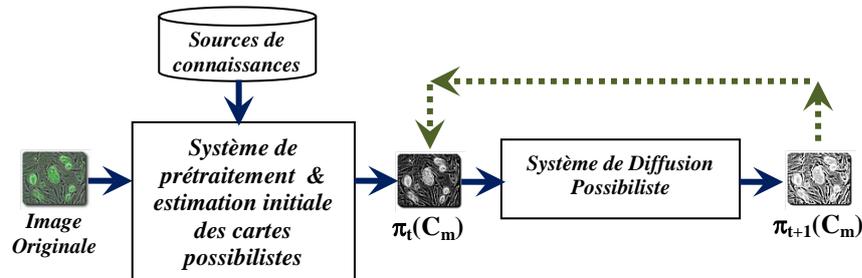


Figure 4.8 Stratégie de diffusion possibiliste Classe/Classe pour une classe thématique

Cette stratégie est, donc, composée de deux étapes indépendantes :

- La première étape consiste à exploiter l'ensemble des sources de connaissances disponibles afin de « transcrire » les mesures physiques, issues du capteur, dans un espace de représentation possibiliste lié à chacune des classes thématiques contenues dans la scène imagée. Par conséquent, la carte, ou image, possibiliste en sortie de cette étape désigne, pour chaque pixel, la valeur de possibilité pour que ce pixel soit d'une classe donnée.
- La deuxième étape consiste à appliquer des méthodes de traitement d'image afin de diffuser les connaissances possibilistes dans le contexte spatial des différents pixels. Cette étape prendra en compte le niveau sémantique de l'information qui est contenue

dans les cartes possibilistes au lieu du niveau physique représenté par les niveaux de gris mesurés en sortie de capteur (i.e. image originale).

Notons que cette stratégie de diffusion s'applique d'une façon indépendante pour chaque classe thématique sans exploiter, d'une façon croisée, les cartes possibilistes des autres classes thématiques.

Plusieurs techniques de lissage d'images peuvent être employées pour réaliser cette diffusion des connaissances. Nous pouvons en citer : le filtrage moyenneur, filtrage de Nagao, la diffusion anisotropique, les filtres de Gabor, ... etc.

Dans ce qui suit, nous présentons le principe de diffusion par filtre de Nagao, étant la technique adoptée pour cette cause.

a) La diffusion par filtrage de Nagao

En filtrage d'image, le filtrage moyen a pour effet intéressant de réduire la variation texturale rendant ainsi les différentes régions de plus en plus homogènes. Néanmoins, un inconvénient majeur de ce filtrage réside dans le fait qu'il affecte les pixels proches des contours et ceux appartenant aux zones homogènes de la même manière. Ceci aura pour effet de rendre les contours des régions de plus en plus flous.

Le filtre de Nagao [NM 1997] constitue l'une des solutions proposées pour remédier à ce problème. Il consiste, d'abord, à considérer neuf « sous-fenêtres » dans le voisinage de chaque pixel. Celle qui est la plus similaire au pixel central, dite sous-fenêtre gagnante, lui donnera la moyenne de ses niveaux de gris. En d'autres termes, il s'agit de restreindre l'application du filtrage moyen à une sous-fenêtre, qui est considérée comme étant la plus « homogène » dans le voisinage de ce pixel, Figure 4.9.

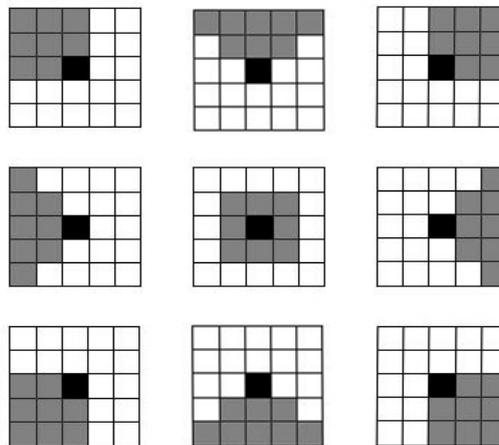


Figure 4.9 La configuration des sous-fenêtres du filtre de Nagao

Notons, $F = \{f_1, f_2, \dots, f_9\}$ l'ensemble des 9 sous-fenêtres contenues dans un voisinage 5x5 de chaque pixel, tel qu'illustré à la Figure 4.9. La sous-fenêtre gagnante, $f_{gagnante}$, est celle ayant la variance minimale. Par conséquent, la diffusion des connaissances possibilistes par le filtre de Nagao est donnée par :

$$\pi_{t+1, (i, j)}(C_m) = \frac{1}{9} \sum_{(\Delta_1, \Delta_2) \in f_{gagnante}} \pi_{t, (i+\Delta_1, j+\Delta_2)}(C_m) \quad \text{Eq4.4}$$

$i \text{ et } j \in \{0, 1, \dots, K-1\}, m = 1, 2, \dots, M$

b) La diffusion par filtrage de Nagao modifié

Le filtre de Nagao classique exploite deux principes :

1. critère de sélection d'une sous-fenêtre « gagnante » qui est fondé sur la variance minimale des valeurs de possibilité ; et,
2. la mise-à-jour des valeurs de possibilités qui consiste à substituer la valeur de possibilité par la valeur moyenne calculée sur la sous-fenêtre gagnante.

En d'autres termes, l'application de ce filtre « transpose » l'approche classique pour mesurer l'homogénéité de l'espace des données numériques issues des capteurs ainsi que la façon de mettre à jour des connaissances à l'espace des connaissances possibilistes contenues dans les sous-fenêtres. En considérant les 9 sous-fenêtres de Nagao comme étant 9 « experts » parmi lesquels on doit choisir l'expert le plus « homogène », alors le critère de variance minimale mesure la variation des connaissances possibilistes portées par les pixels composant cet expert (i.e. sous-fenêtre).

En conservant le critère de sélection de sous-fenêtre gagnante, on note que la mise à jour des connaissances possibilistes n'est qu'une combinaison linéaire des mesures possibilistes contenues dans la sous-fenêtre gagnante. A ce stade, on peut introduire une approche « non-linéaire » de mise-à-jour des connaissances possibilistes, Figure 4.10, qui consiste à appliquer un choix « pessimiste », resp. « optimiste », si la valeur de possibilité du pixel central est proche de la valeur minimale, resp. maximale, des possibilités contenues dans la sous-fenêtre gagnante.

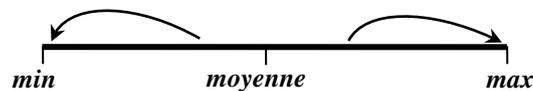


Figure 4.10 Mise à jour non linéaire des connaissances possibilistes

Ce choix peut être, donc, considéré comme un choix qui tient compte du contenu sémantique des connaissances possibilistes véhiculées par l'espace possibiliste.

4. Diffusion croisée

A la différence de la stratégie de diffusion possibiliste Classe/Classe, la stratégie de diffusion croisée, schématisé dans la Figure 4.11, consiste à exploiter l'ensemble des cartes possibilistes $\pi_t(C_m)$, $m = 1, 2, \dots, M$ pour réaliser la mise-à-jour de chacune de ces cartes à l'itération suivante $t+1$:

$$\pi_{t+1}(C_m) = g(\pi_t(C_1), \dots, \pi_t(C_m), \dots, \pi_t(C_M)), \quad m = 1, 2, \dots, M \tag{Eq4.5}$$

où g est la fonction de diffusion.

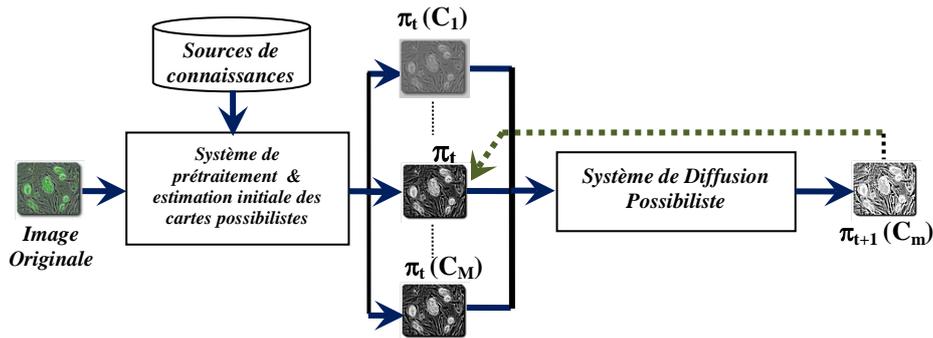


Figure 4.11 Stratégie de diffusion possibiliste croisée

Notons que la différence fondamentale entre cette stratégie et celle de la diffusion possibiliste Classe/Classe réside dans le fait que l'ensemble des cartes possibilistes de toutes les classes thématiques sont considérées comme des sources de connaissances additionnelles pour la diffusion spatiale des connaissances possibilistes au niveau de chaque nouvelle carte possibiliste.

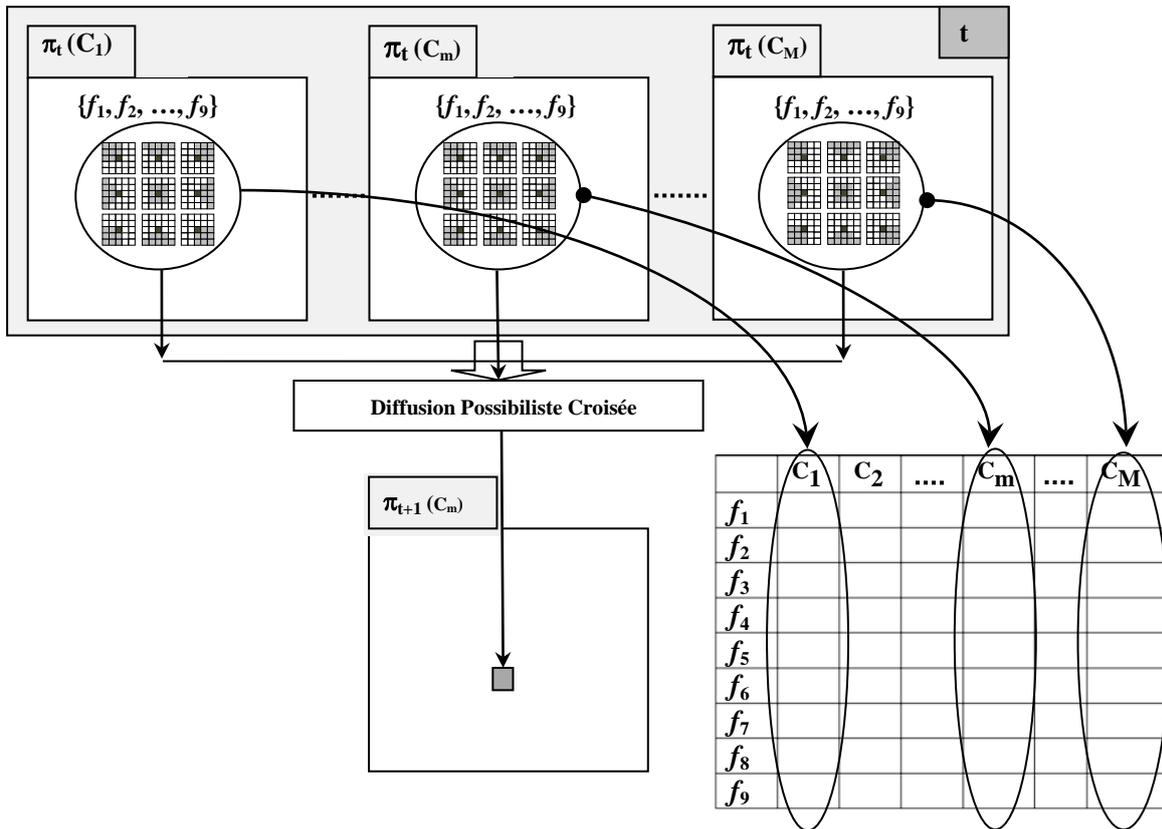


Figure 4.12 Stratégie de diffusion possibiliste croisée basée sur les cliques de Nagao

Nous proposons dans ce qui suit une méthode de diffusion croisée des connaissances possibilistes basée sur l'utilisation des cliques définies par le filtre de Nagao.

En effet, l'approche de diffusion Classe/Classe basée sur l'utilisation des sous-fenêtre définies par le filtre de Nagao, consiste à sélectionner, pour un pixel donné, une sous-fenêtre gagnante au niveau de chaque carte possibiliste et ceci afin de « mettre à jour » les

connaissances possibilistes à l'itération suivante. En revanche, la méthode que nous proposons consiste à identifier (pour chaque pixel) une sous-fenêtre « optimale » parmi l'ensemble des sous-fenêtres de Nagao et d'utiliser cette même sous-fenêtre afin de mettre à jour l'ensemble des cartes possibilistes.

Supposons que les images considérées comportent M classes thématiques $\Omega = \{C_1, C_2, \dots, C_M\}$ et notons, $F = \{f_1, f_2, \dots, f_9\}$, l'ensemble des 9 sous-fenêtres spatiales (i.e. cliques définies par le filtre de Nagao) contenues dans un voisinage 5x5 de chaque pixel, Figure 4.9.

Pour un pixel P donné, chaque carte possibiliste $\pi_t(C_m)$, $m = 1, 2, \dots, M$, apporte 9 informations spatiales correspondant aux neuf cliques. Chacune de ces informations spatiales est résumée par la valeur moyenne des mesures possibilistes véhiculées par la clique considérée et ceci pour une carte possibiliste donnée. Cette représentation permet de résumer l'ensemble des connaissances possibilistes sous la forme d'une matrice possibiliste ayant 9 lignes (correspondant chacune à une clique) et M colonnes (correspondant chacune à une classe donnée), Figure 4.12.

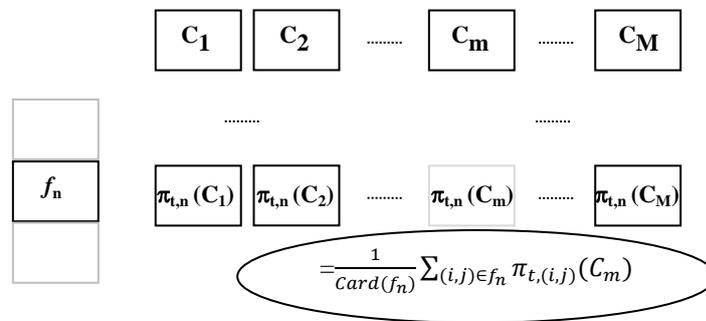


Figure 4.13 Estimation d'une distribution de possibilités définie sur l'ensemble des classes et associée à chaque clique de Nagao

Pour chaque clique f_n , $n = 1, 2, \dots, 9$, nous pouvons associer un degré de possibilité pour chacune des classes thématiques qui est la valeur moyenne des mesures de possibilités calculées sur l'ensemble des pixels de cette clique dans la carte des possibilités correspondante à la classe considérée, Figure 4.13.

En d'autres termes, on considère que chaque clique f_n , $n = 1, 2, \dots, 9$, est une source permettant de définir, à l'itération t , une distribution de possibilités $\pi_{t,n}$ sur l'ensemble des M classes thématiques $\Omega = \{C_1, C_2, \dots, C_M\}$.

Le cœur de la méthode proposée de diffusion croisée des connaissances possibilistes consiste à sélectionner la clique optimale en fonction des distributions de possibilité $\pi_{t,n}$, $n = 1, 2, \dots, 9$.

Considérons une mesure particulière d'incertitude, appelée *Indice de Confiance*, permettant d'évaluer la quantité d'information utile pour la prise de décision dans chacune des neuf distributions de possibilités, correspondant aux neuf cliques de Nagao. Cet indice de confiance est défini de la manière suivante :

Soient Π et N les mesures de possibilité et de nécessité associées à une distribution de possibilités π (i.e. $\Pi(A) = \max_{C_m \in A} \pi(C_m)$ et $N(A) = 1 - \Pi(A^c)$, $\forall A \subseteq \Omega$). Alors, l'indice de confiance en la réalisation d'un événement A est défini par :

$$Ind(A) = \Pi(A) + N(A) - 1 \quad \text{Eq4.6}$$

Notons que :

- $Ind(A) \in [-1,1]$;
- $Ind(A) = -1$, si $N(A) = \Pi(A) = 0$ (ce cas désigne une totale invalidation de la réalisation de A); et
- $Ind(A) = +1$, si $N(A) = \Pi(A) = 1$ (ce cas désigne la totale validation de la réalisation de A).

Alors, à une itération t donnée, et en considérant une clique f_n , $n=1, 2, \dots, 9$, chaque classe thématique sera caractérisée par la mesure de l'indice de confiance :

$$\begin{aligned} Ind_{t,n}(C_m) &= \Pi(\{C_m\}) + N(\{C_m\}) - 1 \\ &= \pi_{t,n}(C_m) + 1 - \Pi(\{C_m\}^C) - 1 \\ &= \pi_{t,n}(C_m) + 1 - \max_{k \neq m} \pi_{t,n}(C_k) - 1 \\ &= \pi_{t,n}(C_m) - \max_{k \neq m} \pi_{t,n}(C_k) \end{aligned} \quad \text{Eq4.7}$$

$Ind_{t,n}(C_m)$ étant une fonction de la clique et de la classe thématique considérées, mesure donc l'écart entre la valeur de possibilité pour que la clique considérée soit associée à la classe C_m et la valeur maximale des possibilités correspondantes aux autres classes.

Si $Ind_{t,n}(C_m)$ est égal à l'unité, alors ceci signifie que la valeur moyenne des possibilités pour la classe C_m est égale à l'unité et que cette valeur moyenne des possibilités est nulle pour toutes les autres classes C_k , $k \neq m$. En revanche, si $Ind_{t,n}(C_m) = -1$ désigne le cas où aucun pixel de la clique n'a de possibilité d'appartenir à la classe et qu'il existe une autre classe pour laquelle toutes les valeurs de possibilités sont égales à l'unité.

Ainsi, on définit un indice de confiance globale associé à chaque clique f_n , $n=1, 2, \dots, 9$, par :

$$Ind_{t,n} = \max_m Ind_{t,n}(C_m) \quad \text{Eq4.8}$$

La clique optimale, qui sera considérée pour la mise-à-jour des connaissances possibilistes, sera celle ayant l'indice de confiance maximale :

$$f_{t,n_{\text{Optimale}}} \leftrightarrow Ind_{t,n_{\text{Optimale}}} = \max_{n \in \{1, 2, \dots, 9\}} Ind_{t,n} \quad \text{Eq4.9}$$

En d'autres termes, la clique $f_{t,n_{\text{Optimale}}}$ est celle qui comporte une classe thématique pour laquelle le couple $(f_{t,n_{\text{Optimale}}}, C_{m_{\text{Optimale}}})$ présente l'indice de confiance le plus élevé sur l'ensemble des 9 cliques et des M classes thématiques.

Finalement, la mise-à-jour des connaissances possibilistes à l'itération « $t+1$ » sera réalisée en affectant la distribution de possibilités $\pi_{t,n_{\text{Optimale}}}$ à l'ensemble des cartes possibilistes et ceci à la position du pixel considéré.

En résumé, l'algorithme de diffusion croisée des connaissances possibilistes est le suivant :

Pour (chaque pixel P de l'image) **Faire**

1. Extraction d'un voisinage 5×5 centré sur P de l'ensemble des cartes possibilistes ;
2. Calcul de la matrice de taille $9 \times M$ des valeurs moyennes des possibilités de l'ensemble des cliques de Nagao et des classes thématiques ;
3. Calcul des indices de confiance pour l'ensemble des cliques de Nagao ;
4. Détermination de la clique optimale ;

5. Mise à jour des connaissances possibilistes dans l'ensemble des cartes possibilistes de l'itération « t+1 » à la position du pixel P.

Fin

5. Contribution : Diffusion croisée avec une représentation possibiliste des connaissances dans une clique

Dans la stratégie de diffusion possibiliste croisée décrite précédemment, la mise à jour des cartes possibilistes se fait par l'attribution, à chaque pixel, d'une valeur de possibilité d'appartenance à la classe thématique associée à la carte, suite à un protocole de choix d'une clique "élue", dite "clique gagnante". Ce pixel prend en fait, la moyenne des valeurs de possibilités portées par la clique gagnante sur chacune des cartes possibilistes. Le protocole de choix employé, consiste à faire représenter chacune des 9 cliques de Nagao, sur chacune des cartes possibilistes, par sa moyenne. La clique qui permet de mieux séparer les cartes, gagne. On peut dire que ceci correspond à la clique la plus catégorique vis-à-vis de la possibilité d'appartenance, particulièrement, à une classe thématique en dépit des autres.

Cependant résumer chacune des 9 informations spatiales correspondant aux 9 cliques par la valeur moyenne de ses mesures possibilistes, neutralise la diversité au sein de la clique et ne donne pas nécessairement la meilleure opinion à son sujet. Par exemple, considérons un tel ensemble: $V = [2,3,149,1,178,6,4]$, c'est-à-dire les valeurs de niveau de gris d'une région de 7 pixels. V doit être classé dans deux classes: C1 comme classe de régions «sombres» et C2 comme classe de régions «brillantes». Nous pouvons voir que V est partagé en deux sous-ensembles $V1 = [2,3,1,6,4]$ et $V2 = [149,178]$. Néanmoins, sa valeur moyenne, moyenne (V) égale à 49, ne peut être considérée comme appartenant ni à $V1$ ni à $V2$, donc, elle ne peut pas être un bon représentant de l'ensemble V . Une telle situation, montrant une ambiguïté dans le contenu informationnel est mieux gérée au moyen d'une modélisation possibiliste.

Conscients qu'une telle situation peut se produire au sein d'une clique, nous proposons de préserver la richesse des informations portées par les cliques, pour permettre à tous leurs pixels de contribuer à la décision d'affecter le pixel discuté à une classe spécifique. En effet, au lieu de présenter une clique par sa valeur moyenne, nous proposons de la représenter par la distribution possibiliste de ses pixels.

a) Description du processus

La matrice possibiliste consistant en 9 lignes (en une rangée pour chaque clique) et M colonnes (en une colonne pour chaque classe donnée), illustrée par la Figure 4.12, n'associe plus pour chaque clique fn , $n \in \{1, \dots, 9\}$, la valeur moyenne de ses mesures de possibilités d'appartenance du pixel à une classe, mais la distribution de possibilités conditionnelle à une classe, estimée par la considération des niveaux de gris des pixels de la clique fn , comme montré dans la Figure 4.14.

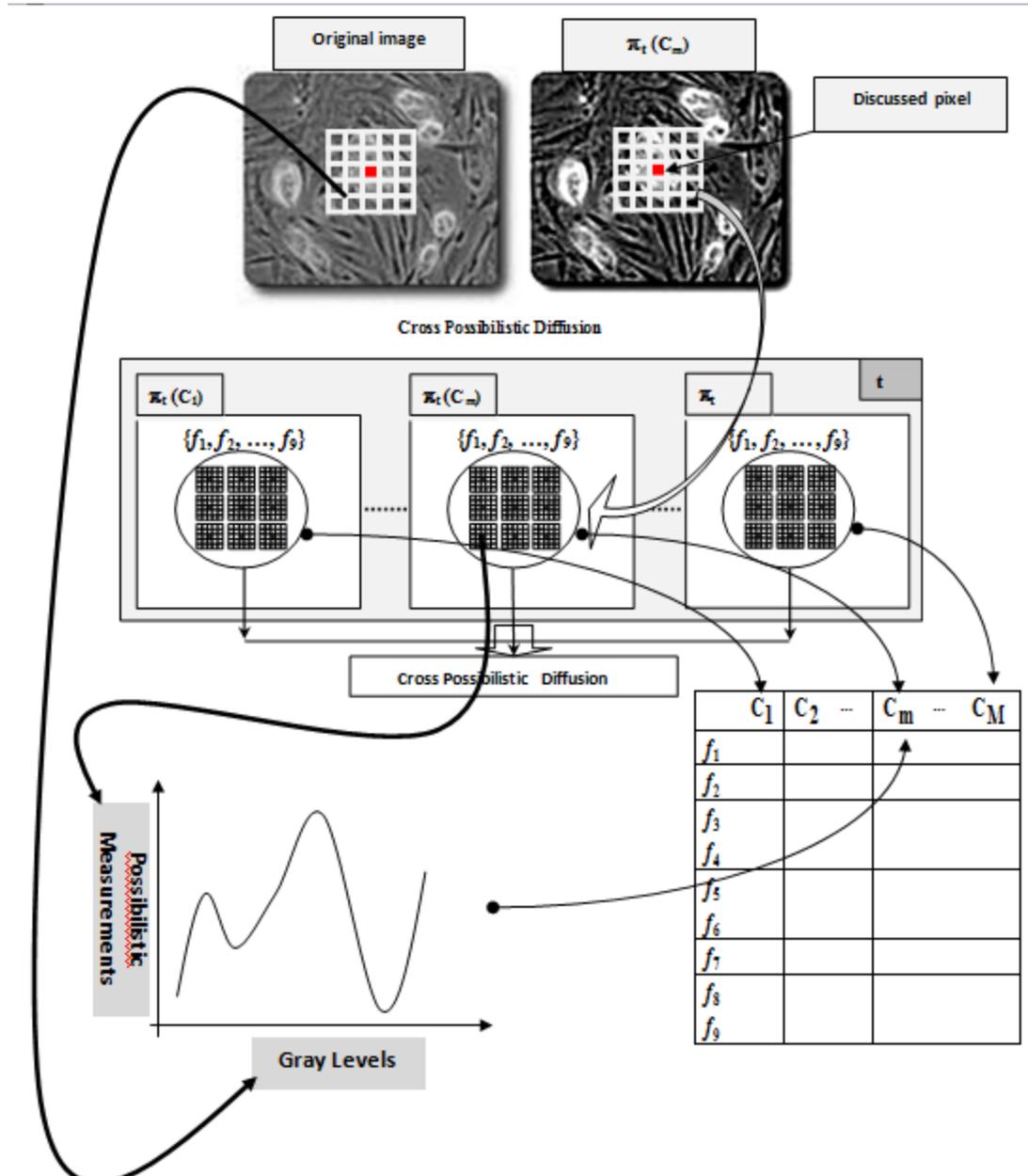


Figure 4.14 Stratégie de la diffusion possibiliste croisée où le contenu informationnel des cliques de Nagao est représenté par une distribution possibiliste des NG de ses pixels.

L'idée est de conserver la richesse de l'information portée par les cliques, de façon à permettre à tous les pixels de la clique de s'exprimer quant à la possibilité d'appartenance à une classe de leur pixel central. Le protocole de choix de la clique gagnante se fait sur deux étapes:

1/ localisation des cliques présentant les plus proches distributions de possibilités sur leurs pixels, à celle considérée initialement sur les niveaux de gris, et ce pour chacune des classes thématiques. Les cliques ainsi localisées sont dites "cliques élues"

2/ Sur l'ensemble des "cliques élues", une seule sera considérée gagnante, c'est celle vérifiant le critère de choix. Deux propositions sont considérées pour ce dernier:

* La clique gagnante est celle marquant un maximum d'écart entre les cartes de possibilités associées aux classes, par rapport aux autres cliques élues: donc c'est celle ayant

une opinion catégorique. L'écart sera calculé sur les moyennes, et la règle de décision associée est nommée « maxmoy »

* La clique gagnante est celle ayant un maximum d'homogénéité entre les opinions de ses constituants (pixels), par rapport aux autres cliques élues: donc c'est celle qui fournit une opinion par l'unanimité et consensus. Pour cela la variance qui est calculée au sein de chaque clique, et la clique qui vérifie un minimum de variance interne, gagne. La règle de décision associée est nommée « minvar ».

La clique gagnante fournira sa moyenne au pixel en question sur chaque carte possibiliste.

b) Simulation sur un échantillon

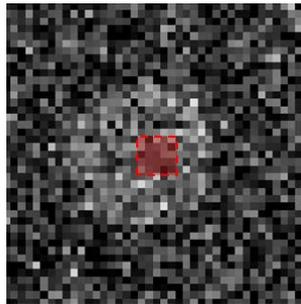


Figure 4.15 Un échantillon de l'image test synthétique

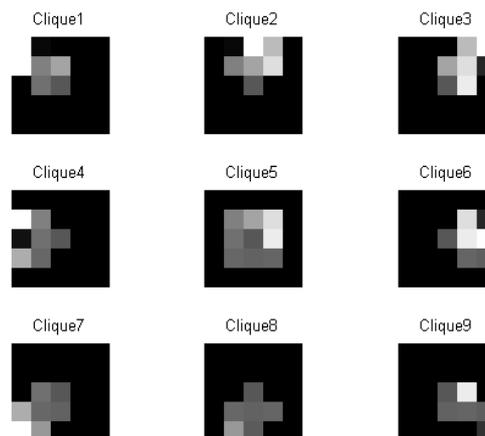


Figure 4.16 Neuf cliques de Nagao, prises sur la carte possibiliste de la classe objet, issue de la première itération

Soit une portion de l'image test synthétique donnée par la Figure 4.15. Le pixel candidat, auquel on désire attribuer une valeur de possibilité d'appartenance pour chacune des classes thématiques, ainsi que son voisinage sont teintés en rouge. Les cliques de Nagao voisines de ce pixel, prises sur la carte possibiliste de la classe objet, issue de la première itération sont données à la Figure 4.16. Cette figure montre que le pixel, bien qu'il appartienne réellement à la classe objet, sa valeur de possibilité d'appartenance à sa classe, est faible (apparue par un niveau de gris sombre).

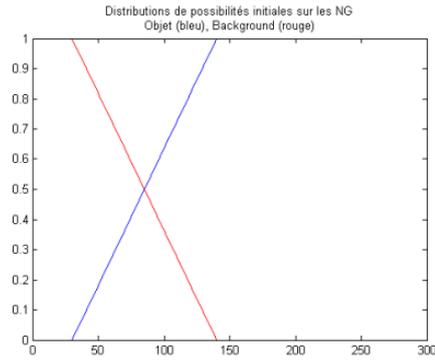


Figure 4.17 Distribution de possibilités considérée initialement sur les niveaux de gris

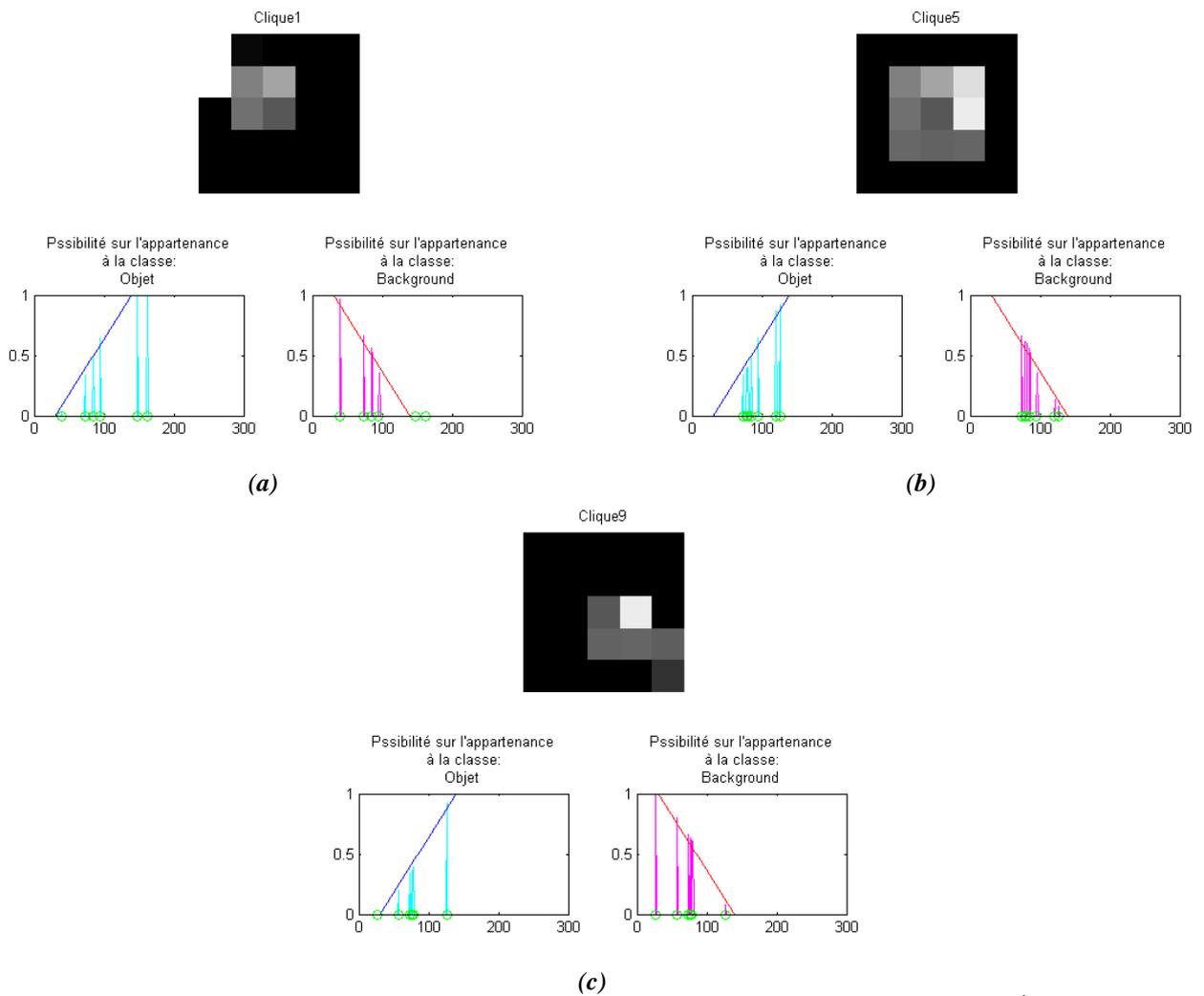


Figure 4.18 Distributions de possibilités sur les cliques (a) 1, (b) 5 et (c) 9, à la 1^{ère} itération

Bleu: distribution de possibilités initiale de la classe objet
 Cyan: distribution de possibilités sur clique de la carte possibiliste objet
 Rouge: distribution de possibilités initiale de la classe fond
 Magenta: distribution de possibilités sur clique de la carte possibiliste fond

Pour chaque clique on mesure le degré de similarité entre la distribution de possibilités fournie par les pixels de la clique et la distribution de possibilités considérée initialement sur les niveaux de gris (Cf. Figure 4.17) et ce conditionnellement à chaque classe thématique. Les distributions des cliques 1, 5 et 9, à la première itération sont illustrées par la Figure 4.18 puis à la 6^{ème} itération par la Figure 4.19. Nous pouvons facilement voir sur la Figure 4.16 que les mesures de possibilités d'appartenance du pixel central à la classe objet, apparues dans les cliques, sont initialement partagées entre fortes et faibles (respectivement vues : claires et sombres). Au bout de la 6^{ème} itération, nous pouvons voir (Cf. Figure 4.19) que les mesures de possibilités sur les cliques, convergent plutôt vers une appartenance à la classe objet.

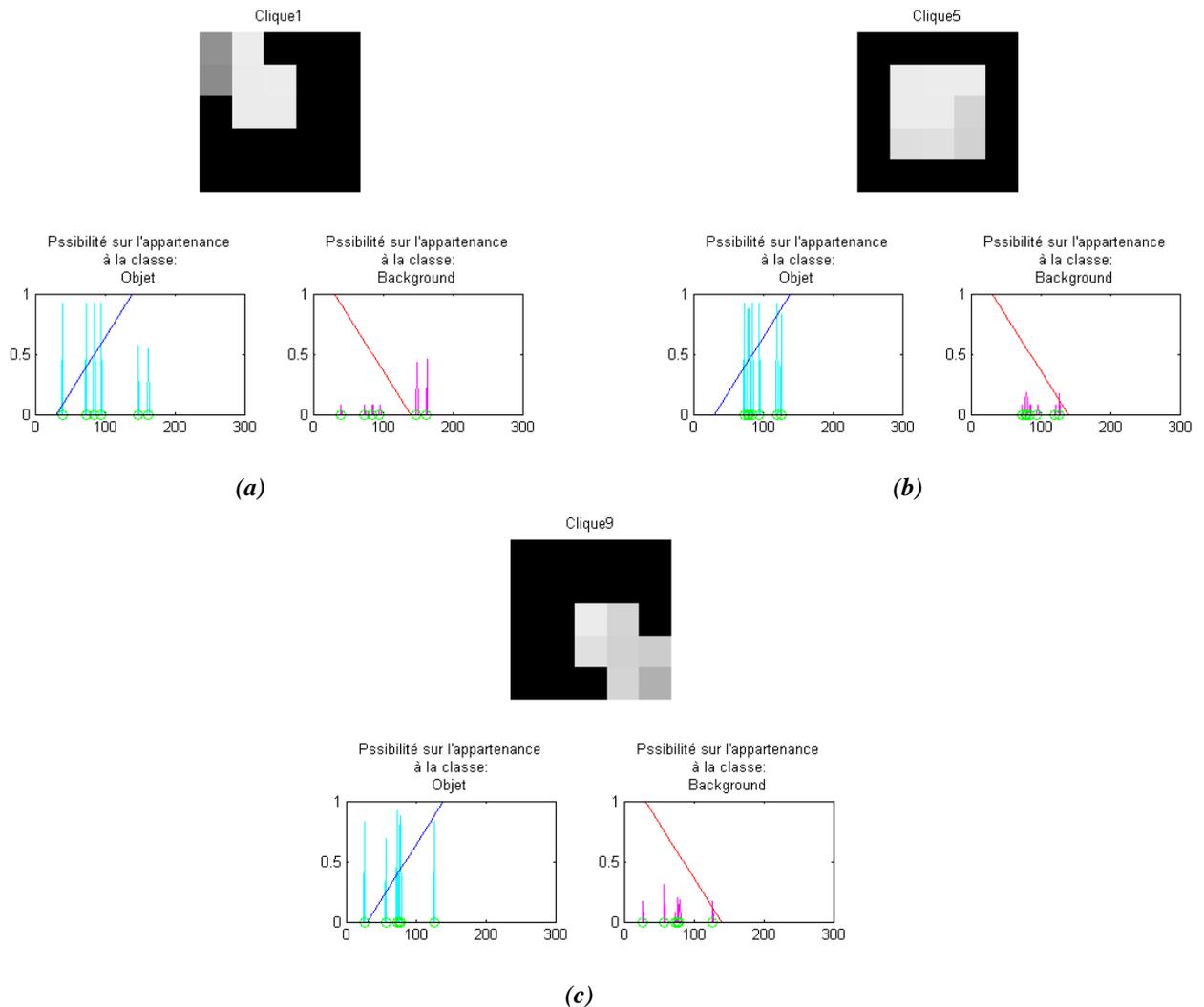


Figure 4.19 Distributions de possibilités sur les cliques (a) 1, (b) 5 et (c) 9, à la 6^{ème} itération

Bleu: distribution de possibilités initiale de la classe objet

Cyan: distribution de possibilités sur clique de la carte possibiliste objet

Rouge: distribution de possibilités initiale de la classe fond

Magenta: distribution de possibilités sur clique de la carte possibiliste fond

6. Résultats et discussions

Dans cette section, nous évaluons les différentes stratégies proposées permettant la diffusion des connaissances possibilistes. Une telle évaluation doit être réalisée d'une manière quantitative et qualitative. Par conséquent, nous reprenons l'image synthétique pour laquelle des indicateurs de qualité sont connus *a priori*. L'image test est illustrée dans la Figure 2.19. Rappelons que cette image est formée d'une imagerie de base comportant quatre disques circulaires représentant un objet d'intérêt de différentes tailles, dupliquée selon cinq niveaux de contrastes différents sur l'horizontale et selon six niveaux de bruit gaussien ajouté sur la verticale.

Notons qu'il s'agit d'une image de synthèse qui représente un contexte, très similaire à celui qui nous concerne dans cette étude (i.e. détection d'un objet d'intérêt) et qui comporte trois indicateurs de qualité : taille d'objet d'intérêt, niveau de contraste et niveau de bruit.

Le plan d'expérimentation que nous proposons consiste, dans un premier temps, à réaliser l'estimation initiale des cartes possibilistes représentant les possibilités, au niveau de chaque pixel, d'appartenir à l'objet d'intérêt ainsi qu'au fond de l'image. Dans un deuxième temps, les différentes stratégies de diffusion itérative des connaissances possibilistes seront évaluées.

a) Estimation initiale des cartes possibilistes

La première étape de notre expérimentation consiste à réaliser l'estimation initiale des cartes possibilistes.

L'approche que nous proposons pour l'estimation des fonctions d'appartenance consiste à considérer un modèle trapézoïdal de ces fonctions puis à appliquer l'algorithme de recuit simulé afin de déterminer automatiquement les différents paramètres définissant ce modèle. L'entropie floue est considérée comme étant la fonction objective à maximiser par l'algorithme de recuit simulé. Ces fonctions d'appartenance sont vues en tant que distributions de possibilités grâce au postulat de Zadeh.

Les résultats obtenus, en termes de cartes possibilistes initiales, sont donnés dans la Figure 4.20.

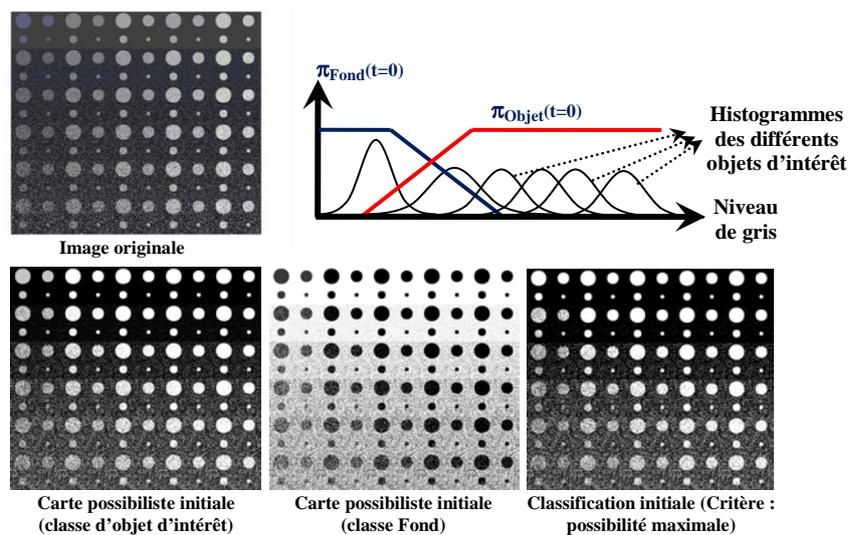


Figure 4.20 Estimation initiale des cartes possibilistes (utilisation des fonctions trapézoïdales) ainsi que les résultats d'une classification initiale

Dans cette même figure, nous avons porté le résultat de la classification pixelique, en deux classes, en utilisant le critère de décision de la possibilité maximale (critère, souvent utilisé pour la classification floue, ou possibiliste, d'images, qui consiste à attribuer chaque pixel à la classe thématique pour laquelle le pixel a la possibilité la plus élevée).

Étant donné qu'il s'agit d'une image de synthèse dans laquelle nous connaissons, *a priori*, la classe d'appartenance de chaque pixel (Objet d'intérêt et fond), nous pouvons calculer l'erreur de classification pour les différents niveaux de contraste et de bruit.

Dans la Figure 4.21, l'erreur de classification « initiale » est donnée pour l'ensemble des 30 imagerie de base correspondant aux cinq niveaux de contraste et aux six niveaux de bruit. Remarquons que pour un fort niveau de bruit et un faible contraste, le taux d'erreur obtenu avec la classification selon le critère de la possibilité maximale atteint des valeurs de l'ordre de 30%. Finalement, le taux moyen d'erreur obtenu avec cette classification initiale est égal à 13%.

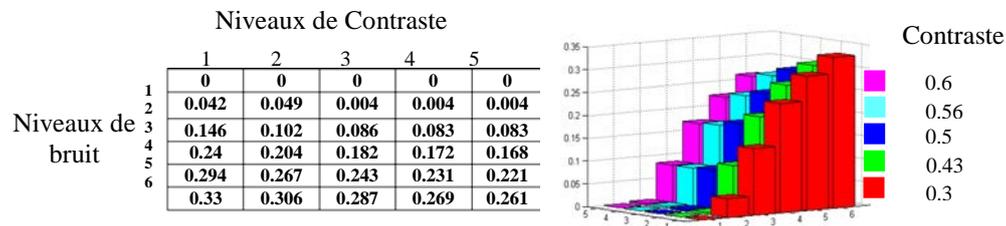


Figure 4.21 Taux d'erreur obtenus par la classification possibiliste initiale en fonction du contraste et du niveau de bruit

Nous remarquons, donc, qu'au niveau pixelique l'approche possibiliste (qui n'est autre que l'approche floue de base) aboutit à des résultats de classification de mauvaise qualité.

b) Diffusion itérative Classe/Classe des connaissances possibilistes

La deuxième étape dans notre expérimentation consiste à considérer les cartes possibilistes initiales, obtenues dans la première étape, et à appliquer d'une façon itérative les différentes approches proposées afin de réaliser la diffusion des connaissances possibilistes du type Classe/Classe. Le premier critère permettant d'apprécier l'intérêt de la propagation des connaissances possibilistes, consiste à observer la « convergence » des connaissances possibilistes issues de l'application itérative de l'opération de diffusion. Cette convergence est évaluée en calculant l'écart quadratique moyen entre les cartes possibilistes « successives » obtenues pour chaque classe. Dans notre expérimentation, nous nous concentrons sur la classe d'intérêt considérée qui est supposée représenter les objets que l'on cherche à détecter dans l'image.

La Figure 4.22, montre l'écart quadratique moyen entre les cartes possibilistes successives représentant la seule classe d'intérêt. Dans cette figure, les trois approches suivantes sont considérées :

- Propagation par filtrage moyen ;
- Propagation par filtrage de Nagao ;
- Propagation par filtrage de Nagao modifié;

Le premier résultat constaté concerne la convergence relativement rapide de toutes ces approches. En effet, nous remarquons que toutes les approches ont atteint un écart quadratique moyen proche de zéro (et ceci pour la classe d'intérêt et pour la classe fond d'image) en moins de cinq itérations.

La Figure 4.23, montre les cartes possibilistes (à la convergence) des différentes approches étudiées, les résultats de la classification possibiliste (basée sur le critère de possibilité maximale), et les taux d'erreur ainsi obtenus.

Comme nous pouvons le remarquer, toutes les approches considérées convergent aussi en termes de taux d'erreur dès la quatrième itération de diffusion. De plus, notons que les erreurs obtenues sont essentiellement des erreurs de « non détection » de pixels appartenant aux objets d'intérêt. Néanmoins, toutes les approches ont permis de détecter, même partiellement, les objets d'intérêt de faibles tailles et ceci pour les différents niveaux de contraste et de bruit.

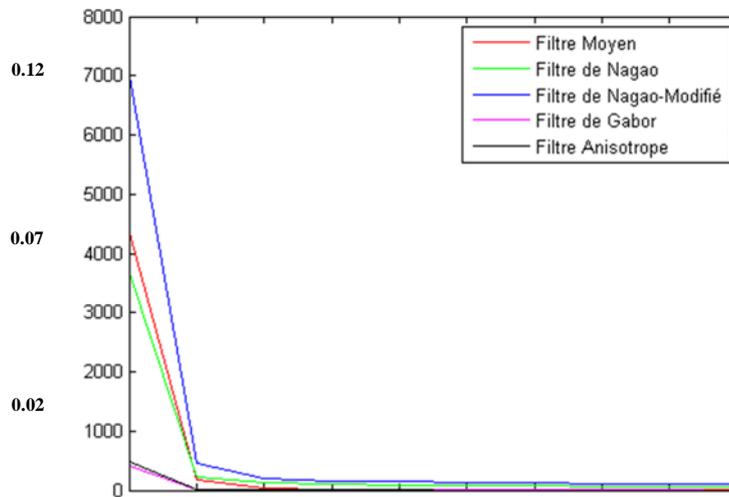


Figure 4.22 Ecart Quadratique Moyen entre les cartes possibilistes successives, pour la classe d'intérêt

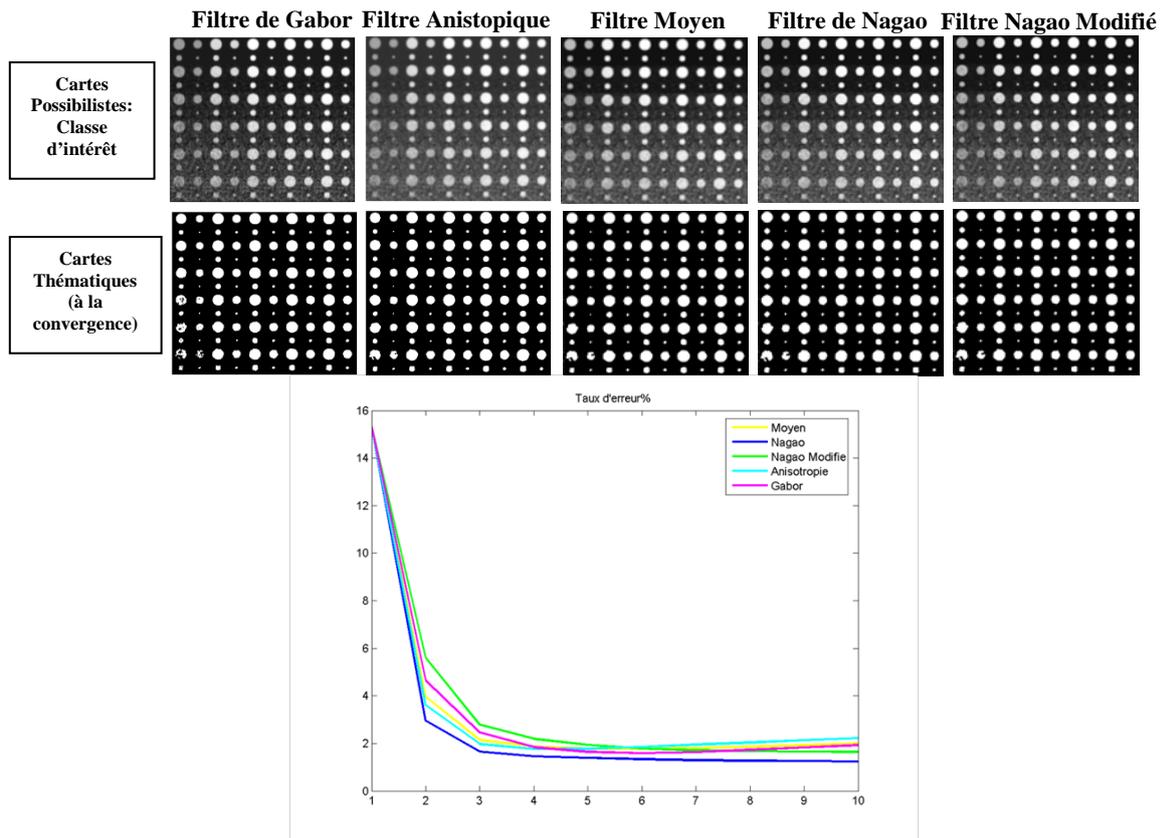


Figure 4.23 Résultats obtenus par l'application des différentes approches de diffusion possibiliste des connaissances : Cartes possibilistes finales, Carte thématiques et taux d'erreur (en fonction du nombre d'itération)

La Figure 4.24, montre les taux d'erreurs obtenus en fonction des différents niveaux de contraste et de bruit, et ceci pour les différentes approches considérées.

Nous remarquons que le filtre de Nagao est celui qui produit les taux d'erreur les plus faibles (en considérant la classification possibiliste maximale) et ceci quel que soit l'itération considérée (Cf. Figure 4.23) et pratiquement pour tous les niveaux de contraste et de bruit.

En vertu de ces propriétés très intéressantes ainsi qu'au fait que ce filtre permet, de par sa conception, la préservation des contours d'objets, nous allons considérer, dans la suite de ce travail, le filtre de Nagao comme support de diffusion des connaissances possibilistes représentatif des méthodes de la stratégie de diffusion Classe/Classe.

A ce stade, on peut se poser une question de principe liée à l'espace de diffusion des connaissances. En effet, ne peut-on pas réaliser la diffusion des connaissances dans l'espace des mesures liées aux capteurs (i.e. espace des niveaux de gris) au lieu de l'espace des connaissances possibilistes !

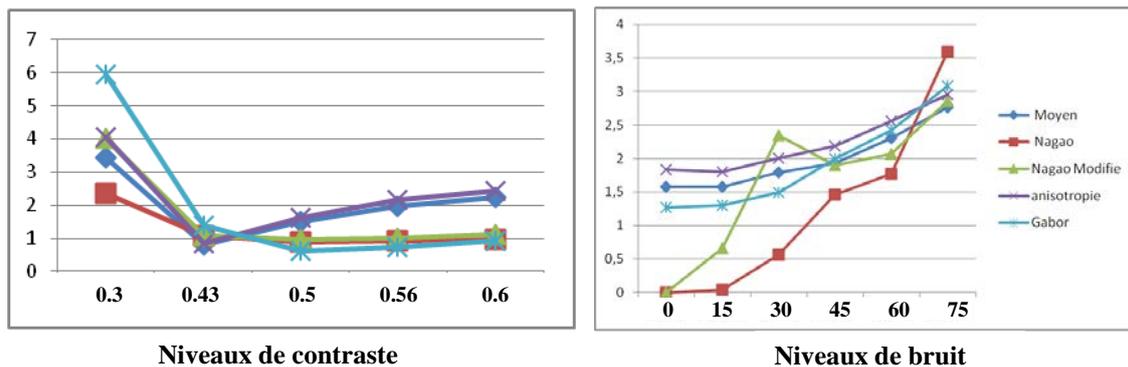


Figure 4.24 Taux d'erreur (%) obtenus en fonction des niveaux de contraste et de bruit

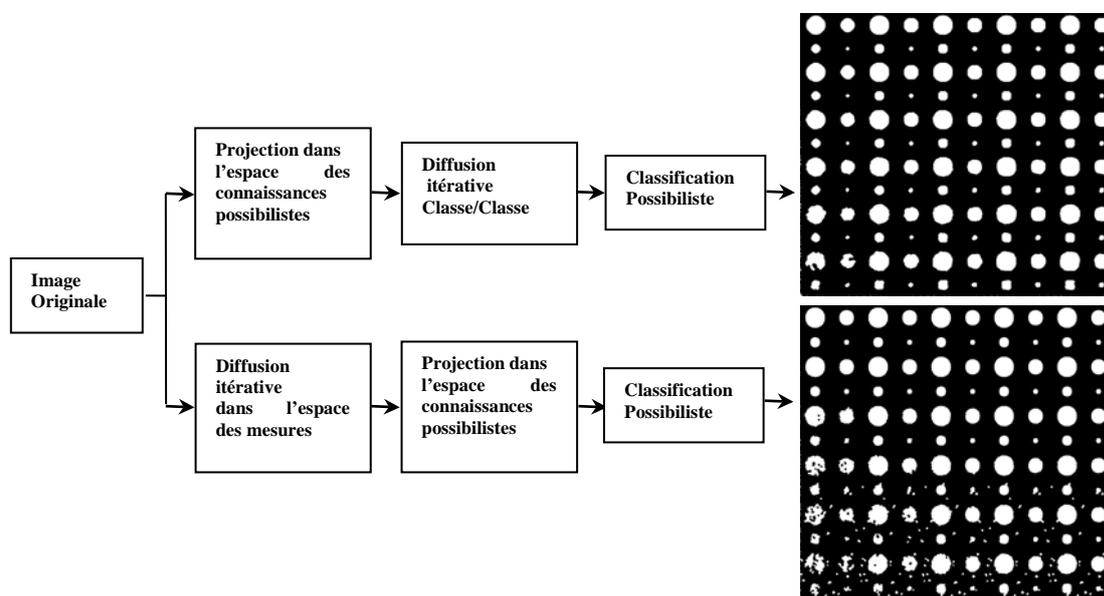


Figure 4.25 Propagation des connaissances dans les espaces des mesures et des connaissances possibilistes (diffusion réalisée par le filtre de Nagao)

Pour répondre à cette interrogation, tout à fait légitime, on peut noter que l'espace des mesures comporte l'ensemble des classes potentielles sans aucune caractérisation particulière, liée à des connaissances externes, de chacune de ces classes. Par conséquent, lorsque l'attention est liée à une classe particulière, cette classe doit être caractérisée via une propriété donnée et ensuite on pourra réaliser la diffusion des connaissances, ce qui revient, en d'autres termes, à réaliser la transformation des connaissances liées aux mesures, en des connaissances possibilistes. Néanmoins, en considérant l'application itérative de l'une des approches considérées dans ce paragraphe sur l'espace des mesures, on peut comparer les performances avec l'approche de diffusion des connaissances possibilistes.

Le schéma d'expérimentation est illustré dans la Figure 4.25. Dans ce schéma, une première branche consiste à réaliser, d'abord, une projection des mesures dans l'espace des connaissances possibilistes et d'appliquer, ensuite, une diffusion itérative des connaissances dans l'espace des possibilités. Dans la deuxième branche, nous proposons d'appliquer, d'abord, la diffusion des connaissances liées aux mesures, et de réaliser la projection des

résultats sur l'espace des connaissances possibilistes. En finale, une classification possibiliste est appliquée en considérant le critère de la possibilité maximale.

En considérant l'approche de diffusion par l'application itérative du filtre de Nagao, nous constatons une très forte dégradation des performances des résultats (en termes de taux d'erreur et de la qualité « visuelle » des résultats obtenus). On constate, aussi, ce même résultat en considérant les autres approches de diffusion des connaissances (i.e. le filtre moyen, filtre de Nagao modifié, filtre de Gabor et filtre anisotropique). Sans que ça soit une preuve théorique, ceci montre l'intérêt pour réaliser la diffusion des connaissances dans l'espace des connaissances possibilistes par rapport à une diffusion des connaissances liées aux mesures brutes en sortie des capteurs.

c) Diffusion itérative Croisée des connaissances possibilistes

Dans ce paragraphe, nous étudions l'approche permettant la propagation des connaissances possibilistes via une diffusion itérative croisée. Etant donné que nous avons retenu l'approche de Nagao comme la plus représentative des méthodes de diffusion Classe/Classe, nous proposons de présenter les résultats d'une façon comparative.

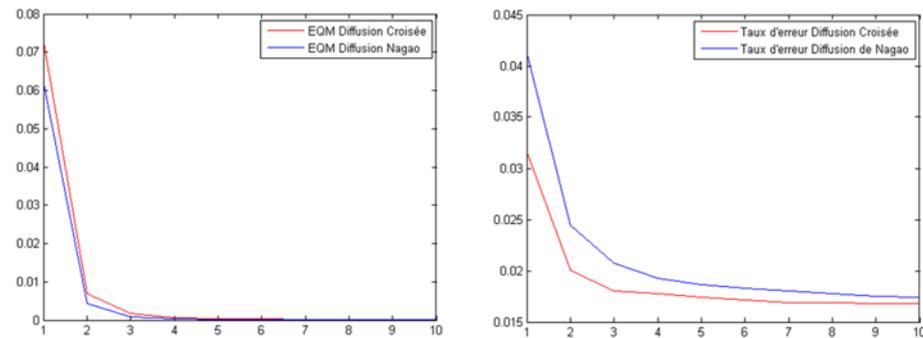


Figure 4.26 Ecart Quadratique Moyen (entre cartes possibilistes successives) et taux d'erreur (classification basée sur la possibilité maximale) des deux modes de propagation des connaissances : Diffusion Croisée et Classe/Classe de Nagao

Dans la Figure 4.26, nous présentons la courbe d'évolution de l'Ecart Quadratique Moyen (EQM) entre les cartes possibilistes successives liées à la classe d'intérêt ainsi que la courbe représentant les taux d'erreur obtenus en considérant le critère de décision de la possibilité maximale. Ces deux courbes sont tracées en fonction du nombre d'itération.

On remarque que les deux approches convergent, en moins de cinq itérations, et ceci en termes d'EQM et taux d'erreur. Les deux premières itérations réalisent la transition la plus importante et les itérations suivantes ne font que « raffiner la convergence ». Notons aussi que les taux d'erreur des deux approches ont un comportement asymptotique (i.e. il existe une barrière basse des taux d'erreur, de l'ordre de 1.7%).

La Figure 4.27, montre les cartes possibilistes et les résultats de classification obtenus à l'issue de la convergence.

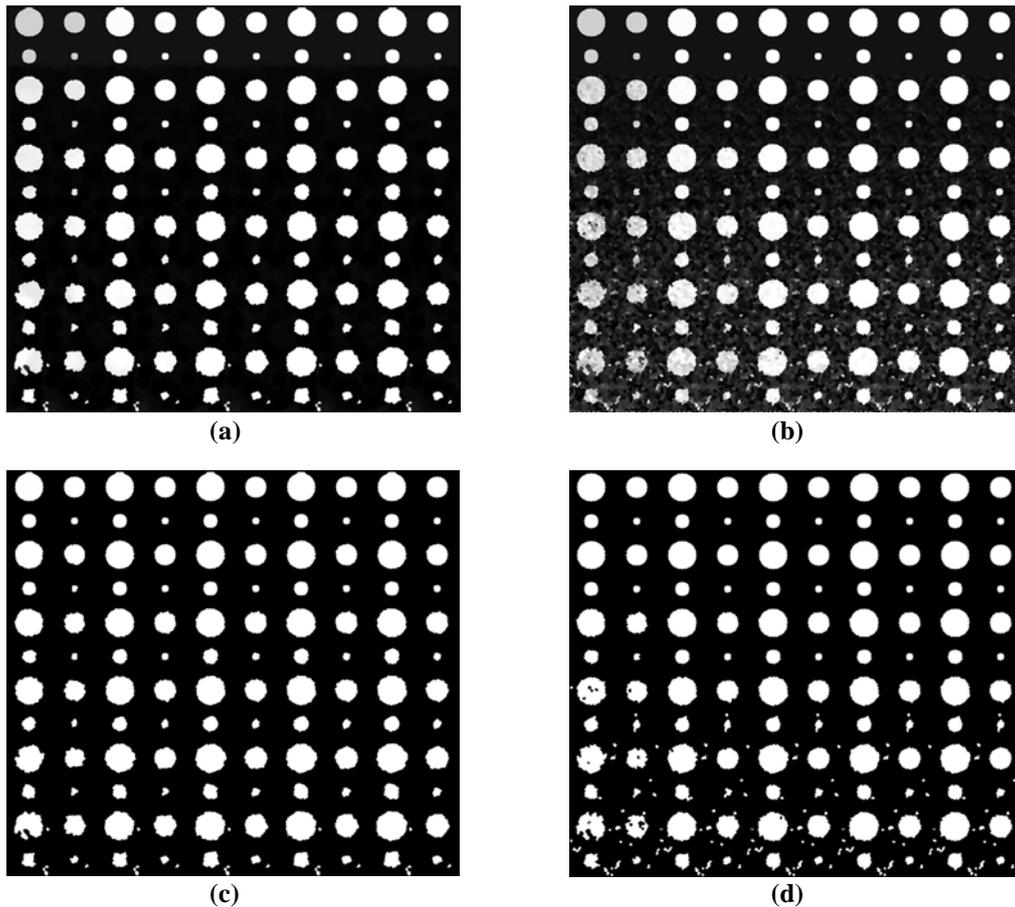


Figure 4.27 Résultats des deux approches de propagation des connaissances possibilistes ;

Carte possibiliste finale (classe d'intérêt) :

(a) Diffusion Croisée

(b) Diffusion Classe/Classe par filtre de Nagao

Résultats de classification (critère de possibilité maximale) à partir des:

(c) Cartes possibilistes obtenues par la diffusion croisée

(d) Cartes possibilistes obtenues par la diffusion Classe/Classe par filtre de Nagao

L'analyse de ces résultats montre que :

- La carte possibiliste obtenue par la diffusion croisée (Figure 4.27-a) est pratiquement une image binaire avec deux niveaux de possibilités d'appartenance, tandis que la carte issue de la diffusion Classe/Classe de Nagao (Figure 4.27-b) comporte plusieurs niveaux de possibilités ;

- La carte possibiliste issue de la diffusion Classe/Classe de Nagao (Figure 4.27-b) présente le même effet observé suite à l'application du filtre de Nagao sur des images numériques : obtention de petites régions mosaïquées et ceci d'une manière croissante en fonction du niveau de bruit ;

- Les zones d'images comportant l'objet d'intérêt sont « bien localisées » par les deux approches de diffusion de connaissances possibilistes (Figure 4.27-c-d);

Notons aussi que les erreurs de classification constatées ne concernent pas les mêmes types de zones.

En effet, la majorité des pixels mal classifiés (en utilisant les cartes possibilistes obtenues par la diffusion croisée, Figure 4.27-c) sont souvent situés sur les zones de contours, tandis que les pixels erronés (en utilisant les cartes possibilistes obtenues par la diffusion Classe/Classe de Nagao, Figure 4.27-d) sont de plus en plus situés dans les zones fortement bruitées.

Cette constatation correspond aux deux principes de base des deux méthodes utilisées, à savoir:

- ✓ le filtre de Nagao vise à préserver les contours, tandis que
- ✓ l'approche croisée met en avant l'homogénéité décisionnelle des cliques considérées.

d) Effet de l'initialisation des connaissances possibilistes

Un aspect important de l'approche de diffusion considérée concerne le choix des distributions des possibilités initiales correspondantes aux différentes classes thématiques. En effet, la question que l'on peut se poser concerne deux aspects : un choix non optimal des distributions initiales et le choix de la forme de base utilisée pour ces distributions.

Rappelons que les expérimentations précédentes ont considéré des distributions initiales du type trapézoïdal avec une optimisation de choix des paramètres liés à ces formes par l'application de l'algorithme de recuit simulé sur l'entropie floue comme fonction objective à maximiser.

Dans un premier temps, analysons l'effet du choix de la forme des distributions possibilistes initiales en adoptant des distributions du type S et en appliquant la même approche d'optimisation du choix des paramètres liés à ces formes.

La Figure 4.28 résume les résultats obtenus en utilisant des distributions initiales du type S et en appliquant, ensuite, les deux approches de propagation des connaissances possibilistes : par diffusion Classe/Classe de Nagao et par diffusion croisée.

Nous constatons, d'après l'analyse de ces résultats, que des performances quasi identiques sont observées, après la convergence, en termes : de vitesse de convergence, de détection d'objets d'intérêt de faibles tailles, de qualité des cartes possibilistes ainsi que des cartes de classification et, finalement en termes de taux d'erreur obtenus.

Ceci montre que pour un choix optimisé des paramètres de forme, les deux types de distributions initiales trapézoïdal et S produisent des résultats de mêmes qualités.

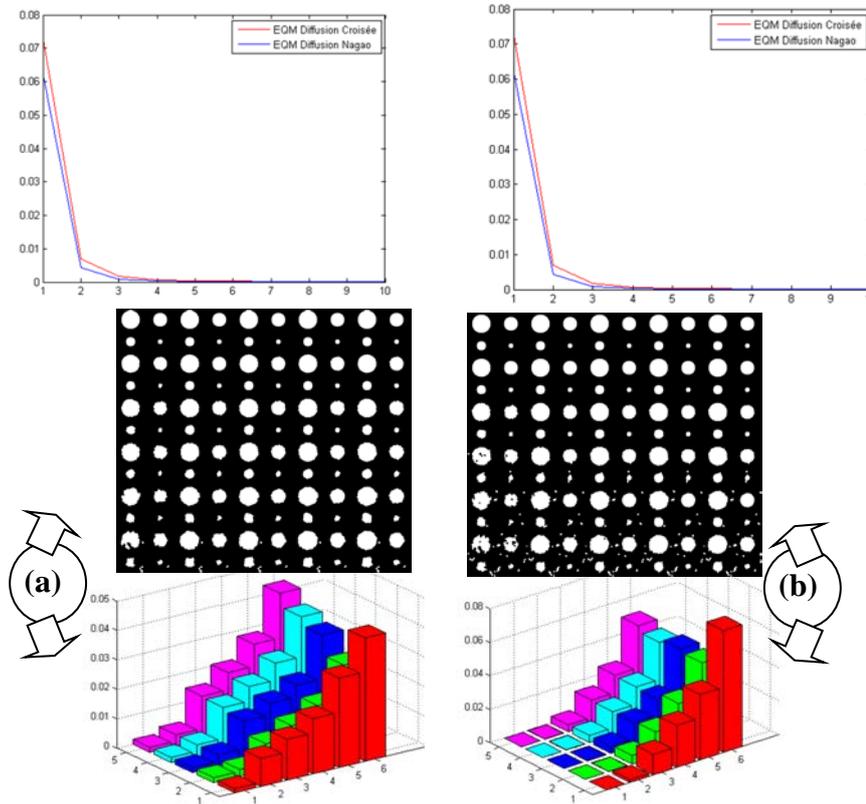


Figure 4.28 Résultats obtenus en utilisant des distributions initiales du type *S*: EQM, Taux d'erreur global, Classification basée sur la possibilité maximale et taux d'erreur (en fonction des niveaux de contraste et de bruit). (a) Propagation par diffusion croisée et (b) Propagation par diffusion Classe/Classe de Nagao

Analysons, maintenant, la situation d'un choix non optimal des paramètres de forme. Rappelons que l'image test considérée comporte un fond d'image du type *sombre* et des objets d'intérêt ayant la forme de disques de différentes tailles et de différents niveaux de gris (afin d'étudier l'effet du contraste) mais tous les disques sont considérés plutôt *brillants* par rapport au fond de l'image. Dans ces conditions, les deux distributions de possibilités de base correspondant aux deux propriétés « ambiguës » de base : *Sombre* et *Brillant* (données dans la Figure 4.29-a constituent les deux distributions les « plus extrêmes » pouvant être utilisées pour représenter les deux classes considérées.

Les résultats de l'application des deux approches de diffusion de connaissances possibilistes : diffusion croisée, Figure 4.29-b et Classe/Classe de Nagao, Figure 4.29-c sont donnés dans cette figure en termes de résultats de classification possibiliste maximale (après convergence).

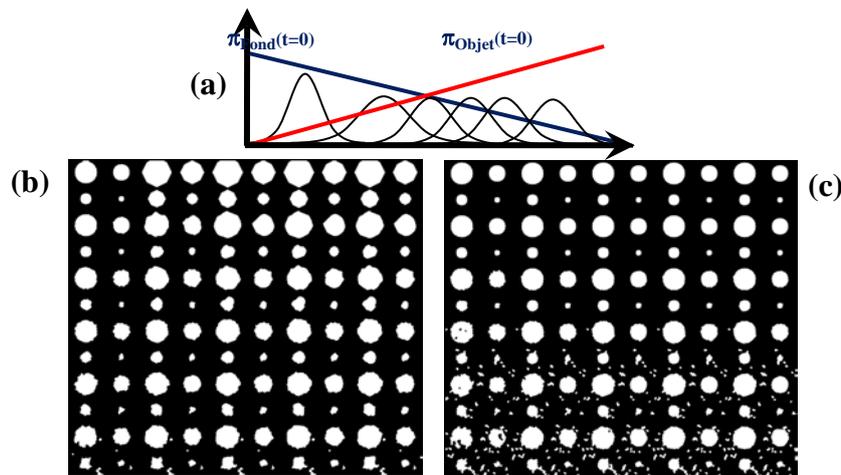


Figure 4.29 Initialisation non optimale

(a) des distributions possibilistes initiales et résultats obtenus en termes de classification basée sur la possibilité maximale

(b) Propagation par diffusion croisée

(c) Propagation par diffusion Classe/Classe de Nagao

On remarque une certaine « dégradation » des résultats tout en maintenant les résultats « majeurs » essentiellement obtenus lors d'un choix optimal des paramètres, à savoir : la détection des objets d'intérêt de faibles tailles (dans les deux modes de propagation), des erreurs de détection dans les zones des contours des objets détectés (dans le cas de la diffusion croisée) et dans des zones de fond de l'image (dans le cas de la diffusion Classe/Classe de Nagao).

A ce stade, il est intéressant d'analyser les courbes de convergence en termes d'écart quadratique moyen (entre cartes de possibilités successives concernant la classe d'intérêt) et les courbes de taux d'erreur obtenues en appliquant le critère de classification possibiliste maximale, Figure 4.30.

En effet, notons qu'après une forte diminution obtenue dès la deuxième itération, la courbe de l'EQM continue sa convergence asymptotique avec une diminution très faible. En revanche, les courbes représentant les taux d'erreur, montrent un comportement particulièrement intéressant à analyser. En effet, le taux d'erreur obtenu via la diffusion des connaissances possibiliste du type Classe/Classe de Nagao montre la même évolution descendante et asymptotique que celle obtenue avec une initialisation optimisée des distributions possibilistes initiales. En revanche, en utilisant l'approche de diffusion croisée des connaissances possibilistes, ce taux d'erreur montre un début de décroissance avec les deux premières itérations et un comportement de « décrochage » dès la troisième itération. Pourtant, l'analyse de la carte thématique (objet d'intérêt) obtenue après convergence ne reflète pas ce décrochage d'une façon visuelle « flagrante ». Néanmoins, on peut constater que les erreurs de classification obtenues sont essentiellement localisées dans le voisinage des zones de contours d'objets d'intérêt. L'interprétation de ce phénomène est quasi immédiate. En effet, l'approche de diffusion croisée réalise l'essentielle de « l'harmonisation » des mesures de possibilité dès les deux premières itérations. En revanche, un choix non optimisé des distributions de possibilités initiales rend, dans les zones voisines des contours, l'écart des

mesures possibilistes entre la classe d'intérêt et le fond de l'image très faible. L'objectif de la diffusion croisée étant d'harmoniser les mesures de possibilité dans les régions homogènes sans se soucier des zones de contours, alors, cette approche aura tendance à faire croître les zones de régions (correspondant à la classe d'intérêt) au détriment des zones de contours et des zones du fond de l'images. Les faibles erreurs observées dans les zones du type fond d'images confirment cette interprétation.

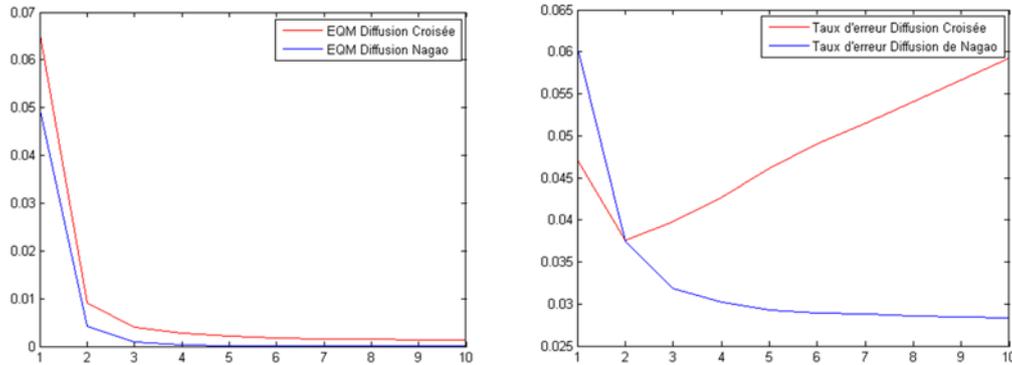


Figure 4.30 Evolution de l'Ecart Quadratique Moyen (entre cartes possibilistes successives) et le taux d'erreur en fonction du nombre d'itérations

e) Utilisation de la diffusion croisée pour la segmentation d'images mammographiques :

L'évaluation des performances de l'approche proposée utilisant la stratégie de diffusion possibiliste classe/classe de Nagao, est effectuée sur des images mammographiques extraites de la base de données d'images MIAS: «Mammographic Image Analysis Society» [Suc 1994]. Ces images présentent deux classes: tumeur et tissus sains (Cf. Figure 4.31). Il convient de noter le faible contraste entre ces deux classes, ainsi que l'importante difficulté «visuelle» pour un radiologue, sur certaines images, pour délimiter les frontières séparant les objets d'intérêt, à savoir les tumeurs.

L'approche de segmentation, basée sur la diffusion possibiliste classe/classe de Nagao proposée, est comparée aux méthodes les plus largement utilisées pour la segmentation d'images mammographiques: la méthode Level set [SLS 2002, DHG 2010], la méthode de Fuzzy C-Mean [SCL 2017, BP 2009, HAN 2004] et la méthode de croissance de région [MRS 2008, ZFK 2004].

Une évaluation quantitative et qualitative des quatre approches, a été menée et analysée dans [KAB 2018], se basant sur des vérités terrain fournies de la part d'un radiologue du CHU de Sfax. La supériorité de la méthode de diffusion des connaissances possibilistes proposée de Nagao, pour la segmentation des images mammographiques, est confirmée, comparée aux trois autres méthodes de segmentation, connues comme les plus utilisées pour les images mammographiques. La Figure 4.31 présente les contours obtenus par les quatre approches sur cinq images préalablement détournées par le radiologue. La Figure 4.32 illustre l'un des tests quantitatifs effectués.

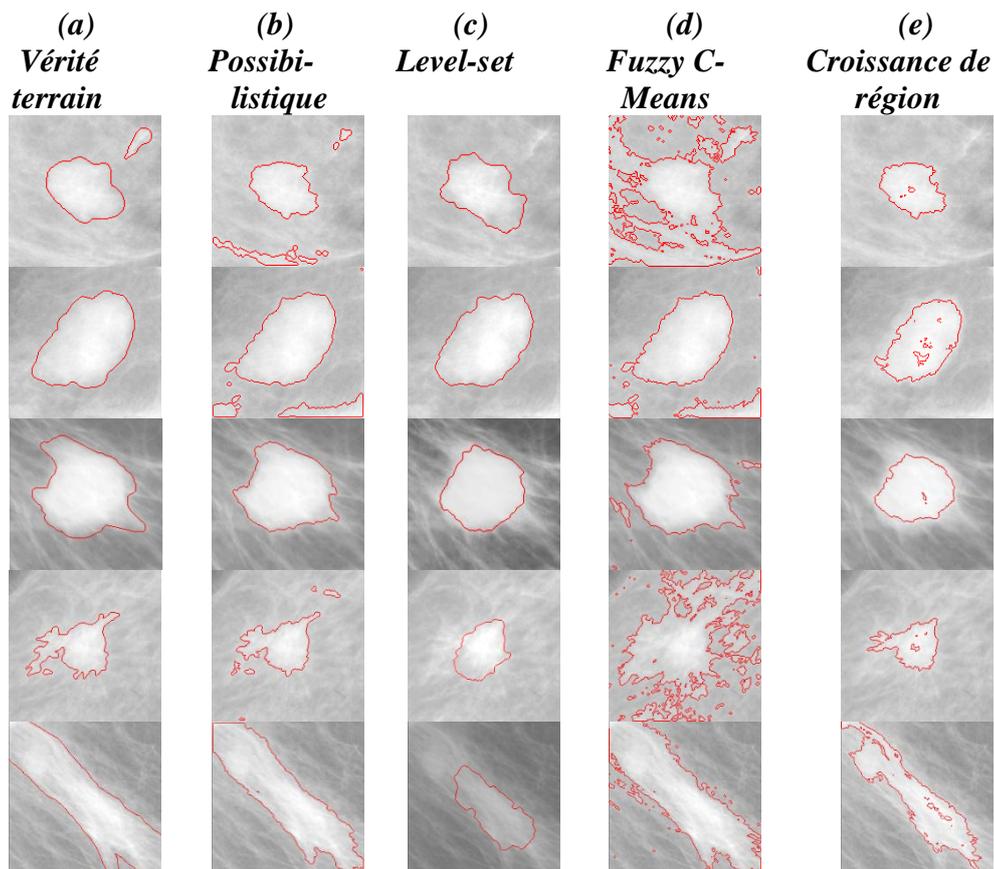


Figure 4.31 Contours des zones suspectes détectées par: (a) le radiologue; (b) approche possibiliste basée sur la diffusion de connaissances par le lissage de Nagao; (c) approche Level-set, (d) approche Fuzzy C-Means; et (e) approche de croissance de régions

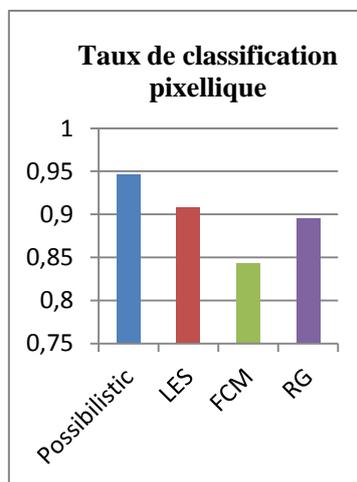


Figure 4.32 Taux de classification pixellique obtenu pour cinq images de la base MIAS pour les quatre approches

f) Effet de la modélisation possibiliste du contenu informationnel de la clique sur la diffusion croisée

La Figure 4.33 illustre les résultats de classification suite à la modification proposée. Les contours de la classification finale, posés sur l'image originale sont donnés à la Figure 4.34. Les taux de reconnaissance, ceux d'erreurs de classification et ceux de rejets sont présentés par la Figure 4.35. Nous donnons l'erreur quadratique moyenne MSE, le rapport signal sur

bruit PSNR et le rapport des normes quadratiques à la Figure 4.36. Ces résultats sont comparés avec ceux de l'approche originale de diffusion croisée que nous notons par la suite "org".

Sur les segmentations finales, nous constatons que les méthodes "minvar" et "maxmoy" donnent visuellement de meilleurs résultats que "org", du fait que les ilots dans le fond qui considérés objet par "org", ont été correctement attribués à la classe fond par "minvar" et "maxmoy". Quelques trous dans les objets ont été également fermés par "maxmoy" (voir celui du disque à la 7ème ligne, 1ère colonne). Nous constatons également que les disques ayant un contraste maximal sans bruit, sont mieux délimités par "maxmoy" et "minvar" que par "org". Cependant, en se confiant aux critères d'évaluation MSE et PSNR, c'est l'approche « org » qui se garde les meilleures performances. Cette observation relève une question importante : « est ce que les critères d'évaluation quantitatifs révèlent une évaluation fidèle à la réalité ? »

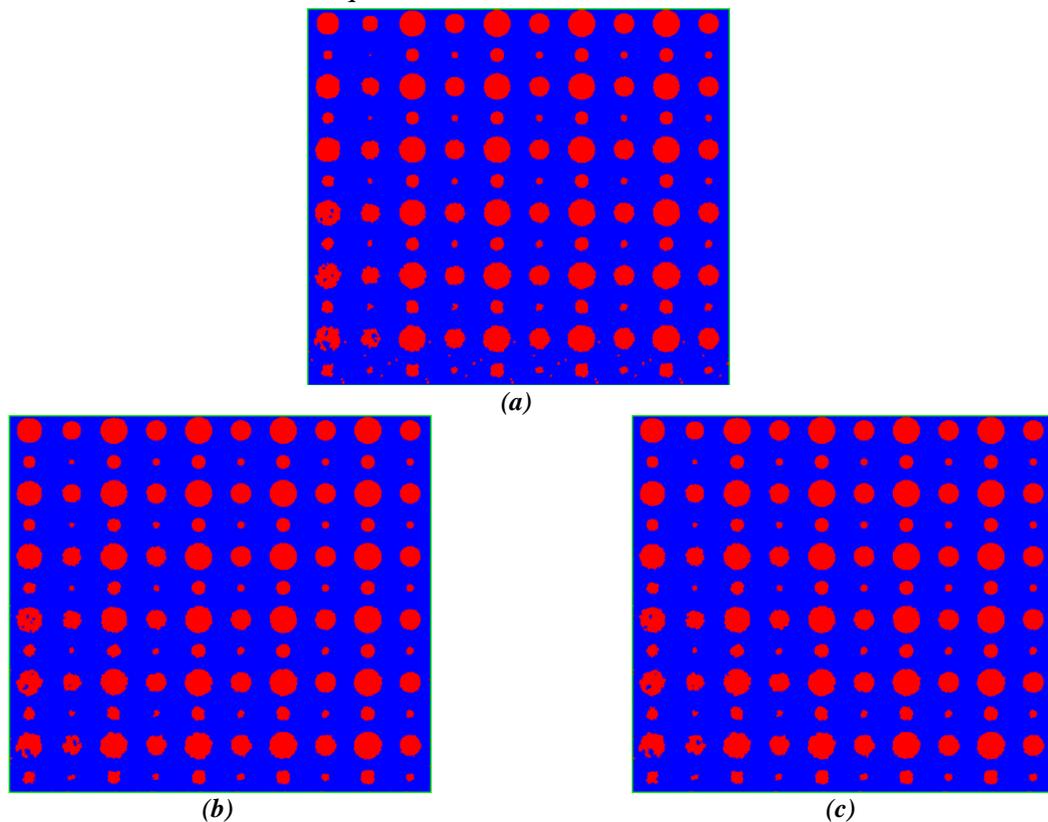
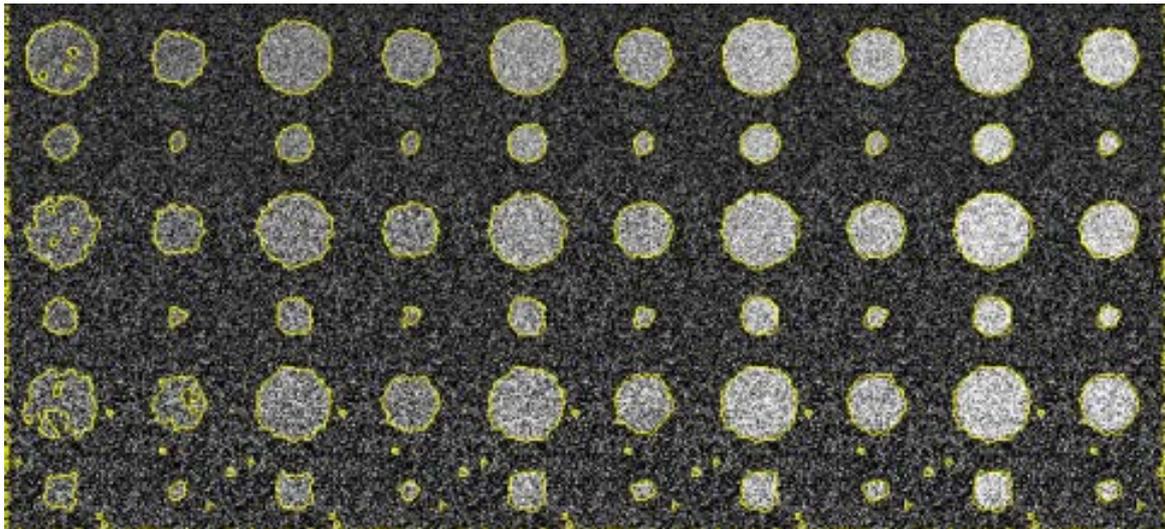
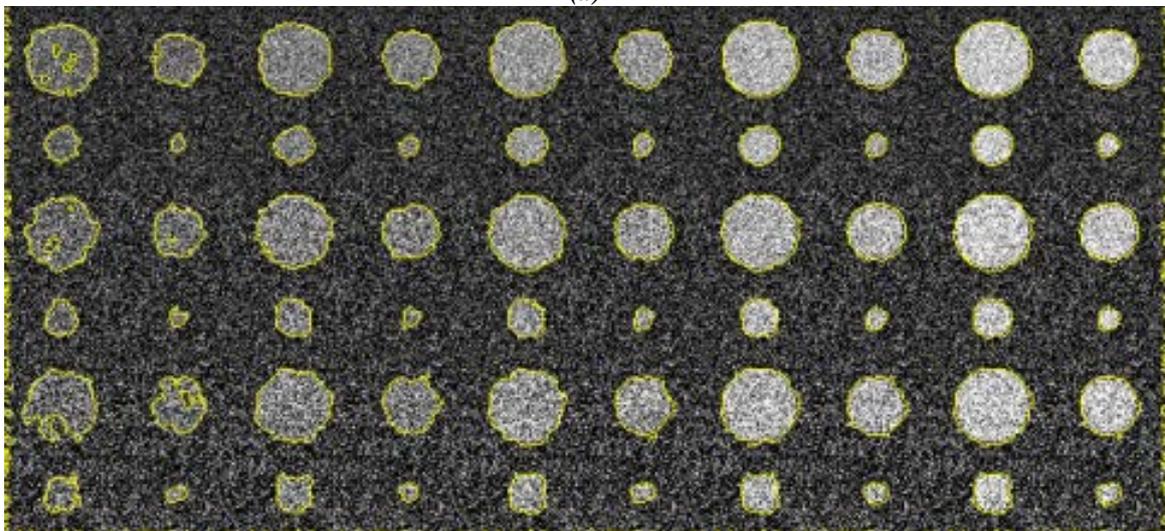


Figure 4.33 Classification finale suite à 30 itérations

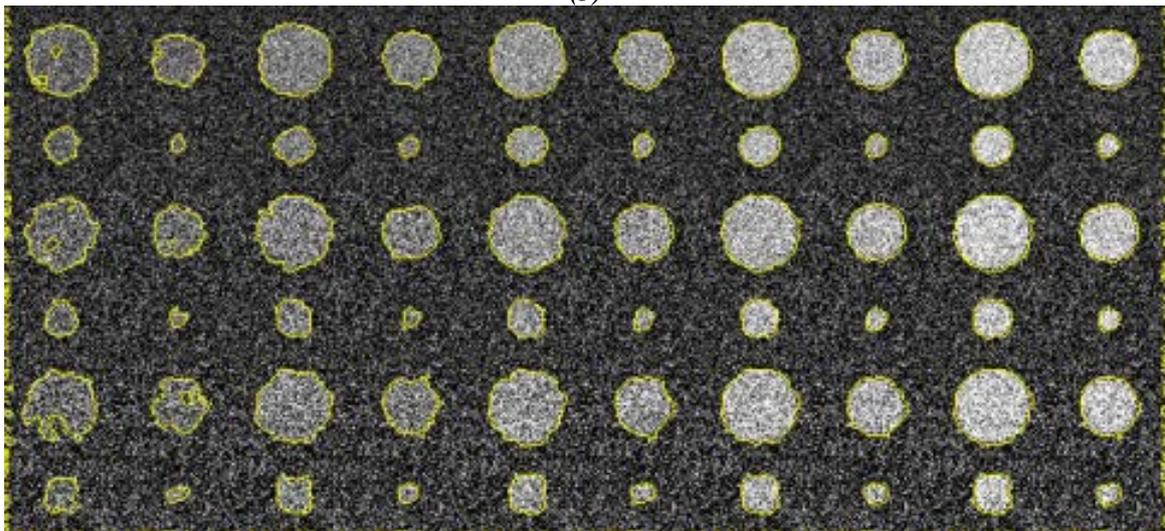
(a) "org", (b) "minvar" et (c) "maxmoy"



(a)

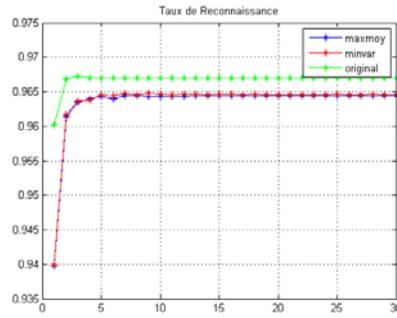


(b)

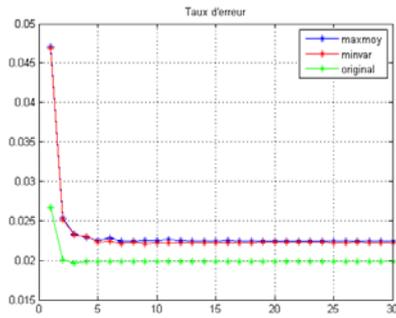


(c)

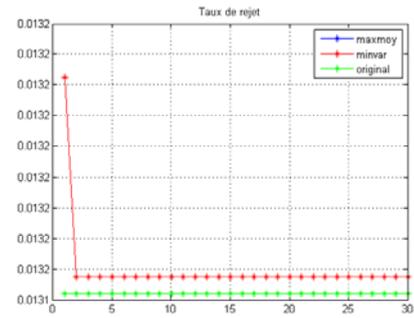
Figure 4.34 Contours de la Classification finale, posés sur l'image originale suite à 30 itérations (zoom sur les lignes de 7 à 12) :(a) "org", (b) "minvar" et (c) "maxmoy"



(a)



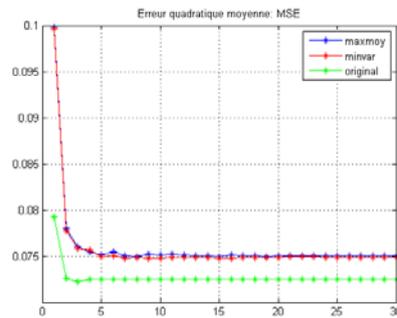
(b)



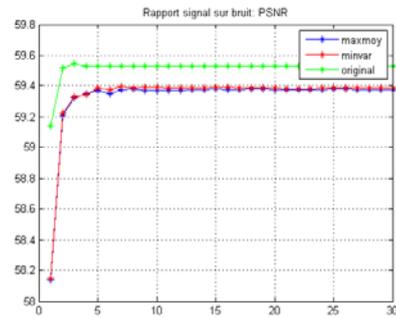
(c)

Figure 4.35 Taux de

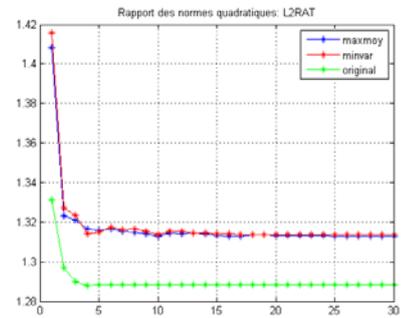
(a) Reconnaissance, (b) Erreur et (c) Rejet



(a)



(b)



(c)

Figure 4.36 Evaluation par:

(a) l'erreur quadratique moyenne MSE,
 (b) le rapport signal sur bruit PSNR et
 (c) le rapport des normes quadratiques

V. Fouille possibiliste d'identité, interprétation de scène biométrique : Notion du carnet d'identité

Il est important de savoir que la fouille de données n'est pas seulement le problème de découverte de modèles dans un ensemble de données pour la prise de décision. En effet, c'est aussi la présentation d'une interprétation de ces modèles, à fin de faciliter leur compréhension et aider la prise de décision. Dans ce cadre nous proposons une approche de fouille possibiliste dans une scène biométrique, permettant, en plus de la reconnaissance de l'identité étant l'objectif classique de la biométrie, la fourniture de renseignements pertinents de l'individu et des modalités biométriques qui lui sont les plus appropriées.

La fouille possibiliste d'identité se base sur des cartes possibilistes biométriques (Cf. Figure 4.37), construites dans le contexte suivant : l'univers Ω est celui des M individus. Chaque individu constitue une classe C^m . Chacune des K modalités biométriques considérées (empreintes digitales, empreintes palmaires, iris, rétine, facial, voix, marche, ...), constitue une source d'information. L'approche de construction des cartes possibilistes biométriques, est celle décrite dans la section V du chapitre 3, à des détails près. Ces détails font toute la différence qui a permis de construire un « carnet d'identité » sur la base de ces cartes biométriques. En effet, dans l'approche originale de la fouille possibiliste, le calcul des poids $\alpha_{Cb_i}^m$ à une itération t , pour la mise à jour de la carte de C^m , se fait sur la base d'un calcul de distances entre les modèles de C^m correspondants aux combinaisons Cb_i ($i = 1 \dots 2^K - 1$) et ceux de la réunion de toutes les classes de l'univers Ω , autres que C^m , correspondants respectivement à ces mêmes combinaisons. Cependant, certains modèles peuvent être efficaces pour séparer C^m uniquement de certaines classes, mais pas de toutes les classes. Spécialement dans le contexte biométrique, l'apport d'un trait biométrique, vis à vis de la reconnaissance, ne présente pas la même efficacité d'un être humain à un autre. Ceci nous invite à construire des cartes de distances entre la carte possibiliste biométrique C^m , et toute carte possibiliste biométrique $C^{m'}$, tel que $m' \neq m$. $M-1$ cartes de distance sont ainsi construites pour la classe C^m , présentant un bilan de ses relations avec toutes les classes de l'univers Ω et selon différents points de vues, exprimés par les différentes combinaisons des sources d'informations considérées. Nous proposons d'appeler ces $M-1$ cartes de distance, le « carnet d'identité » de la classe C^m , (Cf. Figure 4.38).

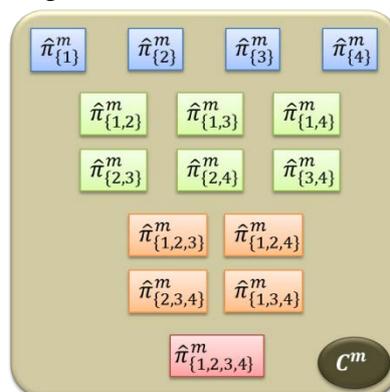


Figure 4.37 Carte possibiliste biométrique de l'individu représenté par la classe C^m , dans le cas de quatre modalités biométriques

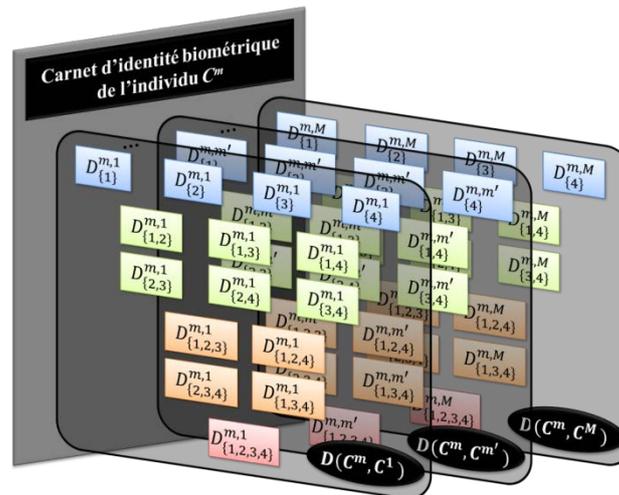


Figure 4.38 Carnet d'identité biométrique de l'individu représenté par la classe C^m dans le cas de quatre modalités biométriques

L'idée de concevoir un carnet d'identité biométrique personnalisé pour chaque individu courant la phase hors ligne d'un processus de reconnaissance biométrique de l'identité, nous paraît d'un grand intérêt. Le système ainsi conçu, agirait d'une façon sur mesure pour l'identification d'un client. La procédure proposée permettrait d'une part une meilleure perception des données biométriques disponibles et d'autre part une économie sur le temps et les moyens. L'économie sur le temps serait due au fait que des traitements inutiles seraient épargnés. L'économie sur les moyens apparaît dans le fait que des traits biométriques considérés comme des faibles sources pour un individu à identifier, ne seraient même pas acquises et nous économisons ainsi sur la durée de vie de l'outil d'acquisition.

VI. Conclusion

Dans ce chapitre, nous avons présenté nos activités de recherche liées à la construction et l'interprétation de scènes.

La construction des scènes a été un choix stratégique dans nos travaux de recherche. En effet, il est toujours convenable d'évaluer les approches conçues sur des bases de données standards, à fin de les comparer à des approches de l'état de l'art testées sur ces mêmes bases, cependant il est plus exaltant de les confronter au monde réel par des données acquises directement sur une scène vécue. La construction de scène est également un défi pris pour concrétiser nos travaux et tester des pistes pour la conception de solutions technologiques spécifiques sur-mesure. Dans ce cadre nous nous sommes aventurés dans la conception de quelques dispositifs, dont trois ont été décrits dans ce chapitre. Les deux premiers sont des dispositifs contribuant à la construction d'une scène Biométrique. Ils assurent l'acquisition de l'iris dans des conditions maîtrisées et conformes aux exigences du comité de biométrie de l'INCITS (International Committee for Information Technology Standards). Le premier iriscope opère dans le spectre visible, tandis que l'autre opère dans le proche infrarouge. Le troisième dispositif est conçu pour un système d'aide à la navigation. Il est équipé par des capteurs ultrason, capteurs de position, et une caméra pour l'acquisition d'une scène environnante. Il est aussi équipé par un système de retour d'information.

Dans l'interprétation de scène, ce chapitre ne se veut pas exhaustif et ne couvre pas toute l'étude effectuée dans cette thématique. Nous avons choisi de présenter deux approches que

nous avons développées pour interpréter une scène médicale pour la première et une scène biométrique pour la deuxième.

L'interprétation de la scène médicale imagée a été assurée par une diffusion itérative de connaissances possibilistes. L'approche a d'abord été proposée dans [Ezi 2012] puis développée dans d'autres travaux [Als 2014, KAB 2018, ASA 2016]. Le principe de cette approche, décrit dans ce chapitre, consiste à utiliser les connaissances *a priori* d'expert, pour construire à partir de l'image, des structures conceptuelles riches en connaissances spécifiques aux différentes classes de l'image. Ces structures permettent une interprétation plus efficace de l'image que lorsqu'on la considère à son état radiométrique. Les connaissances de l'expert sont limitées à la définition de l'ensemble des classes thématiques supposées présentes dans l'image analysée ainsi qu'à la désignation d'une zone germe, représentative pour chaque classe considérée. Ces connaissances ont permis l'estimation des distributions de possibilité des classes. Ces distributions de possibilités sont ensuite utilisées afin de faire une projection de "connaissances" des données radiométriques (c'est-à-dire les niveaux de gris observés) dans les représentations possibilistes des différentes classes. Le résultat de cette projection est fourni, pour chaque classe thématique, sous forme d'une structure conceptuelle richement informative, que nous avons appelée carte de possibilité d'une classe. Un processus itératif de diffusion spatiale des connaissances possibilistes est ensuite appliqué. Deux stratégies de diffusion, ont été proposées dans cette étude : l'une consiste à utiliser une technique possibiliste de lissage d'image au niveau de la carte possibiliste relative à chaque classe de manière indépendante des autres classes, alors que l'autre fait intervenir l'information croisée entre les classes. Les cartes de possibilité des classes, ainsi fournies ont construit un support efficace pour l'interprétation de la scène imagée. La prise de décision qui s'en est suivie est une segmentation de l'image. Après évaluation qualitative et quantitative des performances de segmentation des approches proposées, sur une image synthétique d'abord, puis sur des images réelles, le filtrage de Nagao, étant l'un des cinq techniques de filtrage d'images qui ont été utilisées pour effectuer le processus de diffusion, a été retenu. Une modification au niveau de l'utilisation de ce filtre pour la diffusion croisée, a été proposée. Il s'agit de représenter une clique par la distribution possibiliste de ses pixels, au lieu de la présenter par sa valeur moyenne. Les cliques de toutes les cartes de possibilité des classes thématiques, participent à un « concours » pour fournir la valeur de possibilité au pixel qu'elles représentent. Un protocole de décision a été formulé pour élire la clique gagnante.

Etant un processus itératif, pour étudier la convergence de la diffusion des connaissances, deux critères ont été utilisés : l'erreur quadratique moyenne de l'évolution des cartes possibilistes et le taux d'erreur de la classification par maximum de possibilité. Les résultats de la simulation ont montré que toutes les méthodes de diffusion de connaissances possibilistes présentent un comportement de convergence très rapide (souvent en moins de cinq itérations). Quatre résultats majeurs ont été présentés. Premièrement, il convient de noter la supériorité en termes de performances par rapport au processus de classification basée uniquement sur l'utilisation de distributions de possibilités initiales (qui correspond à l'approche classique de la classification floue). Le deuxième résultat important concerne le fait que la segmentation d'images utilisant la diffusion de connaissances possibiliste dépasse les méthodes de segmentation diffusant les valeurs radiométriques directement. La diffusion croisée de connaissances possibilistes basée sur le filtrage de Nagao a été retenue en tant que

méthode donnant les meilleurs résultats de segmentation. Enfin, nous avons noté que plus on pousse plus loin le raisonnement possibiliste, tel qu'il a été proposé par la représentation possibiliste d'une clique de Nagao, plus la classification fournie est visuellement plus homogène.

Sur la base des résultats obtenus, plusieurs activités de recherche devraient être abordées dans les futurs travaux. Par exemple, il est clair que le critère de décision basé sur le maximum de possibilité ne tient compte ni de l'objectif de la diffusion de connaissances possibilistes ni de la mesure de l'incertitude (pouvant être évaluée au cours de la diffusion). Ce processus décisionnel devrait être raffiné. Il est aussi clair, que les critères d'évaluation usuels ne reflètent pas toujours ce qui est vu directement sur les résultats. Une étude d'évaluation des critères d'évaluation est déjà en cours.

L'approche proposée pour interpréter une scène biométrique, nommée fouille possibiliste d'identité, a été brièvement introduite dans ce chapitre. Il s'agit d'extraire à partir de toutes les sources disponibles, des connaissances caractérisant la classe en tenant compte de la nature de ses relations avec les autres classes d'une part et du degré de fiabilité des différentes combinaisons des sources pour la représenter, d'autre part. Un résumé de ces connaissances est présenté sous forme de deux structures conceptuelles que nous avons nommée respectivement « carte possibiliste biométrique » et « carnet d'identité ». L'ensemble de ces concepts relatifs aux individus pris en charge par un système biométrique, constitue un support pour l'interprétation d'une scène biométrique, que le système consulte à chaque requête d'identification. Les concepts de la carte biométrique et du carnet d'identité ne présentent pas uniquement une référence biométrique d'un individu, mais aussi des informations concernant les modalités à acquérir avec des ordres de confiances respectifs et spécifiques à chaque individu, ainsi que les modalités qui permettent de mieux le distinguer de chacun des autres individus de la base d'une façon personnalisée. Les préférences sur les modalités à employer se répercutent sur l'usage des outils permettant leurs acquisitions, ce qui promeut l'économie sur l'énergie et sur la durée de vie des capteurs.

Chapitre V: **Conclusion générale et perspectives**

Ce manuscrit rapporte mes travaux de recherche réalisés dans l'objectif de contribuer au développement de mécanismes de raisonnement possibiliste pour aider à l'interprétation de scènes et l'aide à la prise de décision dans des contextes affectés d'incertitude épistémique. Un processus automatique d'aide à la décision ne peut opérer que sur une « projection » de la scène réelle, dans un espace de quantifications numériques d'observations ou de mesures physiques prises par des capteurs. Les informations obtenues de cette restriction du monde réel, ne peuvent en aucun cas être parfaites. Elles sont affectées d'incomplétudes de différentes formes, reliées à des facteurs sémantiques (des connaissances a priori nécessaires à la compréhension de la scène), des facteurs sensoriels (le type, la nature et l'état du capteur) et des facteurs techniques (caractéristiques et notions en relation avec le support de l'information). Le traitement d'informations imparfaites collectées à partir de la scène, leurs représentations de façon à prendre en compte leurs imperfections d'une part et les connaissances a priori sur cette scène d'autre part, ainsi que leurs interprétations, constituent des difficultés majeures rencontrées lors de la conception d'un système d'aide à la décision. Afin de faire face à ces difficultés, la théorie des possibilités a été choisie comme cadre de travail commun, pour l'étude et la conception des systèmes d'aide à la prise de décision. La capacité de cette théorie pour la représentation des connaissances et des informations hétérogènes et imparfaites, ainsi que les outils de raisonnement qu'elle fournit, constituent les atouts essentiels justifiant son choix.

A. Recours à la théorie des possibilités comme cadre d'étude

Le premier chapitre a été conçu de façon à présenter le cadre de la recherche, de point de vue contexte et outils. Nous avons d'abord présenté les différents niveaux d'abstraction de l'information afin d'éclaircir leurs apports respectifs et relatifs dans le processus de la compréhension. L'incomplétude étant inévitable dans toute acquisition de l'information, ses différentes formes, encore dites imperfections, ont été présentés. Conscients de la grande nécessité de la prise en compte de l'imperfection dans l'information dans toute procédure d'aide à la prise de décision et d'interprétation de scène, différents cadres mathématiques, les plus reconnus dans le traitement de l'incertitude, ont été présentés. Les quatre principales théories abordées sont: la théorie des probabilités, la théorie des fonctions de croyance, la théorie des ensembles flous et la théorie des possibilités. Leurs manières respectives à considérer les différentes formes de l'imperfection et leurs aptitudes à les modéliser, ont été synthétisés. Cette synthèse a permis de positionner la théorie des possibilités comme un cadre théorique performant pour traiter des informations imprécises et incertaines. Elle permet d'aborder l'incertitude décisionnelle, où les informations disponibles pour la décrire sont de nature ambiguë ou incomplète aussi bien quantitativement que qualitativement (on parle de connaissances obtenues de la part des experts qui sont décrites par le biais de variables linguistiques, ou bien des connaissances provenant de l'utilisation d'un faible nombre d'échantillons extraits de la scène, ou encore des connaissances qui sont obtenues à travers multiples sources de nature hétérogènes). La deuxième partie du chapitre s'est alors concentrée sur cette théorie en exposant les différents concepts et outils qui la distinguent, tels que les mesures de possibilité et de nécessité, la distribution de possibilités et leurs modes de normalisation, des mesures d'incertitudes, des règles de décision possibiliste et des opérateurs de fusion d'informations possibilistes. Les distributions de possibilités étant la clé de la

théorie pour encoder les connaissances liées à la description des classes thématiques présentes dans une scène, une synthèse des différentes approches connues pour les estimer, a été présentée. Ces techniques d'estimation se répartissent globalement en deux catégories : celles qui sont guidées par l'expert et celles qui reposent sur des processus d'inférence. Un intérêt particulier a été porté aux méthodes de transformation d'une distribution de probabilités en une distribution de possibilités car elle permet l'intégration des données statistiques.

Ainsi, d'une part, ce chapitre a permis de montrer que la théorie des possibilités utilise un ensemble relativement simple de stratégies, tout en révélant le comportement complexe de leurs interactions. Ceci a permis à la théorie de répondre au besoin d'un formalisme simple et efficace des situations ambiguës et mal définies. D'autre part, ce chapitre a fourni les outils et notions qui ont servi dans la construction de nos contributions dans les différents niveaux du raisonnement possibiliste, allant de la modélisation jusqu'à l'aide à la prise de décision.

B. Modélisation possibiliste des connaissances

Dans le deuxième chapitre, nous avons présenté nos contributions au domaine de la représentation possibiliste des connaissances extraites à partir des connaissances disponibles. Un modèle n'est jamais parfait, ni totalement représentatif de la réalité mais il est d'autant plus fiable qu'il se rapporte à ce qu'on espère en déduire. C'est la raison pour laquelle il est nécessaire de définir la finalité avant toute modélisation. Deux contributions ont été décrites dans ce chapitre reliées à deux finalités spécifiques.

La première finalité visée consistait à modéliser une classe thématique en disposant juste d'un échantillon de petite taille (un très petit nombre de valeurs). Le nombre réduit de données disponibles entraîne une incomplétude informationnelle. Nous avons modifié une approche d'inférence de distribution de possibilité à partir de distribution de probabilités, construite sur la base d'intervalles de confiance simultanés. L'utilisation des intervalles de confiance était intéressante dans le sens où leur flexibilité permet d'étendre le résultat d'un échantillon à une population, remédiant ainsi à la faible informativité des données. Cependant leur construction se basait sur des tests paramétrés, nécessitant un nombre non réduit de données pour ajuster les paramètres. Notre intervention sur ce point, par l'utilisation d'une technique de construction des ICs, plus approximative et moins gourmande en données, a permis de créer un modèle possibiliste relativement fiable avec des données très pauvres aussi bien en qualité qu'en quantité. Notre apport s'est aussi situé dans la stratégie conçue pour l'évaluation de cette approche par rapport à d'autres de l'état de l'art. Les tests d'évaluation ont été conduits sur trois types de données : synthétiques, benchmarks et acquises d'une scène réelles. Les critères d'évaluation étaient quantitatifs et qualitatifs basés sur les principes de cohérence entre les distributions de probabilité et celles de possibilité d'une part, et sur la mesure de similarité entre les distributions inférées sur la base des échantillons et celles de l'expert, d'autre part. L'appauvrissement des échantillons a été poussé jusqu'au nombre de 5 valeurs et l'approche proposée s'est distinguée par un meilleur degré de similarité et un écart d'incertitude le plus réduit par rapport à la distribution de possibilités de l'expert.

Notre deuxième contribution a visé la représentation des connaissances pixelliques en vue de la classification du pixel dans l'une des classes thématiques d'une scène imagée. Quatre niveaux de représentation des connaissances pixelliques ont été distingués : celui des données

brutes, celui des primitives, le niveau sub-sémantique et finalement le niveau symbolique où l'étiquette d'appartenance est attribuée. Deux approches de modélisation possibilistes ont été proposées : l'une ponctuelle et l'autre contextuelle. La première a un concept basique et ne considère que l'aspect ponctuel du pixel. Le modèle consiste en une projection des attributs calculés au niveau du pixel sur les distributions de possibilité des classes thématiques. En revanche, la deuxième approche révèle un tout nouveau concept. Elle est voulue opérer dans le niveau sub-sémantique, où les connaissances pixelliques sont représentées via un modèle possibiliste, à savoir une distribution de possibilités locale, résumant l'information contextuelle apportée par le voisinage spatial du pixel. Par cette procédure, le pixel étant d'une granularité inférieure à celle de la région (constituant la classe thématique), se trouve représenté par un modèle semblable et comparable à celui de la région. La décision de son attribution à une classe thématique donnée, sera basée sur un test de cohérence entre deux modèles du même calibre informationnel. Les deux approches proposées, ponctuelle et contextuelle, ont été évaluées et comparées à d'autres approches de l'état de l'art, opérant dans les différents niveaux sémantiques de représentation de connaissances pixelliques. L'expérimentation a été menée sur deux types d'images : une image synthétique, de façon à bien contrôler et connaître les modèles de fabrication de l'image et des images optiques médicales qui présentent des contextes réelles d'ambiguïté. La modélisation possibiliste contextuelle s'est avérée beaucoup plus intéressante que la modélisation ponctuelle. L'utilisation d'un modèle local d'estimation se basant sur des primitives contextuelles possibilistes, a pu assurer des performances d'une efficacité supérieure aux autres approches, notamment pour des qualités dégradées de l'image.

Courant nos activités de modélisation de connaissances possibilistes, nous étions confrontés à une question qui s'est toujours alliée à cette problématique, à savoir comment réduire la dimensionnalité de l'espace des attributs à la base desquels sont construits nos modèles, afin d'améliorer la qualité de représentation d'une part et alléger la complexité des calculs d'autre part. Les procédures classiques de sélection d'attributs, ne permettaient pas de profiter de la flexibilité qu'apporte l'approximation possibiliste à ses modèles. C'est pour cette raison que nous avons proposé une nouvelle approche qui intègre la notion de l'incertitude dans le formalisme de Shapley pour la sélection d'attributs. Deux tests de sélection ont été conçus : l'un mettant en épreuve la capacité de représentation d'une classe de l'attribut et l'autre met à l'épreuve la capacité de séparation entre les classes de l'attribut. Les conflits entre ces deux tests ont été gérés et résolus. Suite à des tests d'évaluation, effectués sur des données synthétiques et des benchmarks, la procédure de sélection proposée a montré une bonne optimisation entre nombre d'attributs sélectionnés et performances de classification, comparée à six autres approches.

C. Concepts et mécanismes reliés à l'aide à la prise de décision dans des contextes possibilistes

Après la modélisation, nous nous sommes intéressés à la problématique d'aide à la prise de décision dans le cadre possibiliste. Courant notre recherche plusieurs questions ont été relevées et abordées dans cette problématique, dont certaines ont été traitées et des solutions ont été conçues et testées et d'autres sont encore en cours d'étude et de développement. Nous

avons choisi de présenter certaines d'entre elles, sans nous limiter aux solutions développées, afin de présenter, autre que nos contributions, nos tendances de recherche dans l'aide à la prise de décision dans un contexte possibiliste. Dans cette thématique, nous nous sommes spécialement intéressés à deux concepts sur lesquels se basent la majorité des processus d'aide à la prise de décision, à savoir : la notion de similarité et la notion de fusion. Dans le troisième chapitre de ce rapport, nous avons présenté nos apports reliés à ces deux concepts.

Concernant la similarité, suite à une étude de l'état de l'art, nous avons noté que la majorité des mesures de similarité utilisées dans un contexte possibiliste, consiste en une adaptation des mesures conventionnelles. Spécialement, les mesures de similarité floues sont reprises par la substitution de la fonction d'appartenance par une distribution de possibilités. Il est cependant clair que les modèles possibilistes offrent des spécificités, qu'il convient de considérer et d'exploiter lors de l'évaluation de la nature des liens entre deux modèles possibilistes. Dans ce sens, nous avons commencé par présenter les différentes propriétés d'une mesure de similarité possibiliste, avant de brièvement présenter et analyser six mesures de similarité possibiliste de l'état de l'art. Remarquant que l'aspect morphologique d'une distribution de possibilité renseigne sur la quantité d'information qu'elle modélise, une catégorisation des mesures de similarité, basée sur la forme a été proposée. La prise en compte de la forme dans une mesure de similarité, se fait à travers l'évaluation de la quantité d'incertitude contenue dans chacun des modèles possibilistes à comparer. Remarquant qu'une combinaison entre l'évaluation de l'écart d'incertitudes et l'écart d'amplitudes (point à point) entre deux distributions de possibilités, révèle mieux la similitude existante entre ces deux distributions, nous avons proposé une nouvelle mesure de similarité dans ce sens. Notre proposition nommée « *Infospécificity* » combine une évaluation de la différence de spécificité entre deux distributions, avec une mesure de distance usuelle assurée par la distance de « Manhattan ». Face à la difficulté souvent rencontrée pour l'évaluation d'une mesure de similarité, nous avons proposé une méthodologie se basant sur une évaluation des valeurs attribuées pour des objets similaires et des valeurs attribuées pour les objets différents par une mesure de similarité. L'ensemble des valeurs attribuées pour des objets similaires construit la classe des « similaires » et celui des valeurs attribuées pour les objets différents, construit la classe des « dissimilaires ». Il est évident qu'une mesure de similarité est d'autant plus efficace qu'elle offre une meilleure séparation de ces deux classes. L'évaluation de l'ambiguïté due au « chevauchement » entre ces deux classes a été assurée par des courbes ROC, en considérant l'air sous la courbe ROC comme critère principal d'évaluation. Les tests effectués ont montré l'avantage des mesures de similarité basées sur la forme. Une supériorité a été notée pour les mesures combinant l'évaluation basée sur la forme et celle des distances point à point. Notre approche basée sur la mesure de spécificité a spécialement dépassé les autres dans pratiquement tous les tests, qui ont été conduits sur quatre benchmarks. Un exemple d'exploitation de la mesure de similarité proposée a été présenté. Il s'agit d'un arbre de décision possibiliste qui a servi à la reconnaissance d'objets dans un environnement domestique. Comparée à SVM, l'approche proposée a permis d'avoir un taux global de reconnaissance de 81.87% contre 72.18% pour SVM.

Nos travaux sur la fusion dans un cadre possibiliste ont également été présentés dans le troisième chapitre. Nous avons d'abord décrit notre contribution dans la sélection d'un

opérateur de fusion, adéquatement aux informations à combiner. Un processus de fusion est appelé quand il est question de combiner des informations provenant de sources multiples et/ou de natures distinctes. Les degrés de cohérence et d'accord entre les sources à combiner, influent fortement sur la qualité du résultat de la fusion en terme de fiabilité dans la représentation des connaissances. Ainsi des informations redondantes ne doivent pas être combinées de la même façon que des informations complémentaires. La manière de combiner étant fortement liée à l'opérateur utilisé pour la fusion, un choix raisonné de celui-ci est à faire selon des règles basées sur la nature des informations à fusionner et des relations qui les spécifient. La contribution qui a été proposé dans ce contexte, consistait en une modification apportée à une approche de l'état de l'art. La sélection de l'opérateur de fusion a été assurée conformément à des règles se rapportant à l'évaluation du conflit entre les informations d'une part et de leurs cohérences d'autre, part en se basant sur deux concepts fondamentaux, l'un relié à la théorie des possibilités et l'autre à celle des évidences. En effet, l'interprétation du comportement conflictuel des sources a été faite sur la base d'une mesure de l'inconsistance de la distribution de possibilités résultante de leur fusion. Alors que la cohérence entre les sources a été évaluée par une distance mesurée entre les probabilités pignistiques de leurs fonctions de masses respectives. Une évaluation comparative de notre proposition avec l'approche originale et d'autres opérateurs basiques de la fusion, a été menée sur deux benchmarks, « IRIS » et « Diabète ». L'évaluation de la pertinence de la fusion s'est rapportée aux tests sur la diminution de l'incertitude au sein d'une même classe et la maximisation de la séparation entre les distributions des différentes classes, suite à la fusion. La modification rapportée à l'approche originale a, globalement, fourni de meilleurs résultats.

Deux mécanismes d'aide à la prise de décision, ont été proposés. Le premier, nommé MCDA, se base sur une analyse multicritère de la décision : MCDA. Le système fait combiner plusieurs sources de connaissances, affectée chacune par un facteur de confiance dépendant de degrés de qualité, préalablement estimés pour les sources conditionnellement aux classes. La fusion entre les sources est de ce fait assurée d'une façon « sur mesure », selon la classe à représenter et selon la qualité des informations disponibles. Le système est en cours de développement dans le cadre d'une thèse travaillant sur l'identification biométrique.

Le deuxième mécanisme présenté est basé sur une fouille de connaissances possibiliste. L'extraction des connaissances par le processus proposé, génère une nouvelle structure conceptuelle, nommée « carte possibiliste », pour la représentation concise mais richement informative de chaque classe. Le processus de construction des « cartes possibilistes », basé sur un arrangement spécifique des combinaisons entre les sources, génère des coefficients de confiance pour les couples (combinaison/classe). Ces coefficients sculptent la décision prise étant donné un état de connaissance sur le monde réel. Ce mécanisme a été testé pour une classification de l'état de sol en trois classes : plat, escalier ascendant, escalier descendant, en se basant sur des données ultrasonores. Comparé à une approche de classification basée sur des méthodes conventionnelles d'analyse du signal que nous avons proposée dans [AKS 2014], ce mécanisme de fouille possibiliste temporelle a permis un taux de classification de 90% contre un taux de 75% pour l'autre approche.

D. Construction et interprétation de scènes

L'interprétation de scènes a été l'objet du quatrième chapitre de ce rapport. Le processus d'interprétation de scène dans un contexte possibiliste, vise à identifier le contenu sémantique de la scène observée par l'exploitation conjointe des données radiométriques issues de systèmes d'imagerie et de certaines connaissances a priori. Ces connaissances portent à la fois sur la scène analysée, sur les capteurs disponibles et les conditions d'acquisition, généralement exprimées de manière vague, soit par des experts ou résultant d'observations et de jugements. Il est de ce fait évident, qu'un processus d'interprétation dépende fortement de la scène elle-même. Courant nos travaux, nous nous sommes investis dans la construction de certaines scènes pour tester nos systèmes d'interprétation de scènes et d'aide à la prise de décision. Certains de ces travaux ont été présentés dans le chapitre quatre. Deux scènes ont été visées : biométrique et d'environnement intérieur. Pour la première deux dispositifs d'acquisition d'iris ont été conçus et construits. Pour la deuxième un système d'acquisition multicapteurs a été mis au point. Cette activité nous a permis de bien gérer les contenus et les paramètres de nos scènes, selon nos besoins pour les tests et l'évaluation de nos approches, conformément à nos objectifs. D'autre part, cette activité a constitué une action collective qui a réuni plusieurs membres de notre équipe avec des partenaires industriels, autour d'un but commun, à travers un processus dynamique et évolutif de communication, de décision, d'intervention et d'apprentissage.

Dans la deuxième partie du quatrième chapitre, nous avons proposé deux approches d'interprétation de scènes, médicale pour la première et biométrique pour la deuxième. Les deux approches se réunissent sur un avantage commun, à savoir celui d'exploiter l'information « croisée » pour extraire des connaissances plus pertinentes pour le processus d'interprétation.

Le premier processus consiste en une diffusion itérative de connaissances possibilistes pour une segmentation possibiliste d'images. L'approche est basée sur la construction de structures conceptuelles riches en connaissances spécifiques aux différentes classes de l'image, nommées « cartes de possibilité ». Ces « carte de possibilité » consistent en une projection des données capteurs, dans les représentations possibilistes des différentes classes, préalablement définies par l'expert ou sur la base de ses connaissances. Le principe d'interprétation est basé sur une propagation itérative des connaissances possibilistes sur les cartes construites, afin de faciliter la décision sur la classification pixellique pour la formation de régions homogènes. Un aspect important du raisonnement humain a été modélisé dans ce processus d'interprétation de scène, en intégrant de manière itérative, l'information contextuelle de la scène. En outre, cette approche permet d'évaluer les connaissances possibilistes sous forme de valeurs de possibilités d'appartenance des pixels, afin d'affiner les résultats de segmentation. Deux stratégies de diffusion, ont été proposées dans cette étude : la première effectue la mise à jour de la carte possibiliste relative à chaque classe d'une façon indépendante des autres classes, alors que la deuxième fait intervenir l'information croisée entre les classes. L'approche a été testée, d'abord sur une image synthétique, puis sur des images mammographiques de la base MIAS, préalablement détournées par un radiologue. Une évaluation comparative et critique a été menée en considérant d'autres approches de l'état de

l'art largement utilisées pour les images mammographiques. L'approche a montré sa supériorité dans les conditions de faibles connaissances (contraste et bruit) et un caractère bien intéressant de stabilité.

Le deuxième processus proposé au quatrième chapitre, consiste en une interprétation d'une scène biométrique par fouille d'identité. Le processus proposé est encore en cours de développement. Il vise la génération de deux structures conceptuelles informatives nommées « carte possibiliste biométrique » et « carnet d'identité », pour chaque classe se référant à un individu, résumant des connaissances concernant la classe elle-même ainsi que les couples (classe/comboinaison de sources) et (classe/classe). Ce bilan informatif fournit plusieurs renseignements pertinents et propres à chaque individu et construit ainsi une base pour l'interprétation de son identité biométrique. Le principe de construction des deux structures a été détaillé dans la dernière section du chapitre quatre.

E. Bilan

Après avoir synthétisé les différents travaux réalisés, nous pouvons conclure que l'apport de la théorie des possibilités pour la conception des systèmes d'aide à la décision et d'interprétation de scènes, intégrant différentes sources de connaissances et différents modes de raisonnement, est sans doute très prometteur et ouvre des perspectives très intéressantes. En effet, cette théorie se distingue par une capacité particulière pour la modélisation et la représentation des connaissances incomplètes, mal définies et représentant des concepts ambigus. Elle offre également des outils parfaitement bien adaptés au développement de mécanismes de raisonnement simples mais efficaces. Les expérimentations de nos différentes approches ont été conduites sur trois types de données : synthétiques de façon à intégrer des indicateurs contrôlant la qualité des données selon les besoins et les objectifs des tests, des benchmarks de façon à permettre la comparaison des performances des approches proposées par rapport à d'autres de l'état de l'art et enfin des données réelles collectées selon nos objectifs. Les résultats obtenus par l'application de nos approches proposées sur ces données, ont été très encourageants et ont montré l'efficacité de la théorie des possibilités comme un cadre de raisonnement sur des connaissances incertaines (pauvres en qualité et/ou en quantités, provenant de sources distinctes et/ou de natures distinctes, de type témoignages). Il a été spécialement noté un comportement intéressant de stabilité et une efficacité supérieure dans des conditions de faibles connaissances.

F. Perspectives

En ce qui concerne, les perspectives théoriques reliées aux différentes problématiques abordées, plusieurs aspects pourront être auscultés.

Notre travail sur la modélisation possibiliste des connaissances, par exemple, a soulevé plusieurs questions d'ordre méthodologique. En effet, il est vrai que la théorie de possibilités offre une flexibilité dans la représentation des données, de façon que ses modèles se trouvent bien appropriés dans les cas de modélisation de l'information incomplète. C'est la flexibilité dans leurs présentations des données qui les rendent plus appropriés dans les cas de manque de données. Plusieurs questions se posent à ce niveau, telles que : « Jusqu'à quelle taille de l'échantillon peut-on se permettre de descendre tout en gardant une bonne représentativité de

la classe ? », « Comment garantir la qualité de ces échantillons », « Comment évaluer la représentativité de la classe ? ». Ces questions soulèvent aussi d'autres points importants tels que la mesure de la qualité des données, l'évaluation de la qualité du modèle et la mesure de la performance de l'approche face aux imperfections. Ces points sont objet de nos activités de recherche futures, dont certains sont déjà entamés. En effet, des études sur certaines pistes sont lancées :

- ✓ Etude sur la représentativité d'un modèle possibiliste construit sur la base d'un échantillon de petite taille de données dans un contexte big data.
- ✓ Etude de la notion d'incertitude, spécialement de point de vue significations physiques des différentes mesures connues dans la littérature, afin de définir des règles pour les maîtriser et élucider les informations utilisées pour la modélisation
- ✓ Etude de la notion de la qualité et son exploitation dans la caractérisation de l'information en vue de l'obtention de meilleures fiabilités dans l'exécution de tâches telles que l'aide à la prise de décision ou autres.

Durant nos travaux, nous avons proposé des méthodologies d'évaluation pour nos contributions. Ces méthodologies apparaissent aussi comme des contributions du travail car elles proposent des stratégies d'analyse et ouvre des perspectives de recherche dans d'autres domaines, spécialement celui des critères d'évaluation. En effet, nous avons noté que les critères d'évaluation quantitatifs usuels des systèmes d'aide à la décision, se trouvent parfois incapables de concrétiser une évaluation fidèle à la réalité des performances observées quant à la décision prise. Ceci nous a motivés à lancer une étude sur les critères d'évaluation des systèmes d'aide à la décision, en vue de concevoir une méthodologie pour les évaluer et encore une approche de sélection de critères d'évaluation adéquats à l'objectif de l'évaluation des systèmes d'aide à la décision.

En termes d'aide à la décision, les approches proposées (le système de décision basé sur l'analyse multicritère de la décision, la fouille possibiliste, l'arbre de décision) offrent au décideur des degrés de possibilités pour chacune des alternatives. Nous avons aussi procédé de sorte qu'il est possible non seulement de choisir une alternative mais également de savoir avec quelle force cette alternative est soutenue. Toutefois, cette stratégie constitue une première approche et certains points méritent d'être améliorés. Un premier point à soulever consiste à développer les règles de décision. Un deuxième point concerne une limitation principale qui apparaît lorsque la cardinalité du cadre de discernement devient trop importante. Dans ce cas, les opérations de combinaison deviennent trop gourmandes en complexité algorithmique. Pour obtenir des résultats dans des délais raisonnables, des procédés de diminution des temps de calcul doivent être développés.

Dans nos propositions d'approches pour l'interprétation de scènes, nous n'avons pas représenté l'évolution du temps. Les systèmes proposés jusqu'ici ne permettent pas d'intégrer des connaissances sur l'évolution dynamique des scènes. La représentation de cette dynamique pourrait être étudiée.

Par ailleurs, notre intérêt à continuer à lier nos activités de recherche à des applications, persiste et trouve dans le succès réalisé avec les premières expériences, tant en terme d'atteinte des objectifs techniques qu'en terme d'interaction entre les différents membres de

l'équipe et de collaboration entretenue avec des parties à caractère industriel, une grande motivation.

Références

- [AAG 2013] B. Alsahwa, S. Almouahed, D. Gueriot et B. Solaiman, « Iterative possibility distributions refining in pixel-based images classification framework », ICPRAM 2013 : 2nd International Conference on Pattern Recognition Applications and Methods, Feb 2013, Barcelone, Spain, SciTePress, pp.176-181, 2013.
- [AAT 2005] S. Auenphanwiriyaikul, S. Attrapadung, S. Thovuttikul and N. Theera-Umpon: Breast Abnormality Detection in Mammogram Using Fuzzy Inference System, IEEE. Fuzzy System, pp. 155-160, 2005.
- [ADK 2000] Andres, F., Dessaigne, N., Kawtrakul, A., Martinez, J., Mouaddib, N. et Ono K. (2000) « A Flexible Approach to Retrieve Medicinal Plant Images. », NII J (Natl Inst Inform), no. 1, p. 23-31, 2000
- [AE 2017a] A. Ammar and Z. Elouedi, "From Crisp to Soft Possibilistic and Rough Meta-clustering of Retail Datasets," in International Conference on Intelligent Systems Design and Applications, 2017, pp. 485-495.
- [AE 2017b] A. Ammar and Z. Elouedi, "Selection of Initial Modes for Rough Possibilistic K-Modes Methods," in International Symposium on Methodologies for Intelligent Systems, 2017, pp. 77-86.
- [AKS 2015] S. Ammar Bouhamed, I. Khanfir Kallel, D. Sellami Masmoudi and B. Solaiman "Feature selection in possibilistic modeling", Pattern Recognition, Volume 48 Issue 11, p. 3627-3640, 2015.
- [AIS 2012] Al Sun, M. H. (2012), « Indexation guidée par les connaissances en imagerie médicale », Traitement des images. Télécom Bretagne, Université de Bretagne Occidentale, 2012. Français. <tel-00719587>
- [ALS 2012] Alsun, M. H., Lecornu, L. et Solaiman, B. (2012), « Raisonnement possibiliste en diagnostic médical : application au diagnostic de lésions endoscopiques », IRBM : ingénierie et recherche biomédicale, 33, 29-34, février 2012.
- [Als 2014] Alsahwa, B. (2014), « Représentation d'un environnement par un système multi-capteurs : fusion et interprétation de scène. Vision par ordinateur et reconnaissance de formes », [cs.CV]. Télécom Bretagne; Université de Rennes 1, 2014. Français. <tel-01206283>
- [Amm 2014] S. Ammar, (2014), « Analyse et traitement possibiliste de signaux ultrasonores en vue d'assistance des non-voyants », rapport de thèse soutenue à l'Ecole Nationale d'Ingénieurs de Sfax.
- [ARB 2007] M. Adel, M. Rasigni, S. Bourennane, and V. Juhan, "Statistical segmentation of regions of interest on a mammographic image", EURASIP Journal on Advances in Signal Processing, Vol. 2007, pp. 1-8, Article ID 49482, 2007.
- [AS 2004] V. Athitsos, S. Sclar off, Boosting Nearest Neighbor Classifiers for Multiclass Recognition, Technical Report No. 2004-006, Boston University Computer Science, 2004.
- [ASA 2016] B. Alsahwa, B. Solaiman, S. Almouahed, É. Bossé, and D. Guériot, "Iterative refinement of possibility distributions by learning for pixel-based classification," IEEE Transactions on Image Processing, vol. 25, pp. 3533-3545, 2016.
- [ASB 2016] B. Alsahwa, B. Solaiman, E. Bosse, S. Almouahed, and D. Gueriot, "A method of spatial unmixing based on possibilistic similarity in soft pattern classification," Fuzzy information and engineering, vol. 8, pp. 295-314, 2016.
- [BBE 2004] Ben Amor, N., Benferhat, S., & Elouedi, Z. (2004). « Qualitative classification and evaluation in possibilistic decision trees », In Proceedings of the IEEE International Conference on Fuzzy Systems, 2, 653-657.
- [BD 1995] Bârdossy, A. et Duckstein, L. (1995), « Fuzzy rule-based modeling with applications to geophysical, biological and engineering systems », CRC Press, Boca Raton, Florida.

- [BGE 1996] L. S. Bins, L. M. Garcia Fonseca, G. J. Erthal, F. Mitsuo, “Satellite imagery segmentation: a region growing approach”, In Proceedings of the 8th Brazilian Symposium Remote Sensing, pp.1-4, 1996.
- [Bis 2000] Bisson, G. (2000), « La similarité: une notion symbolique/ numérique », Apprentissage Symbolique/ Numérique, (tome 2), Eds. Moulet Brito, Edition CEPADUES, pp. 169-201, 2000.
- [BL 2013a] K. Bache, M. Lichman, UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/machine-learning-databases/glass>), University of California, School of Information and Computer Science, Irvine, CA, 2013.
- [BL 2013b] K. Bache, M. Lichman, UCI Machine Learning Repository (<http://archive.ics.uci.edu/ml/machine-learning-databases/liver-disorders>), Irvine, CA : University of California, School of Information and Computer Science, 2013.
- [Bla 1993] C. Black, UCI Machine Learning Repository (<https://archive.ics.uci.edu/ml/machine-learning-databases/iris/>), University of California, School of Information and Computer Science, Irvine, CA.
- [Bla 2016] Blachon, D. (2016), « Reconnaissance de scènes multimodale embarquée. Intelligence artificielle », [cs.AI]. Université Grenoble Alpes, 2016. Français. <NNT : 2016GREAM001>. <tel-01298709>
- [Blo 2003] Bloch, I. (2003), « Fusion d’informations en traitement du signal et des images », Hermes Science Publications (Editeur), 2003.
- [BN 1992] Bender, H. et Nuather, W. (1992), « Fuzzy Data Analysis », Dordrecht, Kluwer Academic Publishers, 1992.
- [Bou 2007] MS. 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 2007, Université De Reims Champagne-Ardenne.
- [BP 2009] S.S. Basha, K. S. Prasad “Automatic Detection of Breast Cancer Mass in Mammograms using Morphological Operators and Fuzzy C – Means Clustering”, Journal of Theoretical and Applied Information Technology, vol.5, n.6, pp. 704-709, 2009.
- [BRB 1996] Bouchon-Meunier, B., Rifqi, M. et Bothorel, S. (1996), « Towards general measures of comparison of objects. », Fuzzy Sets and Systems, Elsevier, , 84 (2), pp.143–153. <hal-01072162>
- [BRL 2008] Bouchon-Meunier, B., Rifqi, M. et Lesot, M-J. (2008), « Similarities in fuzzy data mining: from a cognitive view to real-world applications », Springer LNCS, vol. 5050, pp. 349-367, 2008.
- [BS 2016] Bossé, É., Solaiman, B. (2016), « Information fusion and analytics for big data and IoT », Artech House (2016).
- [CC 1997] H. D. Cheng and J.R. Chen: Automatically determine the membership function based on the maximum entropy principle, Inf. Sci, vol. 96, no. 3-4, pp. 163-182, 1997.
- [Cha 2014] M. Chaabene, (2014), « Etude et comparaison d’approches d’estimation de distributions de possibilités dans des contextes ambigus », rapport de Mastère de Recherche en Informatique et Multimédia (I2M), soutenu à l’Institut Supérieur d’Informatique et de Multimédia de Sfax.
- [Char 2014] A. Charfi, (2014), « Evaluation de mesures de similarité en vue de leur utilisation dans des arbres de décision dans un cadre possibiliste », rapport de Mastère de Recherche en Informatique et Multimédia (I2M), soutenu à l’Institut Supérieur d’Informatique et de Multimédia de Sfax.
- [Chr 2006] Christian, S. « Deux prix Nobel pour la théorie des jeux », Revue d’économie politique, 2006/2 (Vol. 116), p. 133-145. DOI : 10.3917/redp.162.0133. URL : <https://www.cairn.info/revue-d-economie-politique-2006-2.htm-page-133.htm>
- [CKA 2014] M. Chaâbène, I. Khanfir Kallel, S. Ammar Bouhamed, M. Bellaj, W. Bouchaala, “CIs based strategy for estimating possibility distribution of small data samples”. 1st IEEE International

- Conference on Advanced Technologies for Signal & Image Processing ATSIP 2014, March 17-19, 2014, Sousse-Tunisia.
- [CL 2011] C. -C. Chang and C. -J. Lin, “LIBSVM: a library for support vector machines,” *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, article 27, pp. 1–27, 2011, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. View at Google Scholar.
- [CS 2017] Couso, I., Sánchez, L. (2017), « Additive similarity and dissimilarity measures », *Fuzzy Sets and Systems*, Vol. 322, pp.35–53.
- [CSM 2006] H. Cheng, X. Shi, R. Min, L. Hu, X. Cai and H. Du “Approaches for automated detection and classification of masses in mammograms”, Elsevier, vol. 39, n.4, pp. 646-668, 2006.
- [CT 1986] Civanlar, M.R. & Trussell, H.J. (1986), « Constructing membership functions using statistical data », *Fuzzy Sets and Systems* 18 (1): 1-13.
- [CWC 2000] Y. H. Chen, W.J. Wang and C.H. Chiu: New estimation method for the membership values in fuzzy sets, *Fuzzy Sets and Systems*, vol. 112, no. 3, pp. 521-525, 2000.
- [Dam 2014] W. Damak, (2014), « Reconnaissance et classification des objets dans un environnement domestique », rapport de Mastère de Recherche en Informatique et Multimédia (I2M), soutenu à l’Institut Supérieur d’Informatique et de Multimédia de Sfax.
- [DB 2004] J. G. Dy, C. Brodley, Feature Selection for unsupervised learning, *J. Mach. Learn. Res.* 5 (2004).
- [Den 2007] Denguir, A. (2007), «Un Cadre Possibiliste pour l’Aide à la Décision Multicritère et Multi-acteurs Application au Marketing et au Benchmarking de sites E-commerce», Thèse de Doctorat, Université de Savoie, France, 2007.
- [DFM 2004] Dubois, D., Foulloy, L., Mauris, G. et Prade H. (2004), « Probability-Possibility Transformations, Triangular Fuzzy Sets, and Probabilistic Inequalities. pp. 273-297, 2004.
- [DHG 2010] R. B. Dubey, M. Hanmandlu and S. K. Gupta “Level set detected masses in digital mammograms”, *Indian Journal of Science and Technology*, vol. 3, n. 1, 2010.
- [DHS 2001] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern Classification*, 2nd edition, John Wiley & Sons, NewYork, 2001.
- [DKR 2014] W. Damak, I. Khanfir and I. Rebai, “Semantic object recognition by merging decision tree with object ontology”. 1st IEEE International Conference on Advanced Technologies for Signal & Image Processing ATSIP 2014, March 17-19, 2014, Sousse-Tunisia.
- [DP 1982] D. Dubois and H. Prade: On several representations of an uncertain body of evidence, in *Fuzzy Information and Decision Processes*, M.M. Gupta and E. Sanchez, Eds., North- Holland, Amsterdam, pp.167-181, 1982.
- [DP 1986] Dubois, D. et Prade, H. (1986), «Fuzzy sets and statistical data. », *European Journal of Operational Research*, 25, 345–356, 1986.
- [DP 1987] Dubois D. et Prade, H. (1987), « An alternative approach to the handling of subnormal possibility distributions: A critical comment on a proposal by Yager », *Fuzzy Sets and Systems*, vol. 24, no. 1, p. 123-126, oct. 1987.
- [DP 1988a] Dubois, D. et Prade, H. (1988), « The treatment of uncertainty in knowledge-based systems using fuzzy sets and possibility theory », *Int. Journal of Intelligent Systems* 3 : 141-165.
- [DP 1988b] Dubois, D. et Prade, H. (1988), « An introduction to possibilistic and fuzzy logics », P. Smets, E.H. Mamdani, Dubois, D. et Prade, H. (éditeurs), *Non-Standard logics for Automated Reasoning* (Academic Press, NewYork): 288-326.
- [DP 1988c] Dubois, D. et Prade, H. (1988), « Representation and combination of uncertainty with belief functions and possibility measures », *Computational Intelligence* 4: 2M-264.
- [DP 1988d] Dubois, D., Prade, H. (1988), « Possibility theory », New York, Plenum Press, 1988.
- [DP 1988e] Dubois, D., Prade, H. (1988), «Théorie des possibilités. Application à la représentation des connaissances en informatique», Masson, 2ème Edition, 1988, 292 pages.

- [DP 1992] Dubois, D., Prade, H. (1992), « Combination of fuzzy information in the framework of possibility theory », *Data Fusion in Robotics and Machine Intelligence*, New York, Academic Press, 1992.
- [DP 1993] Dubois, D. et Prade H. (1993), « On possibility/probability transformations », *Fuzzy Logic*: 103-112,1993.
- [DP 1994] Dubois, D. et Prade, H. (1994), « La fusion d'informations imprécises », *Traitement du signal*, vol. 11, p. 447-458, 1994.
- [DP 1995] Dubois, D. et Prade, H. (1995), « Tolerant fuzzy pattern matching : an introduction ». In *Fuzziness in Database Management Systems*, pages 42-58, 1995.
- [DP 1998] Dubois, D. et Prade, H. (1998), « Possibility theory: qualitative and quantitative aspects. », In *Quantified representation of uncertainty and imprecision* (pp. 169-226). Springer, Dordrecht, 1998.
- [DP 2008] Dubois, D. et Prade, H. (2008), « Panorama des nouvelles méthodes de traitement de l'incertitude et de l'imprécision pour la Maîtrise des Risques industriels », *Institut de Recherche en Informatique de Toulouse (IRIT), CNRS et Université P. Sabatier, France*, 2008.
- [DPS 1993] Dubois, D., H. Prade and S. Sandri, (1993), « On Possibility/Probability Transformations. In: *Fuzzy Logic: State of the Art* », Lowen, R. and M.R. Roubens (Eds.), Kluwer Academic Publication, New York, ISBN-10: 0792323246.
- [DPT 1988] Dubois, D., Prade, H. et Testemale, C. (1988), « Weighted fuzzy pattern matching », *Fuzzy Sets Syst.*, 28(3) :313-331, 1988.
- [DS 1985] B. B. Devi and V. V. S. Sarma: Estimation of fuzzy memberships from histograms, *Inf.Sci.*, vol. 35, no. 1, pp. 43-59, 1985.
- [DSW 2017] S. R. R. Dhara, S. Suryono and C. E. Widodo, « Determination of breast cancer area from mammography images using thresholding method », *International Journal of Innovative Research in Advanced Engineering (IJIRAE)* ISSN: 2349-2163, Issue 06, vol. 4, pp. 80-83, 2017.
- [Dub 1988] Dubois, D. (1988), « Weighted fuzzy pattern matching », *Fuzzy Sets and Systems*, vol. 28, p. 313-331, déc. 1988.
- [Ezi 2012] Eziddin, W. (2012), « Segmentation itérative d'images par propagation de connaissances dans le domaine possibiliste : application à la détection de tumeurs en imagerie mammographique. », *Image Processing. Télécom Bretagne, Université de Bretagne Occidentale*, 2012. French. <tel-00741996>.
- [FG 1996] Ferson, S. et Ginzburg, LR. (1996), « Different methods are needed to propagate ignorance and variability », *Engineering Systems Safety*, vol. 54, 1996, p. 133-144.
- [GA 1997] H. A. Güvenir and A. Akkus, Weighted K Nearest Neighbor Classification on Feature Projections in: *Proceedings of the Twelfth International Symposium on Computer and Information Sciences (ISCIS XII)* S. Kuru, M.U. Caglayan and H.L. Akin (Eds.), Antalya, Turkey, (Oct. 27-29, 1997), 44-51.
- [GFH 2009] M. Gütlein, E. Frank, M. Hall, A. Karwath, Large-scale attribute selection using wrappers, in *Proceedings of the IEEE Symposium on Computational Intelligence and Data Mining*, 2009, pp. 332–339.
- [GG 1984] Geman, S. et Geman, D. (1984), « Stochastic relaxation, gibbs distributions, and the bayesian restoration of images », *IEEE Trans. on Patt. Anal. and Mach. Intell.* 6(6), 721–741.
- [Goo 1965] Goodman, L. (1965), On simultaneous confidence intervals for multinomial proportions, *Technometrics*, 7, 247-254.
- [GW 2002] R. C. Gonzalez and R. E. Woods: “*Digital Image Processing*”, 2ed. Prentice Hall, 2002.
- [GYT 2003] Garmendia, L., Yager, R. R., Trillas, E. et A. Salvador, A. (2003), « On t-norms based measures of specificity », *Fuzzy Sets and Systems*, vol. 133, p. 237–248, janv. 2003.
- [HAN 2004] H. A. Hassanien, J. M. Ali and H. Nobuhara “Detection of Spiculated Masses in Mammograms Based on Fuzzy Image”, *Processing, ICAISC*, vol. 3070, pp. 1002-1007, 2004.

- [HCL 2003] C. W. Hsu, C. C. Chang, C. J. Lin, (2003), « A Practical Guide to Support Vector Classification », Technical Report, Department of Computer Science, National Taiwan University, 2003.
- [HK 1982] Higashi, M. et Klir, G. (1982), « Measures of uncertainty and information based on possibility distributions », *International Journal of General Systems*, vol. 9, no. 1, p. 43-58, 1982.
- [HK 1983] M. Higashi and G. J. Klir, "On the notion of distance representing information closeness: possibility and probability distributions," *International Journal of General Systems*, vol. 9, pp. 103-115, 1983.
- [HS 2010] B. W. Hong and B.S. Sohn, "Segmentation of Regions of Interest in Mammograms in a Topographic Approach", *IEEE Trans. on Information Technology in Biomedicine*, Vol. 14, N°. 1, pp. 129-139, 2010.
- [IMT 2007] K. Idir, H. Merouani and Y. Tlili, "Dual Pyramid of Agents for Image Segmentation", *World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering* Vol: 1, N. 6, pp. 1809-1812, 2007.
- [JBB 2008] Jenhani, I., Ben Amor, N., Benferhat, S., & Elouedi, Z. (2008). SIM-PDT: A Similarity based Possibilistic Decision Tree approach. In *Proceedings of the 5th International Symposium on Foundation of Information and Knowledge Systems*, 348-364.
- [JBE 2007] I. Jenhani, N. Ben Amor, Z. Elouedi, S. Benferhat and K. Mellouli, « Information Affinity: a new similarity measure for possibilistic uncertain information », *ECSQARU'07, Hammamet, Tunisia*, 2007, 840-852.
- [JBE 2008] Jenhani, I., Ben Amor, N., & Elouedi, Z. (2008). Decision trees as possibilistic classifiers. *International Journal of Approximate Reasoning*, 48, 784-807.
- [JBE 2009] Jenhani, I., Benferhat, S., & Elouedi, Z. (2009). On the use of clustering in possibilistic decision tree induction. In *Proceedings of the 10th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, 505-517.
- [Jum 1993] Jumarie G., "A General Approach to Approximate Reasoning via Probability-Possibility Conversion", *Proc. 1993 International Conference on Systems, Man and Cybernetic*, vol. 4, pp. 656-661, 1993.
- [KAA 2018] Khanfir, I. K., Almouahed, S., Alsahwa, B. et Solaiman, B. (2018), « The use of contextual spatial knowledge for low-quality image segmentation », *Multimedia Tools and Applications*, , pp. 1-21, 10.1007/s11042-018-6540-1.
- [KAS 2018] I. K. Khanfir, S. Almouahed, B. Solaiman, and É. Bossé, "An iterative possibilistic knowledge diffusion approach for blind medical image segmentation," *Pattern Recognition*, vol. 78, pp. 182-197, 2018.
- [Kau 1973] Kaufmann, A. (1973), « Introduction à la théorie des sous-ensembles flous à l'usage des ingénieurs », Tome I, *Eléments théoriques de base*. Masson et Cie, éditeurs, Paris.
- [KBC 1994] F. R. Kara, P. Bolon and J. P. Cocquerez, "A region-region and region-edge cooperative approach of image segmentation", *Image Processing, Proceedings, IEEE Int Conf*, Vol. 3, pp. 470-474, 1994.
- [KBW 2009] A. M. Khuzi, R. B. BEng, Z. WMD Wan and N. A. MEngSc, "Identification of masses in digital mammogram using gray level co-occurrence matrices", *Biomedical Imaging and Intervention Journal*, Vol. 5, N. 3, pp.1-13, 2009.
- [KC 2002] C. Kermad and K. Chehdi, "Automatic Image Segmentation System Through Iterative Edge-Region Co-Operation", *Image and Vision Computing*, Vol. 20, N. 8, pp. 541-555, 2002.
- [KG 1991] Kaufmann, A. et Gupta, M.M. (1991), « Introduction to fuzzy arithmetic : Theory and applications », Van Nostrand, New York.
- [KGV 1983] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi: *Optimization by Simulated Annealing*, *Science*, vol. 220, no. 4598, pp. 671 -680, 1983.
- [KH 2007] J. Kozumplik and J. Hrubes: *A Fuzzy Approach to Thresholding of Image Data*, In *Proceedings of the 13th Conference Student EEICT*, pp. 256 (p), 2007.

- [KKB 2004] I. Khanfir, A. Kallel, M. S. Bouhleb, and K. Taouil, "Melanoma-pattern extraction using histogram-thresholding approach", MESM03, Middle Eastern Symposium on Simulation and Modelling, Sharjah, UAE, 2004.
- [Kli 1993] Klir, G.J. (1993), « Information-preserving probability-possibility transformations: recent developments », *Fuzzy Logic*: 417-428, 1993.
- [Kli 1999] Klir, G. J. (1999), « On fuzzy-set interpretation of possibility theory », *Fuzzy Sets and Systems*, vol. 108, no. 3, p. 263-273, déc. 1999.
- [KM 2005] Kaufman, K. A. et Michalski, R. S. (2005), "From Data Mining to Knowledge Mining", Chapter from "Handbook of Statistics" book, published by Elsevier, Vol. 24, pp. 47-75, 2005.
- [KMH 2005] N. Khalifa, A. Malek and K. Hamrouni: Segmentation d'images par contours actifs: Application à la détection du ventricule gauche dans les images de scintigraphie cardiaque, International Conférence: Sciences of Electronic (SETIT), 2005.
- [KMP 2000] Klement, E.P., Mesiar, R. et Pap, E. (2000)« Triangular Norms », Dordrecht, Kluwer Academic Publishers, 2000.
- [Kon 1994] I. Kononenko, Estimating Attributes : Analysis and Extensions of RELIEF, Proc. Of the European Conference on Machine Learning, ECML-94, 171-182, 1994
- [KP 2004] Kikuchi, S. et Perincherry, V. (2004), « Handling Uncertainty in Large Scale Systems with Certainty and Integrity ». MIT Engineering Systems Symposium, Cambridge, 2004.
- [Krou 2006] T. Kroupa, "Application of the Choquet integral to measures of information in possibility theory," *International journal of intelligent systems*, vol. 21, pp. 349-359, 2006.
- [KS 2017] Khanfir, I. K. et Solaiman, B. (2017), « Contextual Possibilistic Modeling Of Pixellic Knowledge For Tumor Segmentation In Mammographic Images », 3rd IEEE international conference on Advanced Technologies for Signal and Image Processing (ATSIP'2017), May 1, 22-24, 2017, Fez, Morocco.
- [KWT 1988] M. Kass, A. Witkin and D. Terzopoulos, "Snakes: Active contour models", *International Journal of computer Vision*, Vol. 1, N° 4, pp.321-331, 1988.
- [KY 1995] G. J. Klir and B. Yuan: *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, NJ. 1995.
- [Las 1999] Lasserre, V. (1999), « Modélisation floues des incertitudes de mesures de capteurs », Thèse présentée devant l'Université de Savoie, 1999.
- [Lec 2005] F. Lecellier, « Les contours actifs basées région pour la segmentation d'images: Application à la segmentation du contour de l'endocarde sur des images d'échographie cardiaque », STIC, pp. 420-423, Rennes-France, 2005.
- [Lev 1981] Levin, B. (1981), «A representation for multinomial cumulative distribution functions », *The Annals of Statistics* 9 (1981) 1123-1126.
- [LH 2007] G. Loum and T. C. Haba, "Caractérisation de textures à l'aide d'un codage directionnel local", *Afrique Science: Revue Internationale des Sciences et Technologie*, 2007.
- [Liu 2007] Liu, W. (2007), « Conflict analysis and merging operators selection in possibility theory », In Khaled Mellouli, editor, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 4724 of *Lecture Notes in Computer Science*, pages 816–827. Springer Berlin Heidelberg, 2007.
- [LKT 2018] Lajili, R., Kalti, K., Touil, A., Solaiman, B. et Essoukri Ben Amara , N. (2018), « Two-step evidential fusion approach for accurate breast region segmentation in mammograms », in *IET Image Processing*, vol. 12, no. 11, pp. 1972-1982, 11 2018. doi: 10.1049/iet-ipr.2018.5325
- [LWL 2001] H. Li, Y. Wang, K.J.R. Liu, S.B. Lo and M.T. Freedman, "Computerized radiographic mass detection C part I: lesion site selection by morphological enhancement and contextual segmentation", *IEEE Trans. Med. Imaging*, Vol. 20 (4) , pp.289–301,2001.
- [Mas 2004] S. Maskey: *Modelling Uncertainty In Flood Forecasting Systems*, Taylor & Francis Group plc, London, Uk, 2004.

- [MCB 2004] Mencar, C., Castellano, G., Bargiela, A. et Fanelli, A. (2004), « Similarity vs. possibility in measuring fuzzy sets distinguishability. », Proceedings of the Fifth International Conference on Recent Advances in Soft Computing (RASC), pp.354-359.
- [MD 2006] Masson, M. et Denoeux, T. (2006), « Inferring a possibility distribution from empirical data », Fuzzy Sets and Systems, vol.157(3), pp.319-340; 2006.
- [Mic 1998] Z. Michalewicz: Genetic Algorithms and Data Structures: Evolution Programs, Springer, 1998.
- [Min 2015] Assistant de Minitab Statistical Software, (2015) « Tests du Khi deux », Livre blanc, 2015, 2017 Minitab Inc.
- [MJ 1997] WL. May, WD. Johnson, «A SAS macro for constructing simultaneous confidence intervals for multinomial proportions. » Computer Methods and Programs in Biomedecine, 53 :153–162, 1997.
- [MKK 1998] S. Medasani, J. Kim, R. Krishnapuram, (1998), «An Overview of membership function generation technique for pattern recognition. » International Journal of Approximate Reasoning 19 pp. 391-417,1998.
- [MKK 2009] J. Mohanalin, P. K. Kalra and N. Kumar, “Tsallis Entropy Based Contrast Enhancement of Microcalcifications”, International Conference on Signal Acquisition and Processing, ICSAP pp. 3-7, 2009.
- [Mou 2002] MS. Mouchaweh, «Conception d’un système de diagnostic adaptatif et prédictif basé sur la méthode Fuzzy Pattern Matching pour la surveillance en ligne des systèmes évolutifs Application à la supervision et au diagnostic d’une ligne de peinture au trempé », Thèse de Doctorat, Université de Reims Champagne-Ardenne, France, 2002.
- [MRS 2008] A. Mencattini, G. Rabottino, M. Salmeri, R. Lojacono and E. Colini “Breast Mass Segmentation in Mammographic Images by an Effective Region Growing Algorithm”, Advanced Concepts for Intelligent Vision Systems, Springer Berlin Heidelberg, vol. 5259, pp. 948-957, 2008.
- [NB 2007] G. Nieradka and B. Butkiewicz: A Method for Automatic Membership Function: Estimation Based on Fuzzy Measures, IFSA, pp. 451-460, 2007.
- [NHR 2011] Z. Noumir, P. Honeine, C. Richard, Classification multi-classe au prix d’un classifieur binaire, GRETSI’11, Bordeaux, 5–8 September 2011.
- [NM 1997] M. Nagao and T. Matsuyama, “Edge Preserving Smoothing”, Computer Graphics and Image Processing, Vol. 9, pp. 394-407, 1979.
- [Non 1991] Nonaka, I. (1991), « The knowledge-creating company », Harvard business review. Nov-dec 1991.
- [NR 2005] Naumann, F., et Rolker, C. (2005), "Assessment Methods for Information Quality Criteria", Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Informatik, 2005.
- [NT 1997] Nonaka, I. et Takeuchi, H. (1997), « La connaissance créatrice : La dynamique de l’entreprise apprenante », Editions De Boeck. Note de lecture des étudiants MIP du Cnam.
- [OMB 2003] Oussalah, M., Maaref, H. et Barret, C. (2003), « From adaptive to progressive combination of possibility distributions », Fuzzy Sets and Systems; vol. 139, 2003, p. 559-582.
- [Ots 1978] N. Otsu, « A threshold selection method from gray-level histogram », IEEE Trans. Systems Man Cybernet. SMC-8, p. 62-66, 1978.
- [Ous 2000] Oussalah, M. (2000), « Study of some algebraical properties of adaptive combination rules », Fuzzy sets and systems, vol. 114, 2000, p. 391-409.
- [PBF 2014] Paggi, H., Bossé, É., Florea M. C., et Solaiman, B. (2014), "On the use of holonic agents in the design of information fusion systems," 17th International Conference on Information Fusion (FUSION), Salamanca, 2014, pp. 1-8.
- [PG 2009] Portet, F. et Gatt, A. (2009), « Towards a possibility-theoretic approach to uncertainty in medical data interpretation for text generation. In KR4HC, volume 5943 of Lecture Notes in Computer Science, pages 155–168. Springer, 2009.

- [PK 1993] Perincherry, V., Kikuchi, S. et Hamamatsu, Y. (1993), « Uncertainties in the analysis of large-scale systems », in , Second International Symposium on Uncertainty Modeling and Analysis, 1993. Proceedings, 1993, p. 216-222.
- [PR 1988] S. K. Pal and A. Rosenfeld: Image enhancement and thresholding by optimization of fuzzy compactness, Pattern Recogn. Lett, vol. 7, no. 2, pp. 77-86, 1988.
- [Pun 1981] T. Pun: Entropic thresholding: a new approach, Computer Graphics and Image Processing, vol. 16, no. 3, pp. 210-239, 1981.
- [RD 2018] Service de rédaction de la revue en ligne La Revue Data (2018), « Quelle différence entre donnée, information et connaissance ? », La Revue Data, mai 3, 2018. URL : <http://larevedata.com/donnee-information-connaissance>
- [RMJ 2002] M. Rukoz, M. Manouvrier and G. Jomier, “Distances de similarité d’images basées sur les arbres quaternaires”, 18èmes Journées Bases de Données Avancées BDA’02, Evry-France, 2002.
- [RNM 2003] Raufaste, E., Neves, R. D. S. et Mariné, C. (2003), « Testing the descriptive validity of possibility theory in human judgments of uncertainty », Artif. Intell., 148(1-2) :197–218, 2003.
- [Rob 2017] Robert, S. (2017), « Les mécanismes de l’intelligence artificielle et le raisonnement », ACFAS 2017, Université du Québec à Montréal, Canada.
- [Roh 2007] Rohmer, J. (2007), « La théorie des possibilités comme outil de représentation des incertitudes épistémiques d’une chaîne de traitement du risque sismique », AFPS 2007 - 7ème Colloque national, Jul 2007, Chatenay Malabry, France. 8 p., 2007.
- [Ros 2002] J. J. Rousselle, “Les contours actifs, une méthode de segmentation. Application à l’imagerie médicale”, PhD, Université François Rabelais, Tour, 2002.
- [SBP 2015] Solaiman, B., Bossé, É., Pigeon, L., Guériot, D. et Florea, M.C. (2015), « A conceptual definition of a holonic processing framework to support the design of information fusion systems », Information Fusion, Volume 21, January 2015, Pages 85-99.<https://doi.org/10.1016/j.inffus.2013.08.004>
- [SCC 1998] Sanguesa, R., Cabos, J., Cortes, U.: Possibilistic conditional independence: A similarity based measure and its application to causal network learning. International Journal of Approximate Reasoning, 18, 145-167(1998)
- [Sch 2006] Schmidt, C. (2006), « Deux prix Nobel pour la théorie des jeux », Revue d’économie politique, 2006/2 (Vol. 116), p. 133-145. DOI : 10.3917/redp.162.0133. URL : <https://www.cairn.info/revue-d-economie-politique-2006-2.htm-page-133.htm>
- [SCL 2017] J. H. Song, W. Cong, and J. Li, "A Fuzzy C-means Clustering Algorithm for Image Segmentation Using Nonlinear Weighted Local Information.", Journal of Information Hiding and Multimedia Signal Processing 8.9 (2017): 1-11.
- [SG 1995] Sison, C.P. et Glaz, J. (1995), « Simultaneous confidence intervals and sample size determination for multinomial proportions », J. Am. Stat. Assoc. 90 (1995) 366-369.
- [Sha 1961] Shackle G.L.S. (1961), « Decision, order and time in human affairs », Cambridge, Cambridge University Press, 1961.
- [Sha 1976] Shafer, G. (1976), « A mathematical theory of evidence », Princeton university press, Vol. 42, 1976.
- [SHC 2009] Shi, Y.F., He, L.H., Chen, J. (2009), « Fuzzy pattern recognition based on symmetric fuzzy relative entropy », Intelligent Systems and Applications, 1, 68-75
- [SIE 2017] Y. B. Salem, R. Idoudi, K. S. Ettaba, K. Hamrouni, and B. Solaiman, "Ontology Based Possibilistic Reasoning for Breast Cancer Aided Diagnosis," in European, Mediterranean, and Middle Eastern Conference on Information Systems, 2017, pp. 353-366.
- [SLS 2002] J.S. Suri, K. Liu, S. Singh, S.N. Laxminarayana X. Zang and L. Reden “Shape recovery algorithms using level sets for 2D and 3D medical imagery: A state of the art review”, IEEE Trans. Info, Tech. in Biomed, vol. 6, n. 1, pp. 8-28, 2002.

- [Sme 1990] Smets, P. (1990), « Constructing the pignistic probability function in a context of uncertainty », *Uncertainty in Artificial Intelligence 5* (M. Henrion, R.D. Shachter, L.N. Kanal, J.F. Lemmer, eds.), North-Holland, 29-39, 1990.
- [Smi 2002] L. I. Smith, A tutorial on principal component analysis, February 2002, Available: <http://csnet.otago.ac.nz/cosc453/studenttutorials/principalcomponents.pdf>.
- [Sol 2001] 53Solaiman, B. (2001), « Information fusion concepts. From information elements definition to the application of fusion approaches », *SPIE proceedings series*, vol. 4385, p. 205-212, avr. 2001.
- [Suc 1994] J Suckling et al (1994), « The Mammographic Image Analysis Society Digital Mammogram Database Exerpta Medica. », *International Congress Series 1069* pp375-378.
- [Tang 1998] X. Tang, "Texture information in run-length matrices", *IEEE Trans. on Image Processing*, Vol. 7, N° 11, pp. 1602-1609, 1998.
- [TB 1994] W. Tao and H. Burkhardt: An Effective Image Thresholding Method Using a Fuzzy Compactness Measure, in *Proceedings of the 12th ICIP*, vol. 1, pp. 47-51, 1994.
- [Tes 1993] Tessem, B. et al. (1993), « Approximations for efficient computation in the theory of evidence », *Artificial Intelligence*, 61(2) :315–329, 1993.
- [Til 1988] J. C. Tilton, "Image segmentation by iterative parallel region growing with applications to data compression and image analysis", *2nd Symposium on the Frontiers of Frontiers of Massively Parallel Computation*, Proceedings, pp. 357-360, 1988.
- [Til 1989] J. C. Tilton, "Image Segmentation by Iterative Parallel Region Growing And Splitting", *Geoscience and Remote Sensing Symposium, IGARSS 89. 12th Canadian Symposium on Remote Sensing*, Vol. 4, pp. 2420-2423, 1989.
- [TXO 2005] A. K. H. Tung, X. Xu, B. C. Ooi, *CURLER : Finding and Visualizing Nonlinear Correlated Clusters*, SIGMOD Conference, 2005.
- [TY 2011] Taroun, A. et Yang, J. (2011) « Dempster-Shafer Theory of Evidence: Potential usage for decision making and risk analysis in construction project management », vol. 4, no. 1, 2011.
- [UCI 2008] Index of Datasets UCI/arff, Available in: <http://repository.seasr.org/Datasets/UCI/arff/>, 2008.
- [Van 2003] Vannoorenberghe, P. (2003), « Belief functions applied to information processing », *Cépaduès, Toulouse, France*, vol. 3, no. 2, p. 9-45, 2003.
- [WH 1999] Y. Y. Wei and W. J.Hua, "Image segmentation based on region growing and edge detection", *IEEE, SMC, 99 Conference Proceedings*, Vol. 6, pp.798-803, 1999.
- [WK 1992] Z. wang and G.J. Klir: *Fuzzy Measure Theory*, Plenum Press, New York, 1992.
- [Yag 1983] R. R. Yager, "Entropy and specificity in a mathematical theory of evidence." *International Journal of General Systems*, 9, 1983, pp. 249-260.
- [Yag 1986] Yager, R. R. (1986), « A modification of the certainty measure to handle subnormal distributions », *Fuzzy Sets and Systems*, vol. 20, p. 317–324, nov. 1986.
- [Yag 1991] Yager, R. R. (1991), « Similarity based specificity measures », *International Journal of General Systems*, vol. 19, no. 2, p. 91, 1991.
- [YC 1997] P. Y. Yin and L. H. Chen, "A fast iterative scheme for multilevel thresholding methods", *Signal Process*, Vol. 60, N° 3, pp. 305-313, 1997.
- [YF 2016] Yager, R. R., et Fred P. (2016), « An intelligent quality-based approach to fusing multi-source probabilistic information. », *Information Fusion* 31 (2016): 127-136.
- [YS 1995] Yuan, Y., & Shaw, M. J. (1995). Induction of fuzzy decision trees. *Fuzzy Sets and Systems*, 69, 125-139.
- [Zac 1988] Zaccour, G. (1988). Valeur de Shapley et partage équitable des ressources. *L'Actualité économique*, 64(1), 96–121. <https://doi.org/10.7202/601438ar>
- [Zad 1965] Zadeh, L.A. (1965), « Fuzzy sets. *Information and Control* », 8: 338-353.

- [Zad 1978] Zadeh, L.A. (1978), « Fuzzy sets as a basis for a theory of possibility », *Fuzzy Sets and Systems* 1 :3-28
- [Zad 1979] Zadeh, L .A. (1979), « A theory of approximate reasoning », J.E. Hayes, D. Mitchie et L.I. Mikulich (éditeurs), *Machine Intelligence*, Vol. 9 (Wiley, New York): 149-194.
- [Zad 1989] Zadeh, L.A. (1989), « Knowledge representation in fuzzy logic », *IEEE Transactions on Knowledge and Data Engineering I* (1): 89-100.
- [Zah 2006] M. Zhao: *Image Thresholding Techniques Based on Fuzzy Partition and Entropy Maximization*, University of Sydney. School of Electrical and Information Engineering, 2006.
- [ZFK 2004] H. Zhang, S. W. Foo, S. M. Krishnan and C. H. Thng “Automated breast masses segmentation in digitized mammograms”, *IEEE. International Workshop on Biomedical Circuits and Systems*, pp. S2/2-S1-4, 2004.
- [ZJ 2004] Z. H. Zhou, Y. Jiang, NeC4.5: neural ensemble based C4.5, *IEEE Trans. Knowl. Data Eng.*16 (2004).
- [ZXZ 2018] W. Zhao, X. Xianze, Y. Zhu and F. Xu "Active contour model based on local and global Gaussian fitting energy for medical image segmentation", *Optik - International Journal for Light and Electron Optics*, 2018.

Résumé

Un processus automatique d'aide à la prise de décision ne peut opérer que sur une « projection » de la scène réelle, dans un espace de quantifications numériques d'observations ou de mesures physiques prises par des capteurs. Les informations obtenues de cette restriction du monde réel, ne peuvent en aucun cas être parfaites. Elles sont affectées d'incertitudes épistémiques. La théorie des possibilités permet d'aborder ce type d'incertitude, grâce à l'emploi d'un formalisme relativement simple, tout en révélant un comportement complexe. La capacité de cette théorie pour la représentation des connaissances et des informations hétérogènes et imparfaites, ainsi que les outils de raisonnement qu'elle fournit, constituent les atouts essentiels justifiant son choix.

Dans ce cadre, divers mécanismes de raisonnement possibiliste, sont développés pour aider à la modélisation des connaissances incomplètes, l'interprétation de scènes et l'aide à la prise de décision. Les expérimentations de nos différentes approches ont été conduites sur trois types de données : synthétiques de façon à intégrer des indicateurs contrôlant la qualité des données selon les besoins et les objectifs des tests, des benchmarks de façon à permettre la comparaison des performances des approches proposées par rapport à d'autres de l'état de l'art et enfin des données réelles collectées selon nos objectifs. Les résultats obtenus par l'application de nos approches proposées sur ces données, ont été très encourageants et ont montré l'efficacité de la théorie des possibilités comme un cadre de raisonnement sur des connaissances incertaines (pauvres en qualité et/ou en quantités, provenant de sources distinctes et/ou de natures distinctes). Il a été spécialement noté un comportement intéressant de stabilité et une efficacité supérieure dans des conditions de faibles connaissances.

Abstract

An automatic decision-making process can only operate on a "projection" of the real scene, in a space of numerical quantifications of observations or physical measurements taken by sensors. The information obtained from this restriction of the real world, in no case can be perfect. It is affected by epistemic uncertainties. The theory of possibilities makes it possible to approach this type of uncertainty, thanks to the use of a relatively simple formalism while revealing a complex behavior. The capacity of this theory for the representation of heterogeneous and imperfect knowledge and information, as well as the reasoning tools that it provides, constitute the essential assets justifying its choice.

In this framework, various mechanisms of possibilistic reasoning, are developed to help to a model of incomplete knowledge, scene interpretation and decision making. The experiments of our different approaches, were conducted on three types of data: synthetic, so as to integrate indicators controlling the quality of the data according to the needs and objectives of the tests, benchmarks, so as to allow the assessment of the proposed approach performances compared to others in the state of the art and finally real data collected according to our objectives. The results obtained by applying our proposed approach, have been very encouraging and have shown the effectiveness of the possibility theory as a reasoning framework on uncertain knowledge (poor in quality and/or in quantities, from various sources and/or distinct natures). It has been especially noted, an interesting behavior of stability and superior efficiency under conditions of poor knowledge.

ملخص

لا يمكن للعملية التلقائية لدعم اتخاذ القرار أن تعمل إلا على "إسقاط" للمشهد الحقيقي، في مساحة من الكميات العددية للملاحظات (الملاحظات) أو القياسات المادية (الفيزيائية) التي تم الحصول عليها باستخدام أجهزة الاستشعار. المعلومات التي تم الحصول عليها من هذا التقييد (أو التحديد) في العالم الحقيقي، لا يمكن أن تكون مثالية بأي حال من الأحوال. أنها تتأثر بالشكوك المعرفية. نظرية الاحتمالات تجعل من الممكن الوصول الى هذا النوع من عدم اليقين، وذلك بفضل استخدام شكليات بسيطة نسبياً، مع الكشف عن سلوك معقد. تشكل قدرة هذه النظرية على تمثيل المعرفة والمعلومات غير المتجانسة والكمال، وكذلك أدوات التفكير التي توفرها، الأصول الأساسية التي تبرر اختيارها.

في هذا الإطار، تم تطوير آليات مختلفة للتفكير التمكيني للمساعدة في نمذجة المعرفة غير المكتملة وتفسير المشهد ودعم اتخاذ القرار. أجريت تجارب باستخدام طرفنا (أو اسالينا) المختلفة على ثلاثة أنواع من البيانات: أولاً بيانات اصطناعية (من أجل دمج المؤشرات التي تتحكم في جودة البيانات وفقاً لاحتياجات وأهداف الاختبارات)، ثانياً بيانات المعايير (للسماح بمقارنة أداء الأسلوب (الطريقة) المقترحة مقارنة مع غيرها من التقنيات الحديثة) وأخيراً البيانات الحقيقية التي تم جمعها وفقاً لأهدافنا وغاياتنا. كانت النتائج التي تم الحصول عليها عن طريق تطبيق مناهجنا (طرفنا أو اسالينا) المقترحة على هذه البيانات مشجعة للغاية وأظهرت فعالية نظرية الاحتمالات كإطار تفكير بشأن المعرفة غير المؤكدة (رداءة بالجودة وكذلك بالكمية). وقد لوحظ سلوك مثير للاهتمام الا وهو الاستقرار والكفاءة العالية في ظل ظروف المعرفة الفقيرة.

**Mécanismes de raisonnement possibiliste pour
l'aide à la décision et l'interprétation de scènes**

Dr. Ing. Imène KHANFIR KALLEL
