

THÈSE

Présentée et soutenue publiquement par

Najla MEGHERBI

le 15 décembre 2006

pour obtenir le grade de

Docteur de l'Université des Sciences et Technologies de Lille

Discipline : Automatique et Informatique Industrielle



Association de données par la théorie des fonctions de croyance - application au suivi audio et vidéo des individus

François CABESTAING	Président	Professeur à l'USTL
Michèle ROMBAUT	Rapporteur	Professeur à l'Université de Grenoble
Daniel JOLLY	Rapporteur	Professeur à l'Université d'Artois
Patrick VANNOORENBERGHE	Examineur	Maitre de Conférences à l'UPS de Toulou
Marion BERBINEAU	Invitée	Directrice de Recherche à l'INRETS
Olivier COLOT	Directeur de thèse	Professeur à l'USTL
Sébastien AMBELLOUIS	Co-Directeur de thèse	Chargé de Recherche à l'INRETS

Remerciements

Je tiens tout d'abord à exprimer ma vive reconnaissance à Olivier COLOT, professeur à l'USTL, Sébastien AMBELLOUIS, chargé de recherche à l'INRETS et à François CABESTAING, professeur à l'USTL pour avoir dirigé ce travail. Je tiens à les remercier pour leurs conseils avisés et leurs critiques constructives qui m'ont guidé le long de cette thèse, pour nos nombreuses et longues discussions scientifiques que nous avons eues et pour leurs encouragements qui m'ont permis de mener à bien ce travail.

Je remercie en particulier François CABESTAING, qui m'a fait l'honneur de présider mon jury et pour la qualité de ses conseils et de ses critiques.

Je tiens à adresser ma plus vive gratitude aux membres de mon jury. Mme. Michèle ROMBAUT, professeur à l'Université de Grenoble et M. Daniel JOLLY, professeur à l'université d'Artois, pour avoir accepté de rapporter sur mon travail. Je les remercie pour le temps fourni pour l'évaluation de ma thèse. Mes remerciements vont également à Mme. Marion BERBINEAU, Directrice de Recherche à l'INRETS et M. Patrick VANNOORENBERGHE, Maître de Conférences à l'UPS de Toulouse, pour m'avoir fait l'honneur de faire partie du jury afin de juger mon travail.

J'adresse également mes remerciements à Mr. Marc HEDDEBAUT, Directeur de Recherche et Directeur du LEOST au moment de mon arrivée à l'INRETS, pour m'avoir accueilli dans le laboratoire.

Je remercie également tous les membres de l'équipe LEOST et toutes les personnes travaillant dans notre laboratoire, pour leur accueil et pour la bonne ambiance qui a permis un travail agréable.

Je remercie infiniment mon mari, Mahieddine. Son soutien apparaît derrière chaque mot, et chaque phrase de ce manuscrit. C'est lui qui a subi avec une patience les inconvénients que peut engendrer la réalisation d'une thèse et qui m'a donné la force nécessaire pour mener à bien cette expérience et pour tenir bon jusqu'à la fin.

Je remercie ma famille qui a su manifester son soutien et m'entourer d'affection malgré la distance géographique qui nous sépare. C'est à elle que je dédie l'effort continu de ces

dernières années. Cette thèse n'aurait jamais pu aboutir sans leur soutien affectif et leurs encouragements. Que ma mère, mon père, mes soeurs et mes frères trouvent ici l'expression de ma plus profonde gratitude.

Je remercie également ma belle famille qui a aussi contribué à la réalisation de ce travail par son aide et son soutien.

J'adresse également mes remerciements à mes amis de Biskra et de Lille pour leur soutien. Faire une liste exhaustive serait bien trop ardue, mais je me contenterai de citer : Leila, Maria, Raihane, Souad et Abida, amies avec lesquelles je garde de bons souvenirs.

Enfin, merci à tous ceux que je n'ai pas pu mentionner par manque d'espace, et qui ont contribué au travail présenté dans cette thèse.

Table des matières

Introduction	9
1 Filtrage / association dans le contexte du suivi multi-objet	13
1.1 Introduction	13
1.1.1 Quelques définitions	13
1.1.2 Description générale du suivi multi-objet	14
1.2 Modélisation du problème	15
1.3 Filtrage/association mono-objet en présence de mesures imprécises	17
1.3.1 Filtrage de Kalman	17
1.3.2 Filtrage de Kalman étendu (FKE)	18
1.3.3 Filtrage de Kalman "sans parfum"(FKU)	19
1.3.4 Conclusion	20
1.4 Filtrage/association en présence de mesures imprécises et incertaines	20
1.4.1 Association de données	21
1.4.2 Réduction de la complexité	23
1.4.3 Cas particulier du filtrage/association mono-objet	23
1.4.3.1 Filtrage bayésien optimal (FBO)	24
1.4.3.2 Filtrage à association de données probabilistes (PDAF)	25
1.4.4 Filtrage/association multi-objet	25
1.4.4.1 Filtre à association probabiliste jointe des données (JPDAF)	26
1.4.4.2 Le filtre à hypothèses multiples (MHT)	27
1.4.5 Conclusion	27
1.5 Conclusion	28
2 Association de données dans le cadre de la théorie des fonctions de croyance	31
2.1 Introduction	31
2.2 Notions de base de la théorie des fonctions de croyance (TFC)	32

2.2.1	Cadre de discernement et référentiel de définition	32
2.2.2	Fonction de croyance	32
2.2.3	Les Mesures de croyance	33
2.2.4	Combinaison des fonctions de croyance	36
2.2.5	Règles de combinaison conjonctive	36
2.2.6	Gestion de la masse conflictuelle	37
2.2.7	Le modèle des croyances transférables	37
2.2.8	Décision	38
2.3	Modélisation des fonctions de croyance	39
2.3.1	Les approches fondées sur le calcul d'une fonction de vraisemblance .	40
2.3.2	Les approches fondées sur le calcul d'une distance	41
2.3.2.1	Modélisation de Denoeux	41
2.3.2.2	Modélisation de Rombaut	42
2.3.2.3	Modélisation de Gruyer	42
2.4	Association crédibiliste de données pour le suivi multi-objet	43
2.4.1	La problématique	44
2.4.2	Construction du cadre de discernement	45
2.4.3	Génération des fonctions de croyance	45
2.4.4	Combinaison des fonctions de croyance	46
2.4.5	La phase de décision	47
2.4.6	Conclusion	49
2.5	Discussions	49
2.5.1	Normalisation de la combinaison des matrices locales d'association . .	49
2.5.2	L'hypothèse du "monde ouvert étendu" : des précautions à prendre . .	50
2.6	Conclusions	50
3	Relation entre le conflit et l'apparition et la disparition d'objets	53
3.1	L'étape d'association sous l'hypothèse du monde ouvert	54
3.1.1	Définition du cadre de discernement et des fonctions de croyance . . .	55
3.1.2	Combinaison des fonctions de croyance	56
3.1.3	Conflit et cadre de discernement non exhaustif	57
3.1.3.1	Les méthodes existantes	57
3.1.3.2	Notre méthode	59
3.1.4	Décision	61

3.1.5	Conclusion	62
3.2	Comparaison entre monde ouvert et monde ouvert étendu	63
3.2.1	Equation des fonctions de croyance combinées	63
3.2.1.1	Hypothèse du monde ouvert étendu (MOE)	63
3.2.1.2	Hypothèse du monde ouvert avec gestion partielle du conflit	64
3.2.2	Les fonctions de croyance initiales	65
3.2.3	Cas 1 : Contexte de décision identique pour les deux méthodes	66
3.2.3.1	Combinaison des jeux de masses	67
3.2.3.2	Discussion	69
3.2.4	Cas 2 : Les deux méthodes présentent des décisions différentes	69
3.2.4.1	Combinaison des jeux de masses	71
3.2.4.2	Discussion	71
3.3	Conclusions	75
4	Application au suivi audio et vidéo des individus	77
4.1	Le suivi audio et vidéo : un état de l'art	78
4.2	Le système proposé	80
4.2.1	Le capteur logique vidéo	82
4.2.1.1	Détection et suivi de visage : l'existant	83
4.2.1.2	L'attribut visuel	85
4.2.2	Le capteur logique audio	86
4.2.3	Localisation de sources sonores	87
4.2.4	Séparation des sources	88
4.2.5	Modélisation des sources	88
4.2.5.1	Etape de paramétrisation	89
4.2.5.2	Etape de classification	91
4.2.6	Le module d'association observations/cibles prédites	92
4.2.6.1	Les modifications apportées	93
4.2.6.2	Mesure de similarité acoustique	94
4.2.6.3	Mesure de similarité visuelle	96
4.2.6.4	Mesure de similarité de position	97
4.2.6.5	Calcul des fonctions de croyance élémentaires	97
4.2.7	Conclusion	98
4.3	Résultats	99

4.3.1	Cas du suivi d'individus proches	101
4.3.2	Cas du suivi d'individus perçu grâce à l'une ou l'autre des deux modalités	102
4.3.3	Cas de l'apparition des individus	108
4.3.4	Cas de l'apparition/disparition, simultanée et à des positions voisines	113
4.4	Conclusion	113
Conclusions et perspectives		117
4.5	Conclusion	117
4.6	Perspectives	118
 Annexes		 123
A Opérateur de gestion de conflit		123
A.1	Introduction	123
A.2	Les règles de combinaison définies dans l'algèbre booléenne	124
A.2.1	Les règles de combinaison normalisées	124
A.2.1.1	Règle de combinaison de Yager	125
A.2.1.2	Règle de combinaison de Shafer	125
A.2.1.3	Règle de combinaison disjonctif	126
A.2.1.4	Règle de combinaison de Dubois et Prade	126
A.2.1.5	Règle de combinaison de Lefevre	126
A.2.1.6	La règle de combinaison unifiée d'Inagaki	127
A.2.1.7	Les règles de combinaison utilisant une mesure d'intersection	127
A.2.2	Les règles de combinaison non normalisées	128
A.2.2.1	Règle de combinaison conjonctif et modèle des croyances transférables	128
A.2.3	Les règles de combinaison définies dans un treillis distributif	129
A.2.3.1	Le modèle libre DS _m	130
A.2.3.2	Le modèle hybride DS _m (Hybrid DS _m model)	131
A.3	Quelques remarques	132
A.4	Conclusion	132
 Bibliographie		 135

Table des figures

1.1	Description des étapes d'un système de suivi multi-objet.	15
1.2	Volume $V^i(k)$ centré autour de la mesure prédite $\hat{Y}_i(k k-1) = H_k^i \hat{X}^i(k k-1)$	24
1.3	Illustration du filtre à association probabiliste des données (PDAF).	26
1.4	Deux mesures sont validées pour les objets i et j	26
2.1	Les mesures de croyance.	35
3.1	Probabilités pignistiques obtenues sur le jeu de masses de Gruyer.	68
3.2	Probabilités pignistiques obtenues sur le jeu de masses de Rombaut.	68
3.3	Probabilités pignistiques obtenues sur le jeu de masses de Rombaut.	72
3.4	Probabilités pignistiques obtenues sur le jeu de masses de Gruyer.	73
4.1	Architecture proposée du système de suivi/association audio-vidéo.	81
4.2	Capteur logique vidéo.	83
4.3	Capteur logique Audio.	86
4.4	Structure d'un système d'identification du locuteur.	89
4.5	Fonction sigmoïde de normalisation : $D_{bh}^{seuil} = 150$, $r_1 = 20$ et $r_2 = 30$	95
4.6	Mesure de similarité de position entre P_1 et P_2	97
4.7	Extraction manuelle de visage.	100
4.8	Suivi de deux individus proches.	104
4.9	Suivi d'individus perçus grâce à l'une ou l'autre des deux modalités.	107
4.10	Apparitions des individus.	112
4.11	Apparition et disparition simultanées de deux individus.	115



Introduction

Contexte applicatif transport

Face à la croissance du sentiment d'insécurité des usagers des systèmes de transport en commun (voyageurs, agents d'exploitation, chauffeur etc.) et à l'augmentation des actes de vandalismes à l'égard des équipements, la dernière décennie a été marquée par la part de plus en plus importante de la vidéo-surveillance.

Ces systèmes sont destinés à observer un environnement afin de déceler et prendre en charge suffisamment rapidement des situations potentiellement critiques. La tâche de surveillance est assurée par un opérateur humain situé au poste de contrôle qui consulte un ensemble d'écrans de visualisation. Aujourd'hui, devant le nombre élevé de caméras vidéo déployées, cette tâche de surveillance est devenue particulièrement difficile pour un individu et ne permet plus d'assurer la meilleure qualité d'intervention en matière de rapidité et d'adéquation entre les moyens d'intervention engagés et l'incident.

Des recherches, visant à détecter certaines situations à partir de l'analyse automatique des images fournies par les systèmes de caméras, sont menées afin de faciliter cette tâche de contrôle et de la rendre plus confortable. Ces systèmes d'interprétation automatiques sont généralement couplés à des moyens de communication et à une logistique humaine importante. Une fois une situation décelée et l'alarme donnée, les images de la scène sont transmises au poste de contrôle pour validation. L'agent reste au centre du processus car lui seul donne son aval et décide de déclencher, ou non, la procédure de prise en charge la plus adaptée à la situation.

Actuellement, les systèmes de vidéo-surveillance sont déployés à grande échelle dans les espaces ouverts des transports publics (quais, couloirs, guichet...). Depuis deux ans, la vidéo-surveillance prend place dans les espaces embarqués (métro, trains, bus, tramway). Cette conquête des espaces et la volonté de mettre en place une politique de transport favorable au développement de la multi-modalité, nécessite que tous ces systèmes soient interconnectés et qu'un agent soit capable de retrouver un ou plusieurs individus donnés

sur l'ensemble du réseau.

Toutes ces fonctionnalités intègrent une procédure de suivi. Cette procédure vise à retrouver d'une image à l'autre, d'une caméra à une autre et d'un moyen de transport à un autre, certaines entités extraites des images. En présence d'occultations totales ou partielles des individus, les algorithmes de suivi présentent encore une robustesse insuffisante.

Dans ce mémoire nous présentons une méthode de suivi/association crédibiliste d'objets que nous plaçons dans le contexte d'un système multi-capteur. Nous considérons un système composé de microphones et de caméras. Cette perception multi-modale du monde permet de décrire les objets contenus dans la scène (des individus) par des informations complémentaires (attributs modaux) et redondantes (coordonnées des positions). A ce jour, aucune réponse ne peut être donnée à la question de la portabilité de ces systèmes dans le cadre des applications transport. De larges problèmes restent encore à résoudre liés, entre autres, aux contraintes liées aux environnements à surveiller. Toutefois, de tels systèmes bi-modaux de perception sont incontournables pour le développement d'environnements interactifs tels que ceux de la téléconférence. Les caractéristiques de ces milieux sont mieux contrôlés et offrent un cadre pertinent pour l'évaluation de la méthode de suivi proposée.

Problématique et solution proposée

Une procédure de suivi-association d'objets est équivalente à un problème de filtrage. Il s'agit d'estimer à chaque instant le processus d'état lié à chaque objet à suivre à partir des mesures fournies par un ensemble de capteurs. La procédure de suivi d'un objet se décompose en 3 étapes : l'association, la prédiction et l'estimation.

L'objectif de l'étape d'association est de mettre en correspondance les nouvelles mesures fournies par les capteurs aux objets connus par le système de suivi. Cette étape est particulièrement importante dès que l'on veut poursuivre plusieurs objets simultanément et lorsque les capteurs fournissent de fausses mesures. L'association peut être menée dans le cadre d'un système multi-capteur percevant ou non les mêmes parties de la scène. La combinaison de plusieurs sources d'informations permet d'assurer finalement une perception plus complète, plus précise, plus robuste et plus fiable de la scène que celle que le système aurait obtenu avec un seul capteur. Afin d'y parvenir, l'association doit tenir compte de la nature imparfaite des mesures. La combinaison des informations en présence de mesures manquantes (objets non perçus), de mesures incertaines (fausses mesures), de mesures

imprécises mène à des situations de conflit ou d'ambiguïté à lever.

Les méthodes les plus courantes décrites dans la littérature proposent de pallier ces problèmes en formulant l'association dans un cadre probabiliste ou dans celui offert par la théorie des fonctions de croyance. Dans notre travail, nous avons choisi d'exprimer le problème de l'association dans le cadre de la théorie des fonctions de croyance qui permet

- de gérer de manière indépendante l'incertitude et l'imprécision liées aux mesures,
- d'exprimer et de manipuler explicitement les informations manquantes grâce à la notion d'ignorance et
- de quantifier le conflit entre plusieurs mesures.

Dans cette théorie, les solutions de l'association sont contenues dans un ensemble appelé cadre de discernement. Les solutions retenues répondent à un critère de décision calculé sur la base du résultat de la combinaison d'un ensemble de fonctions de croyance. La valeur prise par une fonction de croyance reflète la confiance qu'un expert ou une source d'informations place dans une des hypothèses du cadre de discernement.

Traditionnellement, le cadre de discernement est considéré comme exhaustif, c'est-à-dire que toutes les solutions possibles au problème de l'association y sont représentées : l'une d'entre elle, appelée hypothèse de fermeture, représente alors l'ensemble des solutions inconnues. Ce monde est qualifié de "monde ouvert étendu" et la masse de croyance affectée à l'élément de fermeture permet de détecter des phénomènes d'apparition ou de disparition d'objets.

A l'instar de tous ces travaux, nous avons choisi d'étudier ce problème dans le cadre dit du "monde ouvert". Nous considérons que le cadre de discernement est non exhaustif et exprimons la combinaison des informations selon le Modèle des Croyances Transférables (TBM) proposé par Smets. La mise en évidence des phénomènes d'apparition et de disparition d'objets est assurée par une gestion spécifique du conflit engendré lors de la combinaison des informations.

Le mémoire est organisé en 4 parties. La première section est consacrée à la description des méthodes courantes de suivi/association des objets. Elle expose à la fois les méthodes bayésiennes classiques et les méthodes plus récentes fondées sur le filtrage particulière. La seconde partie présente les bases de la théorie des fonctions de croyance et décrit la formulation communément adoptée par la communauté scientifique afin de décrire la fonction d'association dans le cadre de cette théorie. Dans la troisième partie, nous modifions cette formulation en adoptant l'hypothèse du "monde ouvert" et les TBM. La masse conflictuelle

issue de la combinaison des fonctions de croyance est décomposée puis gérée de manière spécifique afin de détecter si les solutions se trouvent ou non dans le cadre de discernement. La décision finale est prise après une étape de raffinement adaptée. La dernière section est consacrée à la validation de la méthode proposée. Nous avons adopté un contexte multi-capteur regroupant des microphones et des caméras dans le but de suivre des individus. Chaque individu est caractérisé par un modèle audio (un vecteur de loi normale) et un modèle vidéo (un histogramme couleur). Les résultats ont été obtenus sur la base de mesures réelles (les histogrammes couleur et les modèles acoustiques) et de données synthétisées (les positions spatiales). Différents types de scénarii ont été simulés incluant le cas d'objets proches, le cas d'objets entrant ou sortant de la scène et le cas d'objets partiellement perçus par les capteurs.

Chapitre 1

Filtrage / association dans le contexte du suivi multi-objet

1.1 Introduction

1.1.1 Quelques définitions

Le suivi multi-objet, appelé aussi poursuite multi-objet, est le processus visant à déterminer, à chaque instant, le nombre d'objets dans une scène et leur état à partir des informations fournies par un ou plusieurs capteurs observant la scène. C'est une étape importante des systèmes de surveillance développés dans le cadre des applications de robotique mobile, des applications militaires et de surveillance d'activités.

Avant d'aller plus loin dans ce chapitre, il convient de définir plusieurs termes que nous utiliserons dans la suite du mémoire.

- Un capteur physique constitue le premier maillon de toute chaîne d'acquisition de données. C'est un dispositif de saisie d'informations qui permet de prélever une grandeur physique à mesurer et de la transformer en une grandeur exploitable par un système électronique. Deux grandeurs sont généralement associées au capteur : la fiabilité et l'imprécision.
- Une mesure est une information brute collectée à la sortie d'un capteur physique.
- La fiabilité et l'imprécision d'un capteur quantifient respectivement la confiance et le manque d'exactitude dans les mesures qu'il délivre.
- Une observation regroupe l'ensemble des mesures provenant des différents capteurs utilisés pour la perception d'un objet à un instant donné.
- Un attribut est une caractéristique extraite des mesures fournies par un capteur physique (ex. la chrominance est un attribut couleur calculé à partir des mesures des composantes rouge, verte et bleue).
- Une cible est la représentation symbolique d'un objet réel par le système de suivi à un instant donné.

- Une piste est la représentation temporelle d’une cible.
- Un capteur logique désigne tout capteur physique associé à un processus de traitement de données qui transforme un ensemble de mesures en une donnée dont la forme est plus adaptée à la tâche que le système doit accomplir. Le capteur peut être caractérisé par un indice de fiabilité ou une imprécision.

1.1.2 Description générale du suivi multi-objet

Le suivi multi-objet est un processus temporel visant à former l’ensemble des cibles représentant les objets présents dans une scène et à estimer leurs trajectoires à partir de leur observation par un ou plusieurs capteurs. La figure 1.1, inspirée de [L3], présente le processus général de suivi multi-objet. Il est composé de trois étapes : une étape d’association, une étape de prédiction et une étape d’estimation.

L’étape d’association de données consiste à mettre en relation l’ensemble des nouvelles mesures fournies par les différents capteurs à un instant t et l’ensemble des cibles dont l’état a été prédit pour le même instant à partir des informations disponibles à l’instant $t - 1$. Nous appelons ces cibles des *cibles prédites*. A l’issue de l’étape d’association, les cibles connues sont confirmées, les nouvelles cibles sont initialisées et les cibles disparues sont éliminées. Afin de réduire le nombre des mesures candidates à l’association, un processus de fenêtrage précède l’étape d’association. Il permet de définir, pour chaque cible prédite, l’ensemble des mesures qui sont susceptibles de lui être associées. L’étape d’estimation s’applique à chaque association *mesures/cible prédite* qui est retenue. Elle permet de mettre à jour l’état prédit de chaque cible en tenant compte des nouvelles mesures auxquelles la cible a été associée.

L’étape d’association est particulièrement importante dans le contexte multi-objet. En effet, rien ne nous indique l’origine des mesures fournies par chaque capteur. Une mesure peut caractériser un autre objet présent dans la scène ou encore être le résultat du mauvais fonctionnement d’un capteur. Dans ce dernier cas, la mesure n’est issue d’aucun objet présent dans la scène et elle est qualifiée de *fausse alarme*. Pour une mesure donnée, la véracité de son association à une cible prédite ou celle d’être une fausse détection est quantifiée par une fonction d’incertitude qui intervient dans les calculs de l’étape d’estimation. On parle alors de mesures incertaines.

Le suivi multi-objet peut être vu comme l’estimation stochastique de l’état réel de chaque objet à partir d’une séquence d’observations bruitées. Dans ce cas, chaque objet

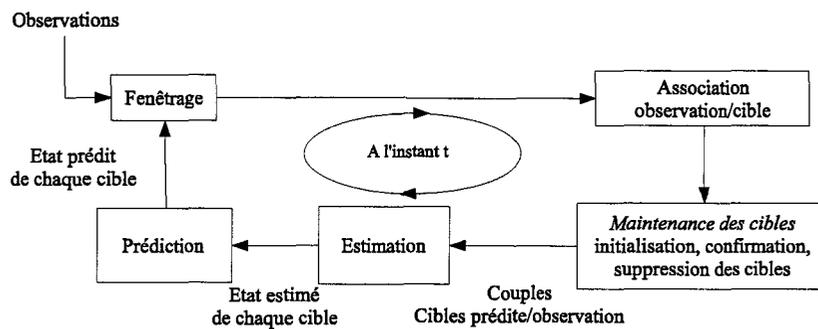


FIG. 1.1 : Description des étapes d'un système de suivi multi-objet.

est représenté par un système dynamique décrit dans un espace d'état. L'état de l'objet à estimer est alors représenté par un vecteur constitué de données cinématiques (la position, la vitesse et l'accélération) et de différents attributs. Le suivi est représenté sous la forme d'un processus de filtrage/association. De nombreuses méthodes ont été proposées dans la littérature selon que le système dynamique est considéré comme linéaire ou non ou selon que le système est en présence de mesures imprécises et/ou incertaines.

Dans la suite de cette section, nous commençons par présenter la modélisation générale du suivi multi-objet sous l'angle d'un processus de filtrage/association. Dans un second temps, nous exposons les méthodes qui permettent de résoudre ce problème dans le cadre particulier d'une application de suivi mono-objet. Dans un troisième temps, nous étendons ces méthodes au contexte du suivi multi-objet. En conclusion, nous présentons une méthode alternative d'association mesures/cible prédite développée dans le cadre de la théorie des fonctions de croyance. Cette méthode a la particularité de ne nécessiter aucun *a priori* sur les fonctions d'incertitude liées aux fausses détections et aux associations mesures/cibles prédites. Par ailleurs, elle peut être intégrée à une grande majorité des méthodes classiques de filtrage.

1.2 Modélisation du problème

Dans cette section, nous nous plaçons dans le cadre général d'un ensemble de mesures imprécises et incertaines. Bien que le système que nous décrivons soit à temps continu, il peut être défini de manière discrète pour chaque instant d'échantillonnage t tel que $t = k.T$ où T est la période d'échantillonnage et $k \in \mathbb{N}$. Dans la suite du mémoire, une variable ou une fonction indiquée par k signifie qu'elle est définie à l'instant d'échantillonnage $t = k.T$.

Nous parlerons de l'instant $k.T$.

Soit M_k le nombre de cibles connues à l'instant $k.T$. Soit $(X^i(k))$ le processus d'état du i^{me} objet suivi. Ce processus aléatoire est markovien et décrit par l'équation :

$$X^i(k) = F_k^i(X^i(k-1), V^i(k)) \quad \forall i = 1, \dots, M_k \quad (1.1)$$

où F_k^i est une fonction d'état non linéaire et $V^i(k)$ est un bruit blanc.

Soit $Y^j(k) = \left\{ Y_1^j(k), \dots, Y_{m_k^j}^j(k) \right\}$ l'ensemble des m_k^j mesures issues du capteur j à l'instant $k.T$. Comme cela a été précisé précédemment, ne connaissant pas l'origine des mesures $Y_l^j(k)$, une étape d'association est indispensable. Cette étape permet de définir de quel objet i provient chaque mesure $Y_l^j(k)$. Il est important de noter que le nombre des mesures m_k^j peut être différent du nombre d'objets suivis M_k .

Le processus $(X^i(k))_{k \in \mathbb{N}}$ est inconnu et correspond au processus aléatoire à estimer. Si, après la phase d'association, nous concluons que la mesure $Y_l^j(k)$ provient de la cible prédite i , alors la relation suivante est vérifiée :

$$Y_l^j(k) = H_k^i(X^i(k), W_l(k)) \quad (1.2)$$

où H_k^i est une fonction d'observation non linéaire et $W_l(k)$ est un bruit blanc indépendant de $V^i(k)$.

Estimer $X^i(k)$, $\forall i = 1, \dots, M_k$, revient à calculer puis à maximiser par rapport à $X^i(k)$ la probabilité *a posteriori* $P(X^i(k)|Y^j(1:k))$ où $Y^j(1:k) = Y^j(1), \dots, Y^j(k) \forall i = 1, \dots, M_k$. Lors de la phase d'association, deux hypothèses sont alors possibles :

1. (H1) A un instant donné, une mesure ne peut provenir que d'au plus un objet
2. (H2) A un instant donné, un objet peut donner naissance à plusieurs mesures

Dans le premier cas, les estimations de chaque $X^i(k)$ sont dépendantes les unes des autres. Dans le second cas, il est possible de calculer chaque estimée de manière indépendante.

Dans les sections suivantes, nous présentons un ensemble de méthodes déterministes permettant de résoudre le problème de filtrage/association modélisé dans cette section. Cette liste n'est pas exhaustive. Nous nous plaçons à la fois dans le cas général des systèmes d'état non linéaires et dans le cas particulier des systèmes linéaires. Dans un premier temps, nous décrivons les méthodes de filtrage/association en prenant soin de séparer

- le cas des mesures imprécises et
- celui des mesures imprécises et incertaines.

Le premier cas regroupe les méthodes suivi mono-objet en l'absence de fausses alarmes. Ces méthodes ne comportent pas d'étape d'association. Le deuxième cas regroupe les méthodes de filtrage/association mono-objet et multi-objet en présence de fausses alarmes. Dans ce dernier cas, l'incertitude de la mesure est liée au fait qu'elle peut être une fausse alarme et qu'elle peut provenir de tout objet présent dans la scène observée. Nous ne présentons que les méthodes qui supposent l'hypothèse (H1).

1.3 Filtrage/association mono-objet en présence de mesures imprécises

Le suivi d'un objet unique est un cas particulier du problème modélisé dans la section précédente. Dans ce cas, la variable M_k prend la valeur 1, le processus aléatoire d'état $(X^i(k))_{k \in \mathbb{N}}$ devient $(X(k))_{k \in \mathbb{N}}$ et l'ensemble $Y^j(k)$ des mesures fournies par le j^{eme} capteur est composé d'une seule mesure $Y_1^j(k)$ que nous noterons $Y^j(k)$ dans la suite de ce mémoire. Nous présentons un ensemble de méthodes qui regroupe le filtre de Kalman, le filtre de Kalman étendu et le filtre de Kalman dit "unscented" que nous traduirons pas filtre de Kalman "sans parfum". Ces méthodes font l'hypothèse que chaque mesure est imprécise mais provient de manière certaine de l'objet suivi; elles ne nécessitent donc pas d'étape d'association.

1.3.1 Filtrage de Kalman

Le filtre de Kalman [A38] a été développé pour le cas des systèmes linéaires en présence d'un bruit additif de type gaussien. Dans ce cas, $X(k)$ et $Y^j(k)$ sont reliés par les deux équations suivantes :

$$\begin{cases} X(k) &= F_k \cdot X(k-1) + G(k) \cdot u(k) + V(k) \\ Y^j(k) &= H_k \cdot X(k) + W(k) \end{cases} \quad (1.3)$$

où F_k et H_k sont respectivement les matrices d'état et d'observation du système, $V(k)$ et $W(k)$ sont deux bruits blancs gaussiens indépendants de variances respectives $\Sigma_{V(k)}$ et $\Sigma_{W(k)}$. Le vecteur et la matrice de commande du système sont désignés respectivement par u et G .

Dans ce cadre gaussien, on montre aisément que $X(k)$ est un processus gaussien. Par conséquent, deux paramètres suffisent à caractériser ce processus : la valeur moyenne $\hat{X}(k)$ et la covariance $P(k)$. La particularité du filtre de Kalman est de calculer de manière récursive l'estimation $\hat{X}(k)$ de $X(k)$ et celle de $P(k)$ en deux étapes : une étape de prédiction

puis une étape de correction. Ces étapes sont décrites respectivement par les systèmes d'équations 1.4 et 1.5. Durant l'étape de prédiction, le filtre calcule $\widehat{X}(k|k-1)$ et $P(k|k-1)$ à partir des estimées courantes $\widehat{X}(k-1)$ de X et $P(k-1)$ de la matrice de covariance. Durant l'étape de correction, la nouvelle estimée de X est déterminée à partir de la nouvelle mesure disponible $Y^j(k)$ et du paramètre K_k appelé gain de Kalman.

Etape de prédiction

$$\begin{cases} \widehat{X}(k|k-1) &= F_k \cdot \widehat{X}(k-1) + G \cdot u(k) \\ P(k|k-1) &= F_k \cdot P(k-1) \cdot F_k^T + \Sigma_{V(k)} \end{cases} \quad (1.4)$$

Etape de correction

$$\begin{cases} \widehat{X}(k) &= \widehat{X}(k|k-1) + K_k \left(Y^j(k) - H_k \cdot \widehat{X}(k|k-1) \right) \\ P(k) &= (I - K_k \cdot H_k) P(k|k-1) \\ K_k &= P(k|k-1) \cdot H_k^T \left(H_k \cdot P(k|k-1) H_k^T + \Sigma_{W(k)} \right)^{-1} \end{cases} \quad (1.5)$$

Une valeur K_k élevée signifie que la nouvelle mesure est privilégiée dans le calcul de l'estimée. A l'inverse, une valeur faible du gain signifie que la prédiction prend plus d'importance devant la mesure. Notons que la covariance prédite $P(k|k-1)$, la covariance corrigée $P(k)$ et le gain K_k ne dépendent pas des mesures. Par conséquent, ils peuvent être calculés avant même l'obtention de chaque mesure, ce qui permet de réduire les temps de calcul.

1.3.2 Filtrage de Kalman étendu (FKE)

Ce filtre a été proposé afin d'offrir une solution au problème des systèmes non linéaires en présence de bruits blancs gaussiens [L11]. Dans ce cas, les équations 1.1 et 1.2 s'expriment au travers du système suivant :

$$\begin{cases} X(k) &= F_k(X(k-1)) + V(k) \\ Y^j(k) &= H_k(X(k)) + W(k) \end{cases} \quad (1.6)$$

où F_k et H_k sont respectivement deux fonctions dérivables et $V(k)$ et $W(k)$ sont deux bruits blancs gaussiens indépendants de variances respectives $\Sigma_{V(k)}$ et $\Sigma_{W(k)}$.

Le filtre de Kalman étendu repose sur la linéarisation des fonctions F_k et H_k . Dans le cadre d'une linéarisation au premier ordre, ces fonctions s'approximent de la manière suivante :

$$\begin{cases} F_k(X) \approx f_k \cdot X + C_k \\ H_k(X) \approx h_k \cdot X + D_k \end{cases} \quad (1.7)$$

et le système 1.6 devient :

$$\begin{cases} X(k) = f_k \cdot X(k-1) + C_k + V(k) \\ Y^j(k) = h_k \cdot X(k) + D_k + W(k) \end{cases} \quad (1.8)$$

où

- f_k et h_k sont respectivement les dérivées premières de F_k et de H_k en $\hat{X}(k-1)$ et $\hat{X}(k)$
- C_k et D_k sont obtenues à partir de l'expression des fonctions 1.7

Les étapes de prédiction et de correction décrites dans la section précédente sont alors applicables au nouveau système ainsi obtenu. Plus la linéarisation est fine, meilleure sera l'estimation de X . Le filtre de Kalman étendu présente de bonnes performances lorsque les fonctions présentent de légères non-linéarités. La qualité de l'estimation peut être mesurée à partir de l'analyse de la trace des matrices $P(k|k-1)$ et $P(k)$ qui quantifient l'écart entre $X(k)$ et respectivement sa prédiction $\hat{X}(k|k-1)$ et son estimée $\hat{X}(k)$.

1.3.3 Filtrage de Kalman "sans parfum"(FKU)

Ce filtre a été proposé par Julier et Uhlmann [A36]. C'est une adaptation du filtre de Kalman au cas des systèmes non linéaires. Comme pour les deux filtres précédents, le processus X à estimer est supposé gaussien. La différence avec le filtre de Kalman étendu est qu'aucune linéarisation des fonctions F_k et H_k n'est appliquée.

Le filtre de Kalman "sans parfum" (FKU) propose d'approximer la loi gaussienne relative au processus X par un nuage de points, appelés *sigma points*, qui évolue de manière récursive selon une étape de prédiction et une étape de correction. Au cours de ces deux étapes, les deux fonctions non-linéaires sont appliquées aux sigma points sans linéarisation.

Au cours de l'étape de prédiction, l'état de chaque sigma point est prédit selon le modèle d'état non linéaire. Les prédictions $\hat{X}(k|k-1)$ et $P(k|k-1)$ sont alors calculées à partir de l'ensemble des prédictions de chaque sigma point. Par ailleurs, à chaque sigma point prédit est associée une mesure calculée sur la base du modèle d'observation non linéaire.

Durant l'étape de correction, le gain K_k , calculé à partir des matrices de variance-covariance des mesures et des sigma points prédits est appliqué afin de déterminer les

estimées $\hat{X}(k)$ et $P(k)$.

Cette démarche permet d'être plus précis dans l'estimation des deux premiers moments du processus aléatoire X . Elle présente de meilleures performances que la méthode par filtrage de Kalman étendu.

1.3.4 Conclusion

Dans cette partie nous avons présenté les méthodes déterministes d'estimation bayésienne de l'état d'un objet à suivre à partir d'un processus aléatoire de mesures dans le cas des systèmes linéaires et non-linéaires. Ces méthodes ont pour objectif de fournir, d'estimer et de maximiser la probabilité conditionnelle liant les deux processus précédents. Ces méthodes ont montré de bon comportements dans le cadre des systèmes linéaires ou très faiblement non-linéaires. Toutefois, leurs performances chutent lorsque les probabilités conditionnelles à estimer sont multimodales.

Afin de résoudre ces problèmes, des méthodes non déterministes ont été décrites dans la littérature ces dix dernières années, telles que les méthodes de filtrage dites de Monte-Carlo, encore appelées filtres particulaires. Ces méthodes permettent de représenter la loi conditionnelle recherchée par une somme finie et pondérée de mesures de Dirac concentrées en certains points de l'espace d'état appelé "particules". Une stratégie d'évolution et de sélection de ces particules permet de ne retenir que les meilleures au sens de la vraisemblance des mesures et finalement de converger vers la solution recherchée. Le lecteur pourra se reporter à la lecture de Carine Hue [L9] qui propose une étude détaillée de ces méthodes.

1.4 Filtrage/association multi-objet en présence de mesures imprécises et incertaines

Dans cette section, nous présentons les méthodes de suivi multi-objet dans le contexte de mesures imprécises et incertaines. Comme cela a été dit dans la section 1.2, dans les environnements bruités et multi-objet, rien ne nous informe sur l'objet qui est à l'origine de chaque mesure et rien ne nous indique si une mesure est une fausse alarme ou non. Or, estimer précisément l'état de chaque cible prédite à partir d'une des méthodes de filtrage présentées dans la section 1.3 nécessite de leur associer avec certitude une mesure parmi celles disponibles. Une mauvaise association entre une mesure et une cible prédite engendre une divergence de la procédure de filtrage et une erreur dans l'estimation de

l'état. A défaut d'avoir la certitude qu'une cible prédite est à l'origine ou non d'une mesure, il s'avère donc nécessaire de faire des hypothèses sur les associations possibles entre la cible prédite et les mesures et de quantifier la véracité de ces associations.

Dans la suite de cette section, nous commencerons par présenter quelques méthodes d'association de données couramment utilisées dans les techniques de filtrage/association. Nous compléterons cette présentation par l'exposé d'un principe de fenêtrage permettant de réduire le nombre d'hypothèses et donc le temps de traitement. Ensuite, nous présenterons des méthodes de filtrage/association adaptées au contexte particulier du suivi mono-objet en présence de fausses alarmes uniquement : le filtre bayésien optimal et le filtre à association de données probabilistes. Nous poursuivrons par la description des méthodes dans le contexte d'une incertitude des mesures liées au contexte multi-objet : le filtre joint à association de données probabilistes et le filtre à hypothèses multiples. Nous ne détaillerons aucune des méthodes en question. Nous invitons le lecteur à se rapporter à la lecture des travaux de Carine Hue [L9] et de Karim Dahia [L7].

1.4.1 Association de données

En se plaçant dans le cadre du système décrit par les équations 1.1 et 1.2 et en reprenant les notations présentées dans la section 1.2, rappelons qu'à l'instant $k.T$:

- $Y^j(k) = (Y_l^j(k))_{l=1, \dots, m_k^j}$ est l'ensemble des m_k^j mesures du capteur j ,
- M_k est le nombre d'objets connus,
- $X^i(k)$ est le processus d'état de l'objet i et $i = 1, \dots, M_k$.

Afin de simplifier les formalismes des présentations qui suivent, nous nous plaçons dans le cas d'un système composé d'un seul capteur ($j = 1$). Nous notons $Y(k) = Y^1(k) = (Y_l(k))_{l=1, \dots, m_k}$ et $m_k = m_k^1$ le nombre de mesures disponibles à l'instant $k.T$.

Comme nous l'avons spécifié à la fin de la section 1.2, nous émettons l'hypothèse (H1) qu'un objet donné peut être à l'origine d'aucune ou d'au plus une mesure. L'étape d'association consiste à calculer la véracité de chaque hypothèse d'association de chaque mesure $Y_l(k)$ à chaque objet connu i à l'instant $k.T$. Nous notons $(\omega_i^l)_{l=1, \dots, m_k}$ l'ensemble des hypothèses d'association des m_k mesures à l'objet i .

Dans la littérature, de nombreuses méthodes ont été proposées pour résoudre le problème d'association. Les méthodes non probabilistes, telles que la méthode du plus proche voisin (NNF) et la méthode du plus proche voisin généralisée (GNNF) [C94], cherchent à retenir, pour chaque objet i , une seule des hypothèses de $(\omega_i^l)_{l=1, \dots, m_k}$. Ainsi, une seule me-

sure est finalement utilisée dans la phase d'estimation de l'état de chaque objet i . Le NNF suppose qu'une mesure peut être associée à plusieurs objets. Les mesures non associées sont considérées comme de fausses alarmes et sont supprimées. Cette méthode ne permet donc pas de créer de nouveaux objets. Le GNNF quant à lui fait l'hypothèse qu'une mesure ne peut provenir que d'un seul objet. Les mesures non associées sont alors utilisées pour créer de nouveaux objets. Le NNF et le GNNF présentent de mauvaises performances en présence d'un grand nombre de fausses détections ou en présence d'objets dont les états sont proches [A9, L3].

A l'inverse, les méthodes probabilistes retiennent plusieurs hypothèses dont ils évaluent la probabilité d'occurrence. Finalement, les mesures $Y_l(k)$ relatives à chaque hypothèse ω_i^l interviennent dans le calcul de l'estimation de l'état $\widehat{X}^i(k)$ de chaque objet i , proportionnellement à la probabilité d'occurrence de ω_i^l . Dans la suite, nous notons :

- $\beta_l^i(k)$ la probabilité d'occurrence de l'hypothèse ω_i^l à l'instant $k.T$ pour $l = 1, \dots, m_k$ et $i = 1, \dots, M_k$ et
- $\beta_l^0(k)$ la probabilité que la mesure $Y_l(k)$ est une fausse alarme.

Dans la suite de cette section, nous présentons uniquement les techniques de filtrage/association fondées sur une modélisation probabiliste de l'occurrence des associations. Les méthodes que nous décrivons supposent que les fonctions de probabilité $\beta_l^i(k)$ et $\beta_l^0(k)$ sont connues et suivent respectivement une loi normale et une loi de Poisson.

Elles se décomposent en quatre étapes successives : une étape d'initialisation, une étape de prédiction, une étape d'association et une étape de correction. Pour toutes, nous faisons l'hypothèse d'un modèle d'état et de mesure linéaire avec des bruits additifs gaussiens, lequel est décrit par le système d'équations suivant :

$$\begin{cases} X^i(k) &= F_k^i X^i(k-1) + V^i(k) \\ Y_l(k) &= H_k^i X^i(k) + W(k) \end{cases} \quad (1.9)$$

où F_k^i et H_k^i sont respectivement les matrices d'état et d'observation du système, $V^i(k)$ et $W(k)$ sont deux bruits blancs gaussiens indépendants de variances respectives $\Sigma_{V^i(k)}$ et $\Sigma_{W(k)}$. Ces méthodes supposent que la loi d'évolution du processus d'état est une gaussienne.

Durant l'étape de prédiction, l'état prédit $\widehat{X}^i(k|k-1)$ et la matrice de covariance associée $P^i(k|k-1)$ sont calculés pour chaque objet i . Ces données sont utilisées durant l'étape d'association afin de calculer les probabilités des associations de chaque objet avec chaque

mesure. Finalement, l'étape d'estimation procède au calcul de $\widehat{X}^i(k)$, l'estimée de l'état de l'objet i , en tenant compte de la probabilité de chaque association.

1.4.2 Réduction de la complexité

La complexité des méthodes d'association croît avec le nombre m_k de mesures et le nombre M_k d'objets à suivre. Afin de réduire l'influence de l'un et de l'autre, un processus de fenêtrage statistique est appliqué aux mesures engagées au début de l'étape d'association afin de ne valider qu'une partie des associations possibles. Ce processus vise à définir, pour chaque objet i , l'ensemble des mesures qui sont susceptibles de provenir de lui.

Ce processus est décrit dans l'espace des observations. Il est fondé sur l'évaluation de la distance de Mahalanobis entre chaque mesure $Y_l(k)$ et la mesure prédite $\widehat{Y}_i(k|k-1) = H_k^i \cdot \widehat{X}^i(k|k-1)$ obtenue à partir de la prédiction de l'état de l'objet i . Cette distance est donnée par l'équation suivante :

$$d_l^i(k) = (Y_l(k) - \widehat{Y}_i(k|k-1))^T S^{-1}(k) (Y_l(k) - \widehat{Y}_i(k|k-1)). \quad (1.10)$$

Dans cette équation, $S(k) = H_k^i \cdot P(k|k-1) \cdot (H_k^i)^{-1} + \Sigma_W$ est la covariance de l'innovation $Y_l(k) - \widehat{Y}_i(k|k-1)$. Une mesure $Y_l(k)$ est validée si $d_l^i(k)$ est supérieure à un seuil noté γ , c'est-à-dire si la mesure $Y_l(k)$ est comprise dans le volume $V^i(k)$ défini par :

$$V^i(k) = \left\{ Y / (Y - \widehat{Y}_i(k|k-1))^T S^{-1}(k) (Y - \widehat{Y}_i(k|k-1)) \leq \gamma \right\}. \quad (1.11)$$

La valeur de γ est fixée de telle sorte que la probabilité d'avoir la bonne mesure à associer à l'objet i dans le volume $V^i(k)$ soit grande. Traditionnellement cette probabilité est fixée supérieure à 0.99. La figure 1.2 illustre un tel volume. Sur cette figure, le volume $V^i(k)$ est centré autour de la mesure prédite $\widehat{Y}_i(k|k-1) = H_k^i \cdot \widehat{X}^i(k|k-1)$ de l'objet i . Nous notons \tilde{m}_k^i le nombre de mesures validées pour l'objet i à l'instant $k.T$.

1.4.3 Cas particulier du filtrage/association mono-objet

Dans ce cas particulier, le nombre d'objets M_k à suivre est connu et égal à 1. On note $X^1(k)$ l'état de cet unique objet à l'instant $k.T$ et $\widehat{X}^1(k)$ l'estimée de cet état. Rappelons que résoudre ce problème d'estimation est équivalent à maximiser la probabilité conditionnelle $P(X^1(k)|Y(1:k))$ relativement à X^1 . Après la phase de validation, le nombre de mesures validées pour l'objet considéré est \tilde{m}_k^1 .

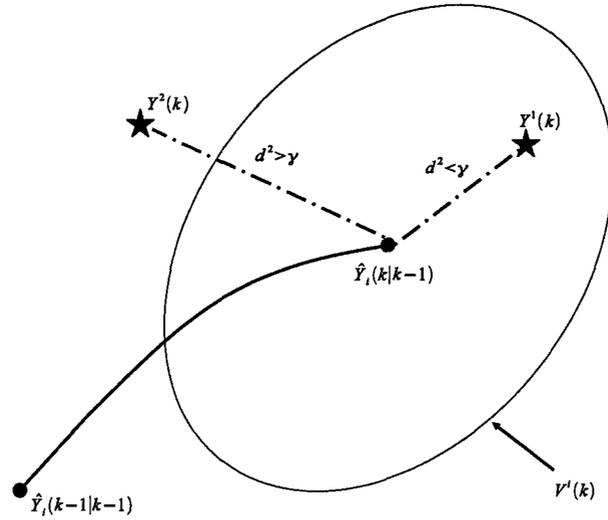


FIG. 1.2 : Volume $V^i(k)$ centré autour de la mesure prédite $\hat{Y}_i(k|k-1) = H_k^i \cdot \hat{X}^i(k|k-1)$.

1.4.3.1 Filtrage bayésien optimal (FBO)

Le FBO suppose que toutes les mesures validées entre ces deux instants peuvent provenir de l'objet. Il propose que toutes les mesures validées entre les instants 1 et $k.T$ pour l'objet à suivre soient prises en compte dans l'estimation de son état $\hat{X}^1(k)$ [A13].

Sous l'hypothèse qu'à chaque instant une seule mesure peut être associée à l'objet, le FBO exploite l'idée qu'une succession de k mesures prises à chaque instant compris entre 1 et $k.T$ permet d'estimer $X^1(k)$. Nous appelons k *successions* cette succession de k mesures. Une k *successions* est l'ensemble des k mesures $\{Y_{l_1}, \dots, Y_{l_k}\}$ tel que Y_{l_j} soit une mesure valide prise à l'instant $j.T$.

Il est important de noter qu'à tout instant, toutes les mesures peuvent être considérées comme des fausses alarmes. Par conséquent, une ou plusieurs mesures peuvent manquer à une k *successions*. Dans ce cas, le cardinal de cet ensemble est inférieur à k . Le nombre $\mu_{1:k}$ de k *successions* est donné par $\mu_{1:k} = \prod_{i=1}^k (1 + \tilde{m}_i^1)$.

Sur la base d'une k *successions*, et dans le cadre d'un processus gaussien et d'un système linéaire, la situation est identique à celle décrite dans la section 1.3 et $\hat{X}^1(k)$ peut être calculée par un filtre de Kalman classique. Un filtre de kalman étant initialisé pour chaque k *successions*, $\mu_{1:k}$ estimation de l'état $X^1(k)$ sont calculées. Si nous notons $\hat{X}_i^1(k)$ l'estimation donnée par la i^{me} k *successions*, alors l'estimation finale $\hat{X}^1(k)$ est donnée par la relation :

$$\hat{X}^1(k) = \sum_{i=1}^{\mu_{1:k}} \hat{X}_i^1(k) \beta_i(k) \quad (1.12)$$

avec $\sum_{i=1}^{\mu_{1:k}} \beta_i(k) = 1$. La fonction $\beta_i(k)$ est la probabilité d'occurrence de la i^{me} k *successions*. Différents modèles ont été proposés dans la littérature pour calculer les $\beta_i(k)$. En général, ils sont fondés sur des lois gaussiennes ou de Poisson.

Ce filtre est en théorie optimal. Toutefois, il présente l'inconvénient d'un croissance exponentielle du nombre de k *successions* ce qui réduit l'efficacité de son implantation et empêche son utilisation pratique.

1.4.3.2 Filtrage à association de données probabilistes (PDAF)

Le PDAF est une approximation du FBO présenté précédemment. Son fonctionnement est illustré à la figure 1.3. Contrairement au filtre précédent, seules les mesures validées à l'instant k pour l'objet à suivre sont utilisées dans le calcul de $\hat{X}^1(k)$ [A13]. Il est une extension du filtre de Kalman dans lequel la phase de correction est remplacée par une combinaison d'estimations pondérées par les probabilités d'occurrence de certaines associations.

Les k *successions* deviennent des singletons composés d'une des \tilde{m}_k^1 mesures validées à l'instant k ou encore d'aucune mesure dans le cas où toutes les mesures sont considérées comme des fausses alarmes.

Le nombre de filtres de Kalman initialisés est sensiblement réduit et égal à $1 + \tilde{m}_k^1$. L'estimation finale $\hat{X}^1(k)$ est donnée par la relation :

$$\hat{X}^1(k) = \sum_{i=1}^{1+\tilde{m}_k^1} \hat{X}_i^1(k) \beta_i(k) \quad (1.13)$$

avec $\sum_{i=1}^{\mu_{1:k}} \beta_i(k) = 1$. La fonction $\beta_i(k)$ est modélisée de la même manière que dans la section précédente.

1.4.4 Filtrage/association multi-objet

Dans le contexte d'un suivi multi-objet, une mesure peut être validée pour plusieurs objets i . Il s'agit d'une situation de conflit. La figure 1.4 en est une illustration. Sous l'hypothèse (H1) régler le conflit consiste à déterminer l'unique objet i qui a donné naissance à cette mesure. Le *filtre à association probabiliste jointe des données (JPDAF)* et le *filtre à hypothèses multiples (HMT)* décrits par la suite sont une solution à ce problème.

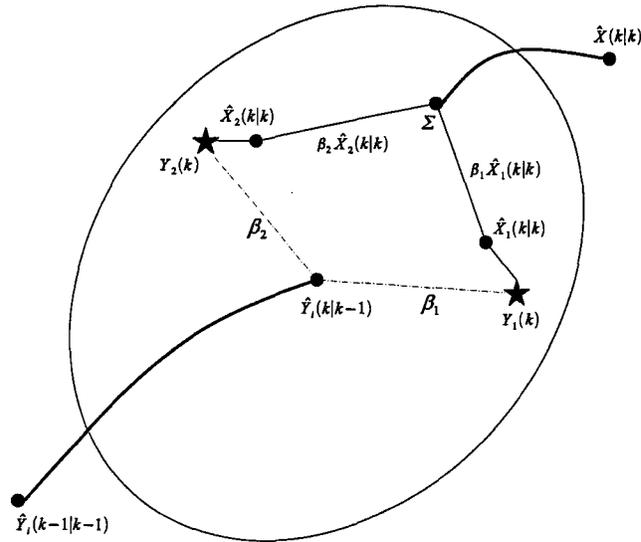


FIG. 1.3 : Illustration du filtre à association probabiliste des données (PDAF).

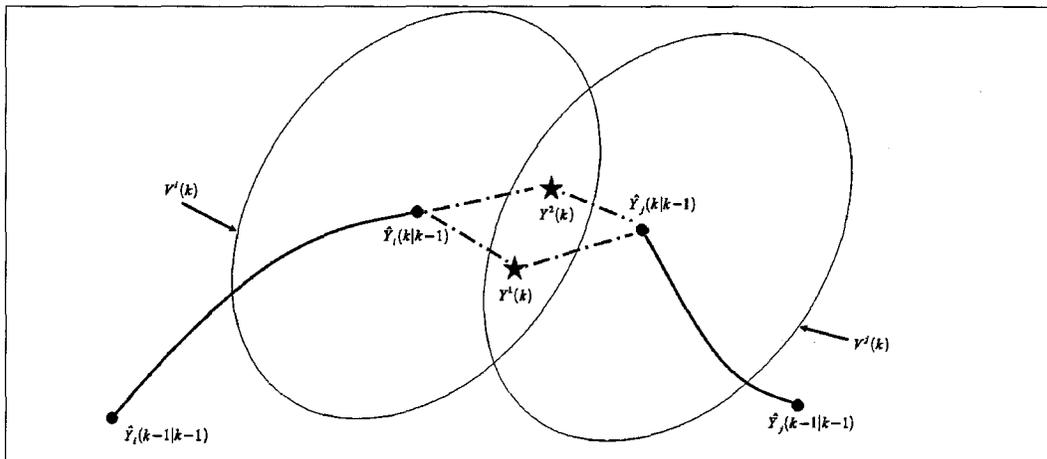


FIG. 1.4 : Deux mesures sont validées pour les objets i et j .

1.4.4.1 Filtre à association probabiliste jointe des données (JPDAF)

On suppose que le nombre d'objets M_k est supérieur à 1, constant et connu.

Le JPDAF est l'extension naturelle du PDAF dans un contexte multi-objet [C6, C7]. Il procède de la même manière en initialisant un PDAF par objet et en instaurant un processus d'exclusion. Sans cette procédure, tous les filtres convergent finalement vers la même solution. L'exclusion est garantie en calculant les probabilités des associations sur la base d'une énumération complète de ces dernières.

Comme les deux méthodes précédentes, le filtre JPDAF ne permet pas d'initialiser des nouvelles cibles. Le filtre à hypothèses multiples (MHT) décrit par la suite permet de ré-

soudre ce problème.

1.4.4.2 Le filtre à hypothèses multiples (MHT)

Ce filtre a été décrit pour la première fois dans [A65, C5]. On suppose que le nombre d'objets M_k est supérieur à 1, mais inconnu et non constant.

Le filtre MHT est en fait une extension du filtre FBO. Sur la base des \tilde{m}_k^i mesures validées pour chaque objet i , le filtre MHT génère l'ensemble des k *successions* qu'il représente sous la forme d'un arbre d'hypothèses. A chaque nouvelle mesure validée, le MHT génère un jeu d'hypothèses d'association qu'il intègre à cet arbre.

Le maintien d'un nombre important d'hypothèses engendre le risque d'avoir une explosion combinatoire exponentielle du nombre d'hypothèses. Son utilisation pratique nécessite donc de mettre en place des méthodes permettant de contrôler le nombre d'hypothèses générées. Ces méthodes incluent le regroupement (clustering) d'hypothèses, la suppression (pruning) des hypothèses et la fusion d'hypothèses (merging) [A9]. Si élaguer cet arbre permet de réduire la complexité algorithmique, il faut garder à l'esprit qu'une hypothèse peut s'avérer à nouveau utile après plusieurs instants d'échantillonnage. Il ne faut donc pas procéder trop vite à son élimination.

D'autres chercheurs ont essayé de contourner ce problème en introduisant le filtre à hypothèses multiples orienté cibles [L3, C48], qui ne maintient pas les hypothèses d'une étape à l'autre. Grâce à ce filtre, il est possible d'initialiser de nouvelles cibles, de les mettre à jour et de les ordonner avant de générer de nouvelles hypothèses.

1.4.5 Conclusion

Dans cette section nous avons présenté des méthodes de filtrage/association tenant compte de l'imprécision et de l'incertitude sur l'origine des mesures. L'incertitude sur l'origine d'une mesure signifie qu'elle peut être une fausse alarme ou que plusieurs objets sont susceptibles de l'avoir créée. Dans le contexte mono-objet, où seules les fausses alarmes interviennent dans la valeur de cette incertitude, les filtres FBO et PDAF offrent une solution bayésienne intéressante. Dans le contexte multi-objet, les FBO et le PDAF se déclinent sous la forme du JPDAF et du MHT afin de traiter le problème de l'incertitude liée à des situations de conflit difficiles à résoudre.

Récemment, des méthodes dites particulières ont été proposées dans la littérature afin de résoudre le problème de l'incertitude liée à la présence de plusieurs objets. Ces mé-

thodes proposent d'approximer la probabilité conditionnelle de l'état de chaque objet i par un nuage de particules donc l'évolution est régie par certaines lois statistiques. Dans ces méthodes, le passé est pris en compte grâce à l'état des particules du nuage. Ceci évite d'énumérer une grande partie des hypothèses d'association et de limiter l'explosion combinatoire du problème liée au nombre d'objets et au nombre de mesures. Nous conseillons la lecture de Carine Hue [L9, C6] pour une description détaillée de ces méthodes.

1.5 Conclusion

Dans ce chapitre nous avons présenté de manière non exhaustive les méthodes de filtrage/association permettant de résoudre le problème d'estimation d'état dans le contexte de mesures imprécises puis en présence de mesures imprécises et incertaines.

Afin de résoudre les problèmes liés à l'imprécision, nous avons décrit les méthodes déterministes que sont le filtre de Kalman, adapté au cas des systèmes linéaires, et ses extensions pour le cas des systèmes légèrement non-linéaires.

L'incertitude d'une mesure est liée au fait qu'elle peut être une fausse alarme et qu'il est difficile de définir de quel objet elle provient (contexte multi-objet). Les méthodes déterministes précédentes ne sont pas applicables directement. Une étape d'association est nécessaire afin de définir les associations possibles entre les mesures et les objets. Ainsi, sur la base des probabilités d'occurrence de certaines associations, un ensemble de filtres (FBO, PDAF, JPDAF, MHT) calcule une estimation de l'état de chaque objet. Toutes ces méthodes supposent la dépendance des associations : si une mesure provient d'un objet alors elle ne peut provenir d'un autre.

Une méthode de suivi multi-objet dans un contexte multi-capteur doit pouvoir rendre compte du nombre et de l'état des objets présents dans une zone surveillée. Ainsi, elle doit

- déterminer les objets entrant dans la zone et estimer l'état de chacun d'entre eux,
- déterminer les objets sortant de la zone,
- estimer l'état de chaque objet y compris dans le cas où les objets ne sont pas perçus par tous les capteurs.

Parmi les méthodes présentées, seul le filtre MHT répond à ces trois critères. Toutefois, il nécessite une modélisation précise des probabilités d'apparition des fausses alarmes et de celle de l'association entre les mesures et les cibles prédites lors du calcul des probabilités d'occurrence des associations.

Nous proposons d'aborder le problème de l'association grâce à la théorie des fonctions

de croyance. Nous nous placerons dans un contexte multi-capteur. La théorie des fonctions de croyance est un cadre général dans lequel la théorie des probabilités apparaît comme un cas particulier. Elle propose un formalisme beaucoup plus souple pour la représentation de données incertaines. Dans le cadre de notre application, cette théorie n'impose pas de choisir un modèle *a priori* des probabilités d'occurrence des associations. L'utilisation de ce formalisme dans le contexte de l'association de données a été décrite à plusieurs reprises dans la littérature [T3, C65].



Chapitre 2

Association de données dans le cadre de la théorie des fonctions de croyance

2.1 Introduction

Depuis une quinzaine d'années, certains auteurs décrivent la procédure d'association de données dans le cadre de la théorie des fonctions de croyance. C'est une nouvelle approche particulièrement intéressante lorsque le système est en présence de données imparfaites (imprécises et incertaines).

Cette théorie, encore appelée théorie de l'évidence (traduction de l'anglais "Evidence theory"), est le résultat de la formalisation par Glenn Shafer [L17] des travaux d'Arthur Dempster [C19] sur les bornes inférieures et supérieures. Elle permet de combiner des informations provenant de plusieurs sources d'informations alors même que leur représentation est mal connue. Ce contexte est très mal appréhendé par la théorie des probabilités qui reste aujourd'hui le cadre de raisonnement le plus utilisé.

Nous trouvons dans la littérature de nombreux travaux qui ont mis en évidence les bénéfices qu'un système d'association multi-objet et multi-capteur pouvait tirer lorsqu'il était formulé dans le cadre de la théorie des fonctions de croyance [C65, C22, C11, C10, C55, T6, T3, C28, C31]. Tous ces travaux sont inspirés du travail de Rombaut [C65]. Plus récemment, Ayoun [A2] décrit le problème de l'association des données en utilisant les modèles de croyance transférables (TBM, Transferable Belief Model), modèles proposés par Smets [A73]. Dans [L2], Blasch modifie la structure du filtre à association probabiliste décrite dans le chapitre 1 en y intégrant la structure d'un filtre crédibiliste (belief filter). Toutes ces méthodes sont appliquées dans le contexte du suivi/association de données dans un contexte multi-capteur.

Dans la section suivante, nous décrivons les notions de base de la théorie des fonctions de croyance. Puis nous présentons comment Rombaut, Gruyer et Royère expriment la pro-

blématique d'association sous une forme crédibiliste. Nous garderons la structure générale de la méthode afin d'obtenir les résultats présentés au chapitre 4.

2.2 Notions de base de la théorie des fonctions de croyance (TFC)

Dans cette section, nous ne décrivons que l'essentiel de cette théorie. Pour davantage de détails, le lecteur pourra se référer à la lecture de [T5].

2.2.1 Cadre de discernement et référentiel de définition

Soit $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$, l'ensemble des N solutions à un problème donné. Cet ensemble est appelé *cadre de discernement*. Un élément du cadre de discernement est appelé une hypothèse. Les hypothèses sont mutuellement exclusives. Lorsque cet ensemble est supposé exhaustif nous le qualifions de monde fermé. Dans le cas contraire, il est qualifié de monde ouvert. L'adjonction d'une hypothèse ω^* est toujours possible afin de le rendre exhaustif. ω^* regroupe l'ensemble des hypothèses inconnues. Ce monde est qualifié de monde ouvert étendu.

Dans le cadre de la théorie bayésienne, toute la connaissance est distribuée sur un ensemble d'hypothèses singletons, que nous pourrions assimiler à Ω . La théorie des fonctions de croyance propose de distribuer la connaissance sur l'ensemble 2^Ω des sous-ensembles du cadre de discernement. Ce nouvel ensemble est appelé référentiel de définition et est composé de toutes les disjonctions des hypothèses singletons : $2^\Omega = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \dots, \{\omega_1 \cup \omega_2\}, \dots, \Omega\}$. Un sous-ensemble A de 2^Ω est appelé une proposition ou encore hypothèse composée. L'élément \emptyset représente la proposition impossible.

2.2.2 Fonction de croyance

Une fonction de croyance regroupe les opinions d'une source d'information (un capteur, un expert, etc.) sur un système observé. Une opinion est un degré de croyance sur la véracité d'une des propositions du référentiel de définition. Le degré de croyance est une valeur réelle comprise entre 0 et 1. Une fonction de croyance m est définie de la manière suivante :

$$m : 2^\Omega \longrightarrow [0, 1] \quad (2.1)$$

La fonction m vérifie :

$$\left\{ \begin{array}{l} m(\emptyset) = 0 \\ \sum_{A \subseteq 2^\Omega} m(A) = 1 \end{array} \right. \quad (2.2)$$

La valeur $m(A)$ représente la masse associée strictement à A indépendamment de la masse associée à ses sur-ensembles ou ses sous-ensembles. Ceci signifie que dans l'état actuel des connaissances, cette masse ne peut pas être attribuée à un sous ensemble plus spécifique de A . Toutefois, dans l'éventualité d'un apport de nouvelles connaissances, la masse de croyance pourra être allouée plus précisément.

Ainsi dans la théorie des croyances, il n'y a pas de relation entre la masse $m(\omega_1 \cup \omega_2)$ et les masses $m(\omega_1)$ et $m(\omega_2)$. La masse $m(\omega_1 \cup \omega_2)$ peut être nulle alors que les masses $m(\omega_1)$ et $m(\omega_2)$ ne le sont pas. Ceci est une différence fondamentale avec la théorie bayésienne, sur le même exemple, si $p(\omega_1)$ et $p(\omega_2)$ sont nulles alors $p(\omega_1 \cup \omega_2) = p(\omega_1) + p(\omega_2) = 0$. Par ailleurs, aucune relation directe n'existe entre $m(A)$ et $m(\bar{A})$. Dans la théorie des probabilités, nous avons : $p(A) = 1 - p(\bar{A})$.

Les sous-ensembles A de Ω tels que $m(A) \neq 0$, c'est-à-dire l'ensemble des propositions pour lesquelles on dispose d'informations, sont appelés éléments focaux de m . Ainsi on définit le noyau du cadre de discernement par :

$$F(m) = \{A \subseteq \Omega, m(A) \neq 0\} \quad (2.3)$$

La masse affectée au cadre de discernement Ω est la part de croyance affectée à l'ignorance. Le cas où $m(\Omega) = 1$ est l'ignorance totale. La source indique que la solution est une hypothèse du cadre de discernement mais elle ne sait pas laquelle. La croyance sur Ω peut être transférée à tous les autres sous-ensembles de Ω au fur et à mesure que de nouvelles évidences apparaissent et donc l'ignorance diminue. C'est par ce biais que nous disons que la théorie des fonctions de croyance manipule explicitement l'ignorance.

2.2.3 Les Mesures de croyance

Il est possible de définir d'autres mesures de croyance à partir d'une fonction de croyance m : la crédibilité, la plausibilité, la communalité et l'implicabilité.

Fonction de Crédibilité

La fonction de crédibilité notée Bel , est définie comme suit :

$$Bel(A) = \sum_{\emptyset \neq B \subseteq A} m(B). \quad (2.4)$$

Elle vérifie les propriétés suivantes :

$$Bel(\emptyset) = 0 \text{ et } Bel(\Omega) = 1. \quad (2.5)$$

La crédibilité d'une proposition A rassemble les masses de croyance de toutes les propositions B comprises dans A, y compris celle de A. La crédibilité de A représente la croyance minimale qui peut être allouée à A. Si les éléments focaux sont des singletons, alors la fonction de crédibilité est égale à la mesure de probabilité. C'est grâce à cette propriété que la théorie des fonctions de croyance est considérée comme une généralisation de la théorie des probabilités. Par ailleurs, il est possible de calculer la fonction de croyance m en appliquant la transformation de Möbius [A39] à la fonction de crédibilité :

$$m(A) = \sum_{B \subseteq A} (-1)^{card(A-B)} Bel(B). \quad (2.6)$$

où $card(A - B)$ est la cardinalité de l'ensemble des éléments de A qui n'appartiennent pas à B.

Fonction de plausibilité

La fonction de plausibilité notée Pl est la fonction duale de la fonction de crédibilité. Elle est définie par l'équation suivante :

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B) \quad (2.7)$$

Elle vérifie les propriétés suivantes :

$$Pl(\emptyset) = 0, \quad Pl(\Omega) = 1 \text{ et } Pl(A) + Pl(\bar{A}) \geq 1. \quad (2.8)$$

La plausibilité $Pl(A)$ prend en compte la part de croyance des propositions qui ont une intersection avec A, c'est-à-dire des hypothèses qui ne contredisent pas A. Elle représente la part de croyance maximale qui peut être transférée à A.

En utilisant les deux fonctions de crédibilité et de plausibilité, nous pouvons définir l'intervalle de croyance $[Bel(A), Pl(A)]$ associé à la proposition A. Cet intervalle encadre la probabilité bayésienne de A. $Bel(A)$ et $Pl(A)$ sont les bornes inférieures et supérieures de

cette probabilité.

Fonction de communalité (communality function)

Cette fonction est utilisée généralement pour simplifier le calcul lors de la combinaison des sources d'information. Elle est définie de la manière suivante :

$$Q(A) = \sum_{A \subseteq B} m(B). \quad (2.9)$$

Cette fonction vérifie les propriétés suivantes :

$$Q(\emptyset) = 1 \text{ et } Q(\Omega) = m(\Omega). \quad (2.10)$$

Fonction d'implicabilité

L'implicabilité de A rassemble les masses de croyance de toutes les propositions qui impliquent A et la masse de croyance placée sur l'ensemble vide. Cette fonction est définie de la manière suivante :

$$b(A) = Bel(A) + m(\emptyset) = \sum_{B \subseteq A} m(B). \quad (2.11)$$

Relations entre les mesures de croyance

La fonction de croyance m peut être retrouvée à partir des cinq fonctions précédentes en utilisant la transformée de Möbius [A39]. La connaissance de l'une quelconque de ces cinq fonctions, permet de déduire de manière unique les quatre autres fonctions. Denoeux dans [C20] représente cette propriété par le schéma suivant :

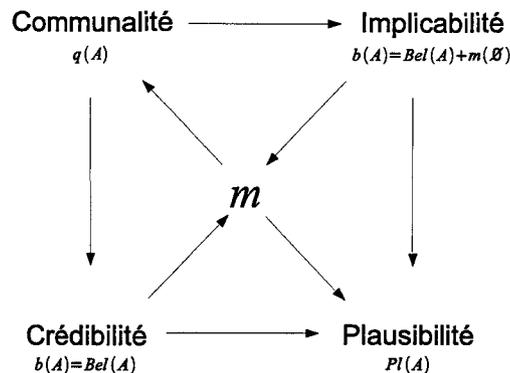


FIG. 2.1 : Les mesures de croyance.

2.2.4 Combinaison des fonctions de croyance

La combinaison est l'étape qui consiste à fusionner de nouvelles connaissances avec la connaissance déjà existante. Son objectif est d'accéder à une information plus "parfaite" du phénomène ou du système observé. Dans la suite nous décrivons la règle de combinaison conjonctive classique et celle proposée par Dempster. Notons qu'il existe d'autres règles de combinaisons (disjonctives et hybrides) dont le lecteur pourra avoir une description en lisant [T2].

2.2.5 Règles de combinaison conjonctive

Soient deux fonctions de croyance m_1 et m_2 , définies sur le même cadre de discernement et issues de deux sources S_1 et S_2 d'information supposées indépendantes et fiables. On note $m_{1\cap 2}$ la fonction de croyance résultant de la combinaison conjonctive de m_1 et m_2 . Cette fonction est définie pour toutes les parties de Ω par :

$$m_{1\cap 2}(A) = \sum_{B\cap C=A} m_1(B).m_2(C). \quad (2.12)$$

Cette règle est associative, commutative et possède un élément neutre. L'associativité et la commutativité de cet opérateur impliquent que le résultat ne dépend pas de l'ordre de combinaison des sources.

La fonction de croyance résultant de la combinaison est non normalisée, c'est-à-dire que la masse de croyance affectée à l'ensemble vide, appelée masse conflictuelle, peut s'avérer être non nulle. Dans le cas du monde fermé, $m_{1\cap 2}(\emptyset)$ doit être nulle. Afin de respecter cette contrainte, Dempster a proposé de normaliser la règle précédente en divisant le résultat par $1 - m_{1\cap 2}(\emptyset)$. Dans ce cas, la règle de combinaison de Dempster, désignée souvent par \oplus , est donnée par la formule suivante :

$$m_{1\oplus 2}(A) = \frac{\sum_{B\cap C=A} m_1(B).m_2(C)}{1 - \sum_{B\cap C=\emptyset} m_1(B).m_2(C)}. \quad (2.13)$$

Le facteur de normalisation $1 - \sum_{B\cap C=\emptyset} m_1(B).m_2(C)$ prend des valeurs comprises entre 0 et 1. Il est introduit afin de redistribuer la masse affectée à l'ensemble vide sur toutes les autres masses.

Tout comme la combinaison conjonctive précédente, la règle de combinaison de Dempster est associative et commutative. Cependant cette règle n'est pas idempotente, c'est-à-dire que $m \oplus m \neq m$. La combinaison de deux sources identiques d'information pourra

donner une information incohérente avec les sources initiales d'information.

2.2.6 Gestion de la masse conflictuelle

La combinaison de fonctions de croyance peut engendrer une masse de croyance non nulle sur l'ensemble vide. Les causes de ce conflit ont différentes origines : un mauvais "fonctionnement" des capteurs ou des experts, une modélisation imprécise des sources d'information etc.

Dans la section précédente, nous avons décrit un opérateur de combinaison qui procède à une étape de normalisation afin d'annuler la masse conflictuelle résultant de la combinaison de fonctions de croyance et de préserver les propriétés des fonctions de croyance. Cette étape est équivalente à une redistribution de toute la masse conflictuelle sur tous les éléments focaux de la fonction de croyance résultante.

Cette règle a montré son efficacité dans plusieurs applications. Toutefois, son utilisation reste valable pour des valeurs faibles du conflit. Dans le cas d'un conflit important, Haenni propose de redéfinir les modèles utilisés lors de la génération des fonctions de croyance [A30].

D'autres auteurs ont proposé d'autres règles de combinaison afin de résoudre le problème du conflit. Ces différentes solutions sont présentées en annexe de ce mémoire. Dans la section suivante, nous présentons la démarche proposée par Smets, développée dans le cadre du modèle des croyances transférables.

2.2.7 Le modèle des croyances transférables

Les modèles des croyances transférables (MCT) est une interprétation non probabiliste de la théorie de Dempster-Shafer proposée par Smets [A72, A73]. Ils doivent leur nom au processus de transfert de masses de croyance d'un sous ensemble à un autre qui ne le contredit pas. Si la masse de croyance est initialement allouée à une proposition A et si, à un instant donné, nous apprenons que la croyance sur une autre proposition B est totale, alors la masse de la proposition A est transférée au sous-ensemble $A \cap B$. Ceci n'est évidemment possible que si $A \cap B \neq \emptyset$. Dans le cas contraire, le transfert n'est plus possible. Le transfert de la masse est effectué grâce à une étape de conditionnement qui est un cas particulier de la règle de combinaison de Dempster.

Le MCT repose sur 2 niveaux : le niveau crédal et le niveau pignistique. Le niveau crédal est composé de deux étapes : une étape de modélisation des fonctions des croyances

et une étape de combinaison. Le niveau pignistique est le niveau dans lequel les décisions sont prises.

Selon Smets [C77], le processus de normalisation contredit la définition de la plausibilité, car en attribuant une partie de la masse conflictuelle à une proposition donnée, la masse résultante peut être supérieure à la masse maximale attendue pour cette proposition (c'est-à-dire sa plausibilité). Dans le MCT, Smets préconise de trouver l'origine du conflit afin de réduire la masse de croyance que le processus de combinaison a attribué à l'ensemble vide. La normalisation des masses est ainsi réalisée, plus tard, au niveau pignistique.

2.2.8 Décision

L'étape de décision a pour objectif de sélectionner la solution au problème posé parmi les éléments du cadre de discernement. Cette étape opère à partir de la fonction de croyance résultant de la combinaison. Les trois principaux critères de décision sont :

- le maximum de crédibilité : c'est le critère de décision pessimiste dans lequel le sous-ensemble du cadre de discernement le plus crédible est choisi comme la solution au problème envisagé. Il s'agit du sous-ensemble du cadre de discernement qui est le plus impliqué dans les autres sous-ensembles du cadre de discernement.
- le maximum de plausibilité : c'est le critère de décision optimiste qui privilégie le sous-ensemble du cadre de discernement le plus plausible, c'est à dire celui qui contredit le moins les informations disponibles.
- le maximum de probabilité pignistique : ce critère est connu comme le critère de décision prudent. Il restreint la prise de décision aux hypothèses singletons du cadre de discernement. Il s'agit du critère employé par Smets au niveau décisionnel du MCF.

Le sous-ensemble résultant d'une décision sur la base des deux premiers critères peut être une hypothèse singleton ou une hypothèse composée.

Probabilité pignistique

La probabilité pignistique est une fonction de probabilité subjective, proposée par Smets [A72] dans le cadre du modèle des croyances transférables. Cette transformation opère en évoquant le principe de raison insuffisante. Ainsi, pour tout élément focal B d'une fonction de croyance m , la masse de croyance portée par B est distribuée de manière équiprobable sur toutes les parties de B . Cette transformation notée $BetP$ est donnée par la formule

suivante :

$$BetP(A) = \sum_{B \subseteq \Omega, A \subseteq B} \frac{1}{|B|} \cdot \frac{m(B)}{1 - m(\emptyset)} \quad \forall A \in \Omega \quad (2.14)$$

Si les fonctions de croyance utilisées sont normalisées, nous avons $m(\emptyset) = 0$ et donc la transformation pignistique devient :

$$BetP(A) = \sum_{B \subseteq \Omega, A \subseteq B} \frac{m(B)}{|B|}. \quad (2.15)$$

Il est important de noter que le passage du niveau crédal au niveau pignistique entraîne une perte d'information. En effet, une fonction de probabilité pignistique est construite de manière unique à partir d'une fonction de croyance mais une infinité de fonctions de croyance permet de construire une même fonction de probabilité pignistique.

Dans la littérature, nous trouvons d'autres transformations probabilistes alternatives à la transformation pignistique qui permettent de transformer une fonction de croyance en une fonction de probabilité ou plus précisément en une fonction de croyance bayésienne [L6, C14, C18]. Parmi ces transformations, nous pouvons citer : la transformation plausibiliste, la transformation probabiliste proportionnelle, la transformation probabiliste prudente, la transformation probabiliste disjunctive.

2.3 Modélisation des fonctions de croyance

Dans un système de fusion utilisant la théorie des probabilités, le choix des probabilités *a priori* joue un rôle primordial. De ce fait, la modélisation des fonctions de croyance constitue un élément fondamental du bon fonctionnement d'un processus fondé sur la théorie des fonctions de croyance. Cette étape d'initialisation est spécifique à l'application visée et au système en présence (type des capteurs ou des experts) et doit quantifier de façon convenable les imperfections liées aux mesures (incomplétude, incertitude, imprécision et fiabilité).

Bien que la modélisation des fonctions de croyance dépende fortement du contexte et de la problématique envisagée, on peut séparer les travaux réalisés dans le domaine en deux grandes classes. D'un côté, les approches fondées sur le calcul d'une fonction de vraisemblance et de l'autre les approches fondées sur le calcul d'une fonction de distance. Dans la suite de ce chapitre, nous présentons quelques unes des approches proposées à ce jour dans la littérature.

Les deux sous-sections suivantes présentent les deux types d'approches dans le contexte d'une application de classification. Soit C l'ensemble des N classes C_i auxquelles nous voulons vérifier l'appartenance d'un vecteur x . Chaque élément ω_i de Ω est l'hypothèse soutenant que x appartient à la classe C_i . Résoudre le problème de classification de x revient à décider en faveur d'une des hypothèses du cadre de discernement.

2.3.1 Les approches fondées sur le calcul d'une fonction de vraisemblance

Parmi ces méthodes, nous pouvons citer la méthode dite séparable proposée par Appriou [A1] et inspirée des travaux de Smets. L'auteur cherche à construire une fonction de croyance m_i^Ω pour chaque hypothèse ω_i du cadre de discernement Ω . Chaque fonction de croyance est alors vue comme une source d'information spécialisée sur une hypothèse donnée. Appriou se place dans le cadre d'un problème de reconnaissance de forme par apprentissage. La fonction de croyance associée à x et spécialisée sur l'hypothèse ω_i est notée $m_i^\Omega \{x\}$.

Il propose deux modèles de génération des fonctions de croyance élémentaires. Ces deux modèles sont établis de façon axiomatique et permettent de prendre en compte toute forme d'imperfection inhérente à une base d'apprentissage. Ils diffèrent de part l'ensemble de leurs éléments focaux.

Le premier modèle répartit les masses de croyance initiales sur les sous-ensembles $\{\omega_i\}$, $\{\bar{\omega}_i\}$ et Ω de la manière suivante :

$$\begin{cases} m_i^\Omega \{x\} (\omega_i) &= \frac{\alpha_i \cdot R_x \cdot P(x|\omega_i)}{1 + R_x \cdot P(x|\omega_i)} \\ m_i^\Omega \{x\} (\bar{\omega}_i) &= \frac{\alpha_i}{1 + R_x \cdot P(x|\omega_i)} \\ m_i^\Omega \{x\} (\Omega) &= 1 - \alpha_i. \end{cases}$$

Le second modèle répartit les masses de croyance initiales sur les sous-ensembles $\{\bar{\omega}_i\}$ et Ω de la manière suivante :

$$\begin{cases} m_i^\Omega \{x\} (\omega_i) &= 0 \\ m_i^\Omega \{x\} (\bar{\omega}_i) &= \alpha_i \cdot R_x \cdot P(x|\omega_i) \\ m_i^\Omega \{x\} (\Omega) &= 1 - \alpha_i \cdot R_x \cdot P(x|\omega_i). \end{cases}$$

Dans les deux modèles, R_x vérifie :

$$R_x \in \left[0, \frac{1}{\max_{i \in \{1, \dots, N\}} \{P(x|\omega_i)\}} \right] \quad (2.16)$$

La fonction de vraisemblance $P(x|\omega_i)$ est obtenue par apprentissage de la distribution de probabilité du modèle de la classe C_i . Par exemple, cette distribution peut être gaussienne ou multi-gaussienne. Le coefficient R_x est un facteur de normalisation. On lui attribue souvent la valeur maximale de l'intervalle dans lequel il prend sa valeur.

L'indice de fiabilité α_i est le facteur de représentativité ou de confiance sur la connaissance de $P(x|\omega_i)$. Il est utilisé comme coefficient d'affaiblissement qui permet de pondérer les masses selon la confiance que l'on accorde à l'apprentissage réalisé de la fonction de vraisemblance.

Avec le premier modèle d'Appriou, $m_i^\Omega \{x\}$ prend une valeur maximale avec $\alpha_i = 1$, $R_x = 1$ et $p(x|\omega_i) = 1$. Cette valeur maximale est égale à 0.5 et elle n'atteint jamais la valeur 1. Selon le deuxième modèle, la valeur 1 est atteinte pour $\overline{\omega_i}$ et nous avons toujours $m_i^\Omega \{x\}(\omega_i) = 0$. Pour le bon fonctionnement de ces deux modèles, le choix du facteur de normalisation R_x est primordial mais reste difficile à réaliser.

2.3.2 Les approches fondées sur le calcul d'une distance

Les fonctions de croyance sont exprimées en fonction d'un critère de distance exprimé dans un espace à plusieurs dimensions. Dans le cadre d'une application de classification, il s'agit de l'espace des caractéristiques retenu pour modéliser un ensemble de classes parmi lesquelles nous voulons classer un vecteur x . Ces approches font l'hypothèse que plus x est proche d'une classe, plus la masse de croyance affectée à l'hypothèse soutenant la classe est élevée.

2.3.2.1 Modélisation de Denoeux

Dans [A22], Denoeux propose de construire la fonction de croyance en gardant la notion de source d'information spécialisée. Soit $T = \{(x^1, C_1), \dots, (x^i, C_j), \dots, (x^M, C_N)\}$ un ensemble d'apprentissage composé de couples (x^i, C_j) où x^i est un vecteur exemple appartenant à la classe C_j . Soit x le vecteur à classer. Pour chaque couple (x^i, C_j) , il construit une fonction de croyance $m_i^\Omega \{x\}$ de la manière suivante :

$$\begin{cases} m_i^\Omega \{x\}(\omega_j) &= \alpha_i \cdot \exp(-\gamma_i \cdot (d_i)^2) \\ m_i^\Omega \{x\}(\Omega) &= 1 - \alpha_i \cdot \exp(-\gamma_i \cdot (d_i)^2) \end{cases}$$

Dans ces équations, $d_i = d(x, x^i)$ est une mesure de distance entre la mesure x et le vecteur exemple x^i dans l'espace des caractéristiques retenu. Le paramètre α_i traduit la confiance

qu'on accorde au vecteur x^i . γ_i est un coefficient réel positif qui permet de spécifier la vitesse de décroissance de la masse en fonction du vecteur d'apprentissage. La fonction de croyance finale est obtenue après combinaison de toute ou partie des fonctions de croyance spécialisée $m_i^\Omega \{x\}$.

En utilisant ce modèle, les masses de croyance sont uniquement attribuées aux sous-ensembles $\{\omega_i\}$ et Ω du cadre du discernement. Aucune masse de croyance n'est attribuée à l'hypothèse $\{\bar{\omega}_i\}$. Cette contrainte empêche la possibilité de créer ou de détruire une classe.

2.3.2.2 Modélisation de Rombaut

La fonction de croyance proposée par Rombaut [C65] est inspirée des travaux de Denoeux. Le jeu de masses diffère de part l'ensemble de ses éléments focaux. Les masses sont distribuées sur les sous-ensembles $\{\omega_i\}$, $\{\bar{\omega}_i\}$ et Ω . Rombaut applique cette méthode dans le cadre de l'association d'un ensemble d'observations à un ensemble de cibles dans un contexte de suivi multi-objet et multi-capteur. Dans cet exemple, l'ensemble T d'apprentissage est composé de couples (x^i, C_i) où x^i est un vecteur exemple appartenant à la classe C_i . Chaque classe C_i correspond à une cible à suivre. La fonction de croyance est définie par les relations suivantes :

$$\begin{cases} m_i^\Omega \{x\} (\omega_i) &= \alpha \cdot \exp(-(d_i)^2) \\ m_i^\Omega \{x\} (\bar{\omega}_i) &= \alpha \cdot (1 - \exp(-(d_i)^2)) \\ m_i^\Omega \{x\} (\Omega) &= 1 - \alpha \end{cases} \quad (2.17)$$

où $d_i = d(x, x^i)$ est une mesure de distance entre la mesure x et le vecteur exemple x^i dans l'espace des caractéristiques retenu et le paramètre α traduit la fiabilité qu'on accorde au capteur ayant donné la mesure x . Comme dans l'approche proposée par Denoeux, la fonction de croyance finale est obtenue après combinaison de toute ou partie des fonctions de croyance spécialisée $m_i^\Omega \{x\}$.

2.3.2.3 Modélisation de Gruyer

La fonction de croyance proposée par Gruyer dans [T3] est inspirée du modèle de Rombaut. Gruyer fait l'hypothèse qu'une source spécialisée ne peut se prononcer à la fois sur l'hypothèse pour laquelle elle est définie et sur son contraire. Ainsi, la fonction de croyance attribue de la masse aux propositions $\{\omega_i\}$ et Ω ou aux propositions $\{\bar{\omega}_i\}$ et Ω . Grâce à ce choix, le modèle proposé permet d'obtenir une masse conflictuelle nulle lors de la combi-

raison de deux fonctions de croyance identiques.

Tout comme Rombaut, Gruyer applique cette méthode dans le cadre de l'association d'un ensemble d'observations à un ensemble de cibles dans un contexte de suivi multi-objet et multi-capteur. Dans cet exemple, l'ensemble T d'apprentissage est composé de couples (x^i, C_i) où x^i est un vecteur exemple appartenant à la classe C_i . Chaque classe C_i correspond à une cible à suivre. Toutefois, contrairement à Rombaut, Gruyer remplace la fonction exponentielle par des fonctions sinusoïdales.

Le jeu de masses de Gruyer est défini selon la valeur prise par la distance d_i entre le vecteur d'observation et le vecteur exemple. Il exprime cette distance en fonction d'un indice de concordance flou \mathfrak{F} qui permet de comparer les vecteurs x^i et x . Les deux modes de fonctionnement sont les suivants :

$$Si \mathfrak{F} \in [0, \tau] \begin{cases} m_i^\Omega \{x\} (\omega_i) = \alpha \cdot \Psi(\mathfrak{F}) \\ m_i^\Omega \{x\} (\bar{\omega}_i) = 0 \\ m_i^\Omega \{x\} (\Omega) = 1 - \alpha \cdot \Psi(\mathfrak{F}) \end{cases} \quad (2.18)$$

$$Si \mathfrak{F} \in]\tau, 1] \begin{cases} m_i^\Omega \{x\} (\omega_i) = 0 \\ m_i^\Omega \{x\} (\bar{\omega}_i) = \alpha \cdot (1 - \Psi(\mathfrak{F})) \\ m_i^\Omega \{x\} (\Omega) = 1 - \alpha \cdot (1 - \Psi(\mathfrak{F})) \end{cases} \quad (2.19)$$

où α est un indice de fiabilité et

$$\Psi(\mathfrak{F}) = \alpha \cdot \left[1 - \left[\frac{\sin(\frac{d_i(\mathfrak{F})}{2}) + 1}{2} \right] \right] \quad (2.20)$$

$$d_i(\mathfrak{F}) = \begin{cases} \pi \cdot \left(\frac{2}{\tau} \cdot (1 - \mathfrak{F}) - 1 \right) & si \mathfrak{F} \in [0, \tau[\\ \frac{\pi}{1-\pi} \cdot (2 \cdot (1 - \mathfrak{F}) - (\tau + 1)) & si \mathfrak{F} \in [\tau, 1]. \end{cases} \quad (2.21)$$

Dans ces équations, le coefficient τ permet de basculer d'un mode de calcul des masses de croyance à l'autre.

2.4 Association crédibiliste de données pour le suivi multi-objet

Cette section est consacrée à la description des travaux antérieurs menés dans le domaine du suivi multi-objet dans un contexte multi-capteur [C65, T3, C50, C55]. Dans ces travaux, Rombaut, Gruyer, Royère et plus récemment Lefevre, proposent une formulation crédibiliste du processus d'association entrant en jeu dans une procédure de suivi (cf. cha-

pitre 1). L'application vise à reconstruire l'environnement routier d'un véhicule.

Le système de perception est composé de plusieurs capteurs extéroceptifs (caméras, radar, lidar, etc.) qui fournissent, à chaque instant, un ensemble de vecteurs d'observation relatifs aux objets situés dans la scène. L'objectif de l'association est de mettre en correspondance chaque objet perçu avec un des objets connus à l'instant de l'acquisition. Outre la fonction de suivi des objets existants, cette étape doit garantir la détection ou la disparition d'un objet. Un vecteur peut renseigner sur la position de l'objet détecté. Un vecteur peut représenter un ensemble d'attributs modaux permettant de caractériser l'objet auquel il se rapporte. Les capteurs peuvent être de natures différentes et fournir ainsi une perception complémentaire de la scène.

Dans un premier temps, nous posons la formulation du problème, les notations proposées par les auteurs et celles retenues dans la suite du mémoire. Dans un deuxième temps, nous définissons les éléments nécessaires à l'application de la théorie des fonctions de croyance (le cadre de discernement, les fonctions de croyance élémentaires, l'étape de décision etc.). Tous ces travaux se placent sous l'hypothèse du monde ouvert étendu.

2.4.1 La problématique

Afin de simplifier la formulation du problème, nous nous plaçons dans le cadre d'un système n'incluant qu'un seul capteur. La suite pourra être étendue à un contexte multi-capteur sans difficulté. Soit $Y(k) = \{Y_1, Y_2, \dots, Y_{m_k}\}$ l'ensemble des m_k vecteurs d'observation fourni par un capteur à l'instant kT . Nous supposons qu'un vecteur Y_l caractérise un seul et unique objet perçu dans la scène. Soit $X(k-1) = \{X_1, X_2, \dots, X_{n_{k-1}}\}$ l'ensemble des vecteurs d'état des n_{k-1} objets connus à l'instant $(k-1)T$. L'objectif de la tâche d'association est de construire l'ensemble $X(k)$ à partir des ensembles $Y(k)$ et $X(k-1)$, c'est-à-dire de déterminer les objets apparus, les objets disparus et d'estimer le vecteur d'état de tous les objets connus à l'instant kT . Afin de prendre en compte les aspects dynamiques du système, une prédiction est appliquée à chaque vecteur d'état de chaque objet connu. Nous notons $X(k-1, k)$ l'ensemble de ces vecteurs.

Gruyer, Royère proposent d'analyser les relations entre les objets perçus et les objets connus à partir de la concordance entre leurs vecteurs d'état respectifs. Ils procèdent en deux étapes. La première étape, appelée association objets perçus/objets connus (APC), consiste à attribuer une croyance aux relations entre chaque objet perçu et tous les objets connus. A l'inverse, la seconde étape, appelée association objets connus/objets perçus

(ACP), attribue une croyance aux relations entre chaque objet connu et tous les objets perçus. La première et la seconde étape permettent respectivement de gérer l'apparition et la disparition des objets. Les deux étapes diffèrent par leur cadre de discernement et par leurs fonctions de croyance élémentaires.

Analyser les relations entre les objets connus et les objets perçus peut être vu comme une procédure de classification. Ainsi, dans la première étape, chaque objet perçu est attribué à l'une des classes créées à partir des objets connus. Inversement, dans la seconde étape, chaque objet connu est attribué à l'une des classes créées à partir des objets perçus. Chaque classe est alors caractérisée par le vecteur d'état (observé ou prédit) de l'objet qui lui correspond (perçu ou connu). Dans la suite de cette section, nous décrivons la méthode d'association sous l'angle de la classification.

2.4.2 Construction du cadre de discernement

Le cadre de discernement regroupe l'ensemble des solutions possibles au problème d'association visé. Les solutions diffèrent selon la phase d'association, APC ou ACP, considérée. Soit Ω_X le cadre de discernement défini dans la phase APC. Ω_X regroupe n_{k-1} hypothèses ω_i et l'hypothèse de fermeture ω^* . L'hypothèse ω_i soutient l'appartenance d'un vecteur d'observation $Y_l(k)$ à la classe C_i caractérisée par le vecteur d'état $X_i(k-1, k)$. L'hypothèse ω^* soutient son appartenance à une classe encore inconnue. Dans ce dernier cas, $Y_l(k)$ peut provenir d'un nouvel objet.

Soit Ω_Y le cadre de discernement défini dans la phase ACP. Ω_Y regroupe m_k hypothèses ω_i et l'hypothèse de fermeture ω^* . L'hypothèse ω_i soutient l'appartenance du vecteur d'état $X_l(k-1, k)$ d'un objet connu à la classe C_i caractérisée par le vecteur d'état $Y_i(k)$. L'hypothèse ω^* soutient que $X_l(k-1, k)$ n'appartient à aucune des classes. Dans ce cas, l'objet connu peut avoir disparu. Dans la suite, nous soulagerons la notation en ne faisant plus apparaître la variable k .

2.4.3 Génération des fonctions de croyance

Les fonctions de croyance élémentaires sont déterminées à partir d'une des méthodes proposées dans la section 2.3. Dans les travaux antérieurs, les modèles de Rombaut et Gruyer sont les plus couramment utilisés. Les résultats que nous exposons dans le chapitre 4 sont obtenus avec la méthode de Rombaut. Ces méthodes sont fondées sur la notion de sources spécialisées. Ces sources diffèrent selon qu'elles sont définies relativement à la

phase ACP ou APC.

Soit Ω_X le cadre de discernement de la phase APC. Soit Y_l un vecteur d'observation et S_i la source spécialisée sur l'hypothèse ω_i . La fonction de croyance construite à partir de l'information fournie par S_i sur Y_l , notée $m_i^{\Omega_X} \{Y_l\}$, est définie par les équations 2.18, 2.19 et 2.17.

Soit Ω_Y le cadre de discernement de la phase ACP. Soit X_l le vecteur d'état de l'objet connu l et S_i la source spécialisée sur l'hypothèse ω_i . La fonction de croyance construite à partir de l'information fournie par S_i sur X_l , notée $m_i^{\Omega_Y} \{X_l\}$, est définie par les équations 2.18, 2.19 et 2.17.

2.4.4 Combinaison des fonctions de croyance

Soient $(m_i^{\Omega_X} \{Y_l\})_{(i,l) \in [1,n+1] \times [1,m]}$ et $(m_i^{\Omega_Y} \{X_l\})_{(i,l) \in [1,m+1] \times [1,n]}$, les fonctions de croyance élémentaires obtenues respectivement dans la phase APC et ACP. Les deux ensembles des masses de croyance sont combinés indépendamment l'un de l'autre. Par ailleurs, la combinaison est réalisée entre les fonctions de croyance relatives à un même vecteur d'observation ou au même vecteur d'état d'un objet connu.

Deux règles de combinaison ont été proposées : la première règle est qualifiée de "combinaison partielle" [T3] et la seconde règle est "une combinaison totale" [T6]. La combinaison partielle proposée par Gruyer est une modification de la règle de Dempster. Elle redistribue toutes les masses de croyance attribuées aux disjonctions d'hypothèses sur la masse de croyance placée sur l'ignorance. Ainsi, à l'issue de cette combinaison, seules les hypothèses du cadre de discernement et le cadre de discernement lui-même peuvent se voir attribuer une masse de croyance. Le résultat de la combinaison est présenté dans le tableau 2.1 pour la phase APC. Un tableau identique est obtenu pour l'autre phase.

La combinaison totale proposée par Royère applique la règle de combinaison conjonctive non normalisée. Aucune redistribution des masses de croyance n'est effectuée. Ainsi une masse de croyance peut être affectée à toutes les propositions du référentiel de définition. Le résultat de la combinaison est présenté dans le tableau 2.2. Un tableau identique est obtenu pour l'autre phase.

Nous constatons aisément que la dimension de la matrice de croyance obtenue par la combinaison totale est plus grande que celle obtenue par la combinaison partielle. La combinaison partielle a tendance à occulter une partie de l'information. La combinaison totale sauvegarde toute l'information jusqu'à l'étape de décision. Ces deux propriétés fixent le

compromis entre la rapidité de la méthode et la représentativité des informations combinées. Dans la suite, ces deux matrices sont appelées *matrices locales d'association*.

$m_{\cap}^{\Omega X} \{.\} (.)$	Y_1	Y_2	...	Y_m
ω_1	$m_{\cap}^{\Omega X} \{Y_1\} (\omega_1)$	$m_{\cap}^{\Omega X} \{Y_2\} (\omega_1)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\omega_1)$
ω_2	$m_{\cap}^{\Omega X} \{Y_1\} (\omega_2)$	$m_{\cap}^{\Omega X} \{Y_2\} (\omega_2)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\omega_2)$
...
ω_n	$m_{\cap}^{\Omega X} \{Y_1\} (\omega_n)$	$m_{\cap}^{\Omega X} \{Y_2\} (\omega_n)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\omega_n)$
ω^*	$m_{\cap}^{\Omega X} \{Y_1\} (\omega^*)$	$m_{\cap}^{\Omega X} \{Y_2\} (\omega^*)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\omega^*)$
Ω	$m_{\cap}^{\Omega X} \{Y_1\} (\Omega)$	$m_{\cap}^{\Omega X} \{Y_2\} (\Omega)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\Omega)$

TAB. 2.1 : Résultats de la combinaison partielle obtenue dans la phase APC.

$m_{\cap}^{\Omega X} \{.\} (.)$	Y_1	Y_2	...	Y_m
ω_1	$m_{\cap}^{\Omega X} \{Y_1\} (\omega_1)$	$m_{\cap}^{\Omega X} \{Y_2\} (\omega_1)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\omega_1)$
...
ω_n	$m_{\cap}^{\Omega X} \{Y_1\} (\omega_n)$	$m_{\cap}^{\Omega X} \{Y_2\} (\omega_n)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\omega_n)$
$\omega_1 \cup \omega_2$	$m_{\cap}^{\Omega X} \{Y_1\} (\omega_1 \cup \omega_2)$	$m_{\cap}^{\Omega X} \{Y_2\} (\omega_1 \cup \omega_2)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\omega_1 \cup \omega_2)$
...
Ω	$m_{\cap}^{\Omega X} \{Y_1\} (\Omega)$	$m_{\cap}^{\Omega X} \{Y_2\} (\Omega)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\Omega)$
ω^*	$m_{\cap}^{\Omega X} \{Y_1\} (\omega^*)$	$m_{\cap}^{\Omega X} \{Y_2\} (\omega^*)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\omega^*)$
\emptyset	$m_{\cap}^{\Omega X} \{Y_1\} (\emptyset)$	$m_{\cap}^{\Omega X} \{Y_2\} (\emptyset)$...	$m_{\cap}^{\Omega X} \{Y_m\} (\emptyset)$

TAB. 2.2 : Résultats de la combinaison totale obtenue dans la phase APC.

2.4.5 La phase de décision

Gruyer décide en appliquant le critère du maximum de croyance. Royère retient l'hypothèse pour laquelle la probabilité pignistique est maximale. Nous décrivons succinctement les deux méthodes.

Maximum de croyance

Afin de prendre en compte d'éventuelles ambiguïtés entre les matrices $m_{\cap}^{\Omega X} \{.\} (.)$ et $m_{\cap}^{\Omega Y} \{.\} (.)$, Gruyer combine les deux matrices, en excluant les deux lignes relatives aux masses de croyance placées sur ω^* et Ω . Le résultat de la combinaison est définie par l'équation suivante :

$$M_{combine}(\omega_i, \omega_j) = m_{\cap}^{\Omega X} \{Y_i\} (\omega_j) \cdot m_{\cap}^{\Omega Y} \{X_j\} (\omega_i) \quad \forall (i, j) \in [1, m] \times [1, n] \quad (2.22)$$

Gruyer applique alors un algorithme d'affectation (Algorithme Hongrois [A42]) sur

$M_{combine}$ afin de maximiser la somme des croyances contenues dans cette matrice. Cet algorithme nécessite une matrice $M_{combine}$ carrée. Si elle ne l'est pas, l'une des deux dimensions est complétée par les éléments virtuels notés Ev . La matrice résultante $M_{combine}^{affecte}$ est composée de 1 et de 0 et a la particularité de ne présenter qu'un seul 1 par ligne et par colonne. $M_{combine}^{affecte}(\omega_i, \omega_j) = 1$ signifie que le vecteur d'observation Y_i et le vecteur d'état X_j peuvent correspondre au même objet. La décision fait intervenir la masse de croyance sur ω^* en s'articulant de la manière suivante :

- Si $M_{combine}^{affecte}(\omega_i, Ev) = 1$ (resp. $M_{combine}^{affecte}(Ev, \omega_j) = 1$) alors Y_i est le vecteur observation d'un nouvel objet (resp. X_j est le vecteur d'état d'un objet disparu).
- Si $M_{combine}^{affecte}(\omega_i, \omega_j) = 1$ et $M_{combine}^{affecte}(\omega_i, \omega_j) > \max(m_{\cap_{partiel}^{\Omega X}} \{Y_i\}(\omega^*), m_{\cap_{partiel}^{\Omega Y}} \{X_j\}(\omega^*))$ alors Y_i et X_j sont deux vecteurs correspondant au même objet ; sinon Y_i est le vecteur d'observation d'un nouvel objet.

L'ignorance intervient après cette phase afin de relativiser la véracité de la décision prise : en présence d'une ignorance importante, la décision proposée peut ne pas être la meilleure décision à prendre.

Maximum de probabilité pignistique

Ces deux ensembles de fonctions de croyance $m_{\cap_{total}^{\Omega X}} \{.\}(\cdot)$ et $m_{\cap_{total}^{\Omega Y}} \{.\}(\cdot)$ sont transformées sous la forme de deux ensembles de fonctions de probabilités pignistiques $P_{\cap_{total}^{\Omega X}} \{.\}(\cdot)$ et $P_{\cap_{total}^{\Omega Y}} \{.\}(\cdot)$ non normalisées. Suite à la transformation pignistique, ces fonctions sont définies exclusivement sur les éléments du cadre de discernement et sur \emptyset .

Afin de prendre en considération les éventuelles ambiguïtés, Royère procède d'une manière analogue à celle de Gruyer en combinant les deux matrices des probabilités pignistiques. Toutefois, Royère effectue la moyenne arithmétique des deux matrices contrairement à Gruyer qui applique un produit terme à terme (cf. équation 2.22). La matrice combinée s'exprime alors par la relation suivante :

$$\begin{cases} M_{combine}(\omega_i, \omega_j) &= \frac{P_{\cap_{total}^{\Omega X}} \{Y_i\}(\omega_j) + P_{\cap_{total}^{\Omega Y}} \{X_j\}(\omega_i)}{2} & \forall (i, j) \in [1, m] \times [1, n] \\ M_{combine}(\omega_i, \omega_j^*) &= P_{\cap_{total}^{\Omega X}} \{Y_i\}(\omega^*) & \forall i \in [1, m] \text{ et } j = i \\ M_{combine}(\omega_i^*, \omega_j) &= P_{\cap_{total}^{\Omega Y}} \{X_j\}(\omega^*) & \forall j \in [1, n] \text{ et } i = j \end{cases} \quad (2.23)$$

Nous pouvons noter que $M_{combine}$ est composée de n hypothèses ω_j^* et m hypothèses ω_i^* supplémentaires de telle sorte que cette matrice soit carrée.

Tout comme Gruyer, Royère applique un algorithme d'affectation à $M_{combine}$ afin d'obtenir une matrice $M_{combine}^{affecte}$ sur laquelle il décide sur les associations. La méthode de Royère

s'avère moins rapide que la méthode de Gruyer en présence d'un nombre important de disjonctions d'hypothèses car elle sauvegarde toute l'information jusqu'au passage au niveau pignistique. Afin de pallier ce problème, il propose de ne calculer qu'une approximation de chaque fonction de probabilité pignistique grâce à un processus itératif qui permet de l'arrêter dès que l'estimation de la distribution des probabilités pignistiques est suffisamment proche de la distribution réelle. L'estimation est réalisée en distribuant la part de croyance portée par les disjonctions d'hypothèses par ordre croissant de leur cardinalité. Cette technique suppose que plus la cardinalité d'une proposition est élevée moins cette proposition influence le calcul de la probabilité pignistique [T6].

2.4.6 Conclusion

Dans cette section nous avons présenté la démarche généralement employée pour implémenter une procédure de suivi/association dans un cadre crédibiliste. Dans la pratique, la mise en place de cette démarche diffère d'un auteur à l'autre en fonction :

- du modèle utilisé pour la génération des fonctions de croyance élémentaires,
- de la règle de combinaison et
- de la règle de décision.

A ce jour, l'activité majeure a été portée sur la manière de formuler l'apparition et la disparition des objets dans le cadre de la théorie des fonctions de croyance. Ces deux phénomènes ont pour conséquence de considérer le cadre de discernement comme non exhaustif. Afin de pallier cette difficulté, les méthodes décrites dans la littérature étudient l'association sous l'hypothèse d'un monde ouvert étendu. Ce monde est caractérisé par un cadre de discernement fermé grâce à une hypothèse qui représente toutes les hypothèses inconnues. Toutes ces méthodes présentent des résultats satisfaisants mais nécessitent de prendre quelques précautions.

2.5 Discussions

2.5.1 Normalisation de la combinaison des matrices locales d'association

Les matrices locales d'association sont normalisées : la somme par colonne des valeurs de chaque matrice vaut 1. Les deux opérations de combinaison proposées par Gruyer et Royère ne conservent pas cette propriété. Il est donc préférable de normaliser la matrice résultante afin que la somme de ses valeurs par ligne et par colonne valent 1. Toutes ces

composantes étant positives, la matrice devient alors doublement stochastique (doubly stochastic matrix).

2.5.2 L'hypothèse du "monde ouvert étendu" : des précautions à prendre

Comme les méthodes décrites dans la littérature l'ont montré [C65, C22, C10, C55, T6, C28], travailler dans le cadre du monde ouvert étendu est une solution pour gérer l'apparition et la disparition d'objets dans la scène. L'hypothèse de fermeture ajoutée au cadre de discernement regroupe toutes les solutions qui sont inconnues juste avant l'instant où les nouvelles observations sont mises à disposition du système. Bien que cette hypothèse supplémentaire soit exclusive par rapport à toutes les autres hypothèses du cadre de discernement, aucune information ne nous est donnée sur la cardinalité de cette hypothèse. Dans [A74], Smets nous met en garde contre cette donnée inconnue dans le cadre d'une décision de type pignistique en perturbant la distribution finale des probabilités pignistiques.

2.6 Conclusions

Dans la première partie de ce chapitre, nous avons présenté les bases de la théorie des fonctions de croyance. Cette théorie a l'avantage de modéliser et de combiner de manière adaptée des informations imparfaites. Cette théorie n'intègre aucune méthode générique de génération des fonctions de croyance et la gestion de la masse conflictuelle reste un problème ouvert (modèles incorrects, dysfonctionnement des sources d'information, cadre de discernement non exhaustif etc.). Smets propose un cadre particulier de la théorie, le modèle des croyances transférables, qui permet de prendre en compte chacun de ces problèmes. Ce cadre se décompose en deux niveaux : le niveau crédal et le niveau pignistique. L'objectif est de déceler l'origine du conflit au niveau crédal afin de le réduire et de rendre possible une décision par maximisation des probabilités pignistiques.

Dans la seconde partie, nous avons décrit la démarche proposée par quelques auteurs afin de développer une procédure de suivi/association d'objets dans un contexte multi-capteur. Toutes ces méthodes travaillent dans le cadre du monde ouvert étendu qui définit un cadre de discernement fermé par une hypothèse supplémentaire. Cette hypothèse permet de gérer avec une certaine efficacité l'apparition et la disparition des objets. L'association est réalisée en deux phases : une phase d'association objet perçu/objet connu et une phase d'association objet connu/objet perçu. Une procédure est mise en place afin d'éliminer d'éventuelles ambiguïtés entre ces deux associations locales. La décision est fondée

soit sur le critère du maximum de crédibilité, soit sur le critère du maximum de probabilité pignistique.

Dans la suite de ce mémoire, contrairement à toutes les méthodes proposées à ce jour, nous proposons de poser le problème de l'association dans le contexte d'un cadre de discernement ouvert et non-exhaustif. Nous appliquons le modèle des croyances transférables de Smets afin de gérer le conflit et de détecter les phénomènes d'apparition et de disparition d'objets.

Chapitre 3

Gestion du conflit dans le cadre de la théorie des fonctions de croyance dans le contexte du monde ouvert. Gestion de l'apparition et de la disparition d'objets

La gestion de l'apparition et de la disparition d'objets constitue une étape essentielle de tout processus de suivi-association d'objets. Dans de nombreux travaux menés dans le cadre de la théorie des fonctions de croyance, ces deux phénomènes sont reliés à la non-exhaustivité du cadre de discernement : l'apparition d'un nouvel objet se traduit par la création d'une hypothèse supplémentaire et, a contrario, la disparition d'un objet équivaut à la suppression de l'hypothèse correspondante.

Dans [C65, C28, C55], les auteurs se placent dans le contexte d'un monde ouvert étendu c'est-à-dire dont le cadre de discernement est complété par une hypothèse de fermeture qui représente toutes les hypothèses inconnues. Ce nouvel élément est exclusif par rapport aux autres hypothèses du cadre de discernement. Après la phase de combinaison des fonctions de croyance, la masse de croyance attribuée à l'hypothèse de fermeture permet d'identifier les objets qui sont entrés dans le champ de perception des capteurs ou qui en sont sortis.

Lorsque les masses de croyance initiales sont générées à partir des informations fournies par des sources spécialisées, la masse de croyance attribuée à cette hypothèse supplémentaire est non nulle uniquement lorsque toutes les sources ont soutenu l'hypothèse contraire pour laquelle elles étaient spécialisées. Il suffit que l'une d'entre elles ne respecte pas cette condition pour que cette masse de croyance soit nulle. Dans ce cas, une décision prise sur la base du maximum de crédibilité [C31, T3] ne peut conclure à l'apparition ou à la disparition d'un objet. Dans [C32], les auteurs proposent d'appliquer une transformation pignistique afin de faire apparaître une probabilité sur l'élément de fermeture et permettre une décision en faveur de l'apparition et la disparition d'un objet.

Dans ce chapitre nous proposons de formuler le problème des apparitions et des disparitions dans le contexte du monde ouvert non étendu. Dans ce monde, le cadre de discerne-

ment est considéré comme non exhaustif mais ne comporte pas l'hypothèse de fermeture. Dans une première partie, nous poserons le formalisme de la méthode et tout particulièrement la manière dont nous proposons de résoudre le problème de la création et la suppression d'une hypothèse du cadre de discernement à partir de la masse de croyance conflictuelle. Dans une seconde partie, nous comparerons notre méthode à celles formulées sur la base d'un cadre de discernement ouvert étendu.

Bien que nous ayons proposé cette méthode afin de résoudre une partie des problèmes liés à une application de filtrage/association multi-objet, il faut indiquer qu'elle peut être appliquée dans tout contexte de classification dont l'association est un cas particulier. La méthode de gestion du conflit proposée permet de gérer la création ou la destruction de classes du cadre de discernement.

3.1 L'étape d'association sous l'hypothèse du monde ouvert

Rappelons que l'hypothèse du monde ouvert est préconisée par Smets [A74]. Sous cette hypothèse, la masse de croyance sur l'ensemble vide peut être non nulle et une partie de cette masse conflictuelle peut provenir d'une situation où la solution à notre problème d'association n'est pas comprise dans le cadre de discernement. En effet, dans le cadre de MCT, le conflit engendré lors de la combinaison a au moins trois origines :

- le cadre de discernement est non exhaustif, c'est-à-dire que des propositions alternatives ont été oubliées,
- deux sources fiables rapportent sur deux objets différents,
- deux sources fiables rapportent sur un même objet selon deux modèles imprécis.

Lors de la manipulation des fonctions de croyance, le conflit doit être pris en compte afin de conduire à la meilleure décision. Les règles de combinaisons proposées dans la littérature opèrent de trois manières différentes :

- en procédant à une phase de normalisation (combinaison de Demspter [C19]),
- en distribuant la masse de croyance conflictuelle sur un ensemble spécifique d'hypothèses (opérateurs de Dubois et Prade [A24] et de Yager [A85]),
- en réduisant la masse de croyance conflictuelle, soit en affaiblissant la fiabilité de certaines sources d'informations y ayant contribué, soit en recherchant son origine exacte [A74].

Afin d'avoir une liste assez complète des opérateurs et une description plus précise de chacun d'entre eux, le lecteur pourra se référer à l'annexe 1 de ce mémoire.

Notre méthode s'inscrit dans la troisième catégorie. Nous proposons d'analyser l'évolution de la masse de croyance conflictuelle afin de décider ou non de la modification du cadre de discernement et plus particulièrement de l'ajout ou de la suppression d'une hypothèse traduisant l'apparition ou la disparition d'un nouvel objet. La suite de cette section décrit les étapes suivantes :

1. Définition du cadre de discernement et des fonctions de croyances initiales
2. Combinaison des fonctions de croyance
3. Gestion du conflit et mise à jour du cadre de discernement
4. Décision

Il est important de noter que le contexte de l'association est identique à celui décrit dans le chapitre précédent et que nous nous plaçons sous l'hypothèse (H1) du premier chapitre qui suppose qu'un objet réel ne peut être à l'origine que d'au plus une mesure.

3.1.1 Définition du cadre de discernement et des fonctions de croyance

Soit $Y(k) = (Y_l(k))_{l=1, \dots, m_k}$ l'ensemble des m_k mesures fournies par un capteur à l'instant $k.T$. Soit $X_i(k)$ le processus d'état de l'objet O_i^k de l'ensemble $O(k)$ des objets connus. $X(k) = (X_i(k))_{i=1, \dots, n_k}$ est l'ensemble des processus d'état des n_k objets connus à l'instant $k.T$. L'objectif est d'associer chaque mesure $Y_l(k)$ à l'un des objets connus afin de mettre à jour son état.

A chaque instant $k.T$, l'association est équivalente à un problème de classification dans lequel chaque objet connu O_i^k constitue une classe ω_i^k dont le modèle est défini par les caractéristiques liées à son état. Une mesure $Y_l(k)$ est alors associée à l'objet O_i^k si $Y_l(k)$ appartient à la classe ω_i^k . Puisque les variables sont définies pour un même instant, nous simplifierons les notations en ne renseignant plus la variable k .

Soit $\Omega_X = \{\omega_i\}_{i=1, \dots, n}$ l'ensemble composé des classes ω_i auxquelles chaque mesure peut appartenir. Soit $c(x)$ la fonction définie sur Y , à valeurs dans Ω_X , telle que $c(Y_l) = \omega_i$ lorsque Y_l appartient à la classe ω_i . L'ensemble Ω_X constitue le cadre de discernement et chaque élément ω_i de cet ensemble est appelé une proposition (ou une hypothèse). Associer une mesure Y_l à un objet O_i revient à vérifier la véracité de la proposition " $c(Y_l) = \omega_i$ ". Si cette proposition est vraie nous disons que "l'hypothèse ω_i est vraie".

Toutes les hypothèses de Ω_X sont disjointes. Nous nous plaçons dans le cadre du monde ouvert et le cadre de discernement est non exhaustif. Une ou plusieurs hypothèses sont donc susceptibles d'apparaître ou de disparaître. Le cadre de discernement peut se voir

attribuer une hypothèse supplémentaire dans le cas où une mesure n'a été attribuée à aucune classe. Au contraire, une hypothèse peut être éliminée lorsque aucune mesure n'a été attribuée à la classe correspondante.

Soit $m_i^{\Omega_X} \{Y_l\}$ la fonction de croyance élémentaire (Basic Belief Function) associée à l'objet O_i et définie pour chaque mesure Y_l . Cette fonction est définie sur l'ensemble 2^{Ω_X} (référentiel de définition) et prend ses valeurs dans l'intervalle $[0, 1]$. Elle vérifie les propriétés suivantes :

$$m_i^{\Omega_X} \{Y_l\} (\emptyset) = 0 \quad (3.1)$$

$$\sum_{A \subseteq \Omega_X} m_i^{\Omega_X} \{Y_l\} (A) = 1. \quad (3.2)$$

$m_i^{\Omega_X} \{Y_l\} (A)$ représente la masse de croyance associée à l'hypothèse A du référentiel de définition.

Comme nous l'avons indiqué dans le premier chapitre et rappelé dans le second chapitre, l'association est réalisée entre l'ensemble des mesures et l'ensemble des objets connus dont l'état est prédit pour l'instant $k.T$, instant auquel le capteur fournit les mesures. Les fonctions de croyance initiales sont obtenues en appliquant l'une des méthodes présentées dans le chapitre précédent. $n \times m$ fonctions de croyance initiales sont alors obtenues. Chacune d'entre elles constitue une source d'informations à fusionner.

3.1.2 Combinaison des fonctions de croyance

A l'issue de cette phase d'initialisation, les n fonctions de croyance initiales $(m_i^{\Omega_X} \{Y_l\})_{i=1, \dots, n}$ définies pour la même mesure Y_l sont combinées. L'opérateur que nous utilisons est l'opérateur conjonctif non normalisé noté \cap . Ainsi le résultat de la combinaison des fonctions de croyances $m_1^{\Omega_X} \{Y_l\}$ et $m_2^{\Omega_X} \{Y_l\}$, notée $m_{1 \cap 2}^{\Omega_X} \{Y_l\}$, est définie par :

$$m_{1 \cap 2}^{\Omega_X} (A) = \sum_{B \cap C = A} m_1^{\Omega_X} \{Y_l\} (B) . m_2^{\Omega_X} \{Y_l\} (C) \quad \forall A \subseteq \Omega_X. \quad (3.3)$$

Cet opérateur est associatif. Il fournit un résultat qui ne dépend pas de l'ordre de combinaison des fonctions de croyance. Par ailleurs, contrairement à l'opérateur de combinaison de Dempster, la masse de croyance sur l'ensemble vide peut être non nulle :

$$m_{1 \cap 2}^{\Omega_X} \{Y_l\} (\emptyset) \neq 0.$$

Nous pouvons étendre l'équation 3.3 à la combinaison, notée $m_{\cap}^{\Omega_X} \{Y_l\}$, des n fonctions de croyance initiales $(m_i^{\Omega_X} \{Y_l\})_{i=1, \dots, n}$:

$$m_{\cap}^{\Omega_X} \{Y_l\} (A) = \sum_{A_1 \cap \dots \cap A_n = A} \left(\prod_{i=1}^n m_i^{\Omega_X} \{Y_l\} (A_i) \right) \quad \forall A \subseteq \Omega_X. \quad (3.4)$$

A partir de 3.4, il est aisé de généraliser l'expression de la masse conflictuelle résultante de la combinaison des n fonctions de croyance initiales :

$$m_{\cap}^{\Omega_X} \{Y_l\} (\emptyset) = \sum_{A_1 \cap \dots \cap A_n = \emptyset} \left(\prod_{i=1}^n m_i^{\Omega_X} \{Y_l\} (A_i) \right). \quad (3.5)$$

Cette masse conflictuelle reflète le degré du désaccord qui existe entre les sources combinées. Comme l'a montré Lefevre dans [T5], l'opérateur de Dempster est très sensible au coefficient de normalisation. Il est donc nécessaire de surveiller la valeur atteinte par la masse conflictuelle afin de ne pas procéder à une normalisation systématique des masses de croyance résultantes et de préserver la nature de l'information qu'elles représentent.

3.1.3 Conflit et cadre de discernement non exhaustif

3.1.3.1 Les méthodes existantes

L'origine du conflit est de plusieurs types. Il peut être dû à des hypothèses manquantes dans le cadre de discernement ou à une autre origine, comme une modélisation imparfaite des fonctions de croyances initiales (voir annexe 1 de ce mémoire). Dans le cas où la masse conflictuelle résultant uniquement d'un cadre de discernement non exhaustif, un nouveau cadre de discernement peut alors être défini.

Soit Ω_X^* ce nouveau cadre de discernement. Il est défini par $\Omega_X^* = \Omega_X \cup \{\omega^*\}$ où l'hypothèse ω^* regroupe toutes les propositions inconnues à l'instant de l'association. Les fonctions de croyance initiales associées à Ω_X^* sont calculées à partir de celles définies sur Ω_X . Il faut donc décider pour chaque partie A de Ω_X , les masses de croyance à attribuer à A et à $A \cup \{\omega^*\}$. Cette répartition est particulièrement difficile à justifier lorsque nous ne possédons aucune information sur ω^* . Afin de pallier cette difficulté, il est possible d'adopter la solution la moins informative ("the least committed solution") ou de se placer dans le

contexte du monde ouvert étendu.

La règle du "the least committed solution" propose d'allouer toute la masse de croyance d'une partie A de Ω_X à $A \cup \{\omega^*\}$ [A74]. Cette règle est appliquée aux fonctions de croyance initiales avant même leur combinaison. Ainsi, l'intersection des éléments focaux de chaque fonction de croyance étant non nulle, la masse de croyance conflictuelle résultant de la combinaison devient nulle. L'exemple 3.1 illustre cette propriété.

Exemple 3.1 :

Soit le cadre de discernement $\Omega_X = \{\omega_1, \omega_2\}$. Soient $m_1^{\Omega_X}$ et $m_2^{\Omega_X}$ deux fonctions de masses de croyance initiales définies sur Ω_X et non nulles pour toute partie A de Ω_X . Soit $m_1^{\Omega_X^*}$ et $m_2^{\Omega_X^*}$ les fonctions de croyance initiales obtenues après l'application de la règle du "the least committed solution". $m_1^{\Omega_X^*}$ et $m_2^{\Omega_X^*}$ sont non nulles sur les parties de Ω_X^* suivantes : $\{\omega_1\} \cup \{\omega^*\}, \{\omega_2\} \cup \{\omega^*\}$ et $\Omega_X \cup \{\omega^*\}$. L'intersection deux à deux de ces éléments focaux est non nulle et aucun conflit n'apparaît après la phase de combinaison.

Il est important de noter que "the least committed solution" revient à attribuer toute la masse de croyance conflictuelle résultante obtenue sur Ω_X à l'hypothèse ω^* . Après l'étape de combinaison, l'ensemble des éléments focaux de la fonction de croyance résultante n'est composé d'aucun singleton (cf. exemple précédent). Par conséquent, seule une décision selon le maximum de probabilité pignistique est possible. L'utilisation d'un autre critère (maximum de crédibilité ou de plausibilité) est possible mais nécessite une étape supplémentaire afin d'attribuer une masse de croyance aux singletons de Ω_X^* . Par ailleurs, comme le souligne Smets, la décision pignistique nécessite de prendre des précautions car la cardinalité de ω^* n'étant pas connue, la décision peut s'avérer caduque.

Dans [C65, C28, C55], dont nous avons présenté les travaux dans le chapitre 2, les auteurs se placent sous l'hypothèse du monde ouvert étendu et définissent les fonctions de croyance initiales directement sur $2^{\Omega_X^*}$. Après la phase de combinaison, toute la masse de croyance est répartie sur l'ensemble des sous-parties de Ω_X^* . Cette solution diffère de la règle du "the least committed solution" sur les points suivants :

1. les éléments focaux peuvent être disjoints et la masse de croyance conflictuelle résultante peut être non nulle,
2. la masse de croyance résultante attribuée aux singletons de Ω_X^* peut être non nulle ; dans ce cas, l'ensemble des règles de décision est applicable et la mise en garde de Smets concernant l'influence de la cardinalité de ω^* sur la qualité de la décision pignistique reste valable.

3.1.3.2 Notre méthode

Les deux solutions précédentes sont fondées sur l'expression des fonctions de croyance initiales sur Ω_X^* . Ceci se traduit par une disparition totale ("the least committed solution") ou partielle (monde ouvert étendu) de la masse de croyance conflictuelle après la phase de combinaison. A l'inverse, la méthode que nous proposons opère après la phase de combinaison.

Nous étendons la règle proposée par Smets au cas des sources d'informations "faiblement non fiable et non précis". Dans ce contexte, nous supposons que la grande partie de la masse conflictuelle est due à des hypothèses manquantes et que la partie restante n'a pas d'influence lors de la phase de décision. Nous justifions cette hypothèse de travail en mentionnant bien qu'il est toujours difficile de déterminer la fiabilité de chacune des sources à combiner et les modèles à partir desquels les fonctions de croyance élémentaires sont calculées. Dans [A30, C33], Haenni précise d'ailleurs que la phase de modélisation des problèmes est primordiale. Elle doit être menée de telle sorte à ce que les modèles soient les plus représentatifs des phénomènes observés. Ainsi, la masse conflictuelle résultante de l'application de la règle de combinaison de Dempster est réduite, ce qui évite d'introduire l'utilisation d'autres opérateurs qui ne vérifient pas les propriétés d'associativité ou de commutativité.

Soit $m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset)$ la masse conflictuelle définie par l'équation 3.5. Nous supposons que cette masse peut s'écrire de la manière suivante :

$$m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset) = m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset^{un}) + m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset^*) \quad (3.6)$$

où $m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset^*)$ représente le conflit engendré par un cadre de discernement non exhaustif et $m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset^{un})$ correspond au conflit dont l'origine est indéterminée. Notre méthode propose de prendre une décision sur un nouveau cadre de discernement Ω_X^* tel que :

- la masse de croyance conflictuelle due à la non exhaustivité est nulle et
- la masse de croyance conflictuelle d'origine inconnue est conservée.

Soit $m_{\cap}^{\Omega_X^*} \{Y_i\}$ la fonction de croyance relative à la mesure Y_i , définie sur $2^{\Omega_X^*}$. Nous la calculons à partir de la fonction de croyance $m_{\cap}^{\Omega_X} \{Y_i\}$ de la manière suivante :

$$\begin{cases} m_{\cap}^{\Omega_X^*} \{Y_i\} (A) & = m_{\cap}^{\Omega_X} \{Y_i\} (A) \quad \forall A \subseteq \Omega_X - \{\emptyset\} \\ m_{\cap}^{\Omega_X^*} \{Y_i\} (\emptyset) & = m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset^{un}) \\ m_{\cap}^{\Omega_X^*} \{Y_i\} (\omega^*) & = m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset^*) \end{cases} \quad (3.7)$$

La valeur de $m_{\cap}^{\Omega_X} \{Y_i\} (\emptyset^*)$ est donnée par l'équation 3.5 appliquée à toutes les parties A_i de Ω_X telles que :

$$\begin{cases} \text{card}(A_i) = \text{card}(\Omega_X) - 1 \\ \omega_i \notin A_i \end{cases}$$

Ces deux conditions signifient que chaque fonction de croyance $m_i^{\Omega_X}$ indique que la mesure Y_i peut ne pas appartenir à la classe ω_i à laquelle la source d'information est associée. Finalement, une partie de l'information portée par les n fonctions de croyance indique que la mesure Y_i peut n'appartenir à aucune des classes ω_i . Dans ces conditions, la mesure Y_i peut conduire à la création d'une nouvelle classe traduisant l'apparition d'un nouvel objet. Cette classe est représentée par ω^* .

Exemple 3.2 : Illustration de notre méthode.

Afin d'illustrer notre méthode, prenons l'exemple d'un cadre de discernement composé de deux classes auxquelles peut appartenir chaque mesure. Soit $\Omega_X = \{\omega_1, \omega_2\}$ ce cadre de discernement. Soit $\Omega_X^* = \Omega_X \cup \{\omega^*\} = \{\omega_1, \omega_2, \omega^*\}$ le cadre de discernement complété par l'hypothèse de fermeture ω^* . Soient Y_1 une mesure et $m_1^{\Omega_X} \{Y_1\}$ et $m_2^{\Omega_X} \{Y_1\}$ les deux fonctions de croyance relatives à la mesure Y_1 . Ces deux fonctions de croyance initiales sont définies sur les trois parties suivantes de Ω_X : $\{\omega_i\}$, $\{\bar{\omega}_i\} = \Omega_X - \{\omega_i\}$ et $\{\Omega_X\}$.

Le tableau 3.1 regroupe les sous-parties de Ω_X^* sur lesquelles sont distribuées les masses de croyance après la combinaison de $m_1^{\Omega_X^*}$ et de $m_2^{\Omega_X^*}$ dans le contexte du monde ouvert étendu.

Le tableau 3.2 regroupe les sous parties de Ω_X sur lesquelles sont distribuées les masses de croyance après la combinaison de $m_1^{\Omega_X}$ et de $m_2^{\Omega_X}$ dans le contexte de notre méthode.

$m_{1 \cap 2}^{\Omega_X^*}$		$m_2^{\Omega_X^*}$		
		ω_2	Ω_X^*	$\bar{\omega}_2$
$m_1^{\Omega_X^*}$	ω_1	\emptyset	ω_1	ω_1
	Ω_X^*	ω_2	Ω_X^*	$\omega_1 \cup \omega^*$
	$\bar{\omega}_1$	ω_2	$\omega_2 \cup \omega^*$	ω^*

$m_{1 \cap 2}^{\Omega_X}$		$m_2^{\Omega_X}$		
		ω_2	Ω_X	$\bar{\omega}_2$
$m_1^{\Omega_X}$	ω_1	\emptyset^{un}	ω_1	ω_1
	Ω_X	ω_2	Ω_X	ω_1
	$\bar{\omega}_1$	ω_2	ω_2	\emptyset^*

TAB. 3.1 : Parties de $2^{\Omega_X^*}$ résultantes de la combinaison de $m_1^{\Omega_X^*}$ et de $m_2^{\Omega_X^*}$ dans le contexte du monde ouvert étendu.

TAB. 3.2 : Parties de 2^{Ω_X} résultantes de la combinaison de $m_1^{\Omega_X}$ et de $m_2^{\Omega_X}$ dans le contexte de NM.

D'après ces deux tableaux, nous remarquons que la différence entre le monde ouvert étendu et notre méthode réside dans la définition des éléments focaux résultant de la combinaison des deux fonctions de croyance initiales.

Dans la méthode du monde ouvert étendu, l'expression des masses de croyance initiales sur Ω_X^* engendre l'apparition de l'hypothèse ω^* dans la composition des sous-ensembles $\bar{\omega}_i$. Comme nous allons le démontrer dans la suite, l'introduction de l'hypothèse ω^* dans le cadre de discernement

perturbe la distribution réelle des masses de croyances sur les sous-ensembles du cadre de discernement et engendre une décision imprécise et incertaine.

Avec notre méthode, le conflit engendré après la combinaison des sources d'information est séparé en deux composantes : un conflit dû à un cadre de discernement non exhaustif représenté par $m_{1\cap 2}^{\Omega_X}(\emptyset^*)$ et un conflit inconnu $m_{1\cap 2}^{\Omega_X}(\emptyset^{un})$ dont l'origine est indéterminée. Ceci et comme le montre le tableau 3.2 n'entraîne pas une modification des sous-ensembles du cadre de discernement.

Dans cet exemple, la valeur du conflit $m_{1\cap 2}^{\Omega_X}(\emptyset^*)$ est donnée par :

$$m_{1\cap 2}^{\Omega_X}(\emptyset^*) = m_1^{\Omega_X}(\bar{w}_1) \times m_2^{\Omega_X}(\bar{w}_2) \quad (3.8)$$

Ce conflit est créé puisque les deux sources d'information $m_1^{\Omega_X}\{Y_1\}$ et $m_2^{\Omega_X}\{Y_1\}$ relatives aux deux hypothèses ω_1 et ω_2 confirment respectivement à travers les masses quantifiant la non-association, $m_1^{\Omega_X}(\bar{w}_1)$ et $m_2^{\Omega_X}(\bar{w}_2)$ que la mesure Y_1 n'appartient pas à aucune des deux classes. Ceci indique que cette mesure correspond à un nouvel objet inconnu du système de suivi. Ainsi, le cadre de discernement n'est pas exhaustif.

3.1.4 Décision

La décision est prise au niveau pignistique sur les singletons du cadre de discernement Ω_X . La probabilité pignistique $BetP_{\Omega_X}\{Y_l\}(\omega_i)$ sur $\omega_i \in \Omega_X$ s'exprime de la manière suivante :

$$BetP_{\Omega_X}\{Y_l\}(\omega_i) = \sum_{A \in 2^{\Omega_X}, \omega_i \subset A} \frac{1}{|A|} \frac{m_{\cap}^{\Omega_X}\{Y_l\}(A)}{1 - m_{\cap}^{\Omega_X}\{Y_l\}(\emptyset)}$$

La partie $m_{\cap}^{\Omega_X}\{Y_l\}(\emptyset^*)$ de la masse conflictuelle est affectée à l'hypothèse ω^* et la partie $m_{\cap}^{\Omega_X}\{Y_l\}(\emptyset^{un})$ est affectée à l'ensemble vide de Ω_X^* . Par conséquent, la distribution des masses de croyance vérifie la relation suivante :

$$\sum_{i=1}^n BetP_{\Omega_X}\{Y_l\}(\omega_i) + BetP_{\Omega_X}\{Y_l\}(\omega) = 1$$

Cette contrainte peut être retrouvée directement à partir de l'expression de la probabilité pignistique sur les éléments de Ω_X .

Démonstration :

Soit $BetP_{\Omega_X}\{Y_l\}(\omega_i)$ la probabilité pignistique de $\omega_i \in \Omega_X$. Cette distribution de probabilité sur Ω_X vérifie :

$$\sum_{i=1}^n BetP_{\Omega_X}\{Y_l\}(\omega_i) = 1 \quad (3.9)$$

Selon la décomposition du conflit proposée par l'équation 3.6, l'égalité 3.9 devient :

$$\frac{1}{1 - m_{\Omega}^{\Omega_X} \{Y_l\} (\emptyset^{un}) - m_{\Omega}^{\Omega_X} \{Y_l\} (\emptyset^*)} \sum_{i=1}^n \sum_{A \in 2^{\Omega_X}, \omega_i \subset A} \frac{1}{|A|} m_{\Omega}^{\Omega_X} \{Y_l\} (A) = 1 \quad (3.10)$$

En séparant $m_{\Omega}^{\Omega_X} \{Y_l\} (\emptyset^{un})$ et $m_{\Omega}^{\Omega_X} \{Y_l\} (\emptyset^*)$, la relation 3.10 est équivalente à :

$$\frac{1}{1 - m_{\Omega}^{\Omega_X} \{Y_l\} (\emptyset^{un})} \sum_{i=1}^n \sum_{A \in 2^{\Omega_X}, \omega_i \subset A} \left[\frac{1}{|A|} m_{\Omega}^{\Omega_X} \{Y_l\} (A) + m_{\Omega}^{\Omega_X} \{Y_l\} (\emptyset^*) \right] = 1$$

En posant

$$BetP_{\Omega_X^*} \{Y_l\} (\omega^*) = \frac{m_{\Omega}^{\Omega_X} \{Y_l\} (\emptyset^*)}{1 - m_{\Omega}^{\Omega_X} \{Y_l\} (\emptyset^{un})}$$

et sachant que $2^{\Omega_X^*} = 2^{\Omega_X} \cup \{\omega^*\}$ nous avons

$$\sum_{i=1}^n BetP_{\Omega_X} \{Y_l\} (\omega_i) + BetP_{\Omega_X^*} \{Y_l\} (\omega^*) = 1$$

Dans notre méthode, l'intégralité de la masse de croyance sur \emptyset^* est attribuée à ω^* . Par ailleurs, puisque nous supposons que chaque mesure Y_l ne provient de la perception que d'un seul objet, la cardinalité de ω^* est égale à 1. Par conséquent, aucune indétermination n'intervient dans le calcul de la probabilité pignistique et la décision prise sur Ω_X^* n'est pas dégradée.

3.1.5 Conclusion

Nous avons présenté une nouvelle méthode de gestion du conflit dans un contexte de sources d'informations "faiblement non fiable". Cette contrainte signifie que la majorité du conflit est la conséquence d'une hypothèse manquante dans le cadre de discernement. Nous avons isolé de l'ensemble des masses de croyance obtenu après la phase de combinaison, la partie de la masse de croyance indiquant la présence d'un tel évènement. Nous avons alors proposé de modifier la granularité du cadre de discernement en y intégrant une hypothèse supplémentaire à laquelle nous avons attribuée la masse de croyance retenue. Cette méthode permet de réduire la masse de croyance conflictuelle et de garantir une meilleure intégrité des probabilités obtenues lors de la transformation pignistique.

3.2 Comparaison des méthodes décrites sous l'hypothèse d'un monde ouvert et celles décrites sous l'hypothèse du monde ouvert étendu

3.2.1 Equation des fonctions de croyance combinées

3.2.1.1 Hypothèse du monde ouvert étendu (MOE)

Les équations suivantes expriment de manière générale les masses de croyance résultantes de la combinaison de n fonctions de croyance initiales dans le cadre du monde ouvert étendu. Ces masses de croyance sont distribuées sur les hypothèses singletons $\{\omega_i\}$, l'hypothèse de fermeture $\{\omega^*\}$, l'ensemble vide \emptyset et la disjonction des hypothèses du cadre de discernement. Remarquons que ces dernières regroupent l'hypothèse de fermeture $\{\omega^*\}$. Pour plus de détails sur l'obtention de ces équations générales, nous pouvons orienter le lecteur intéressé aux travaux de Gruyer et Royère [T3, T6].

Masse de croyance résultante sur les singletons

$$m_{\cap}^{\Omega^*} \{Y_l\} (\omega_i) = m_i^{\Omega^*} \{Y_l\} (\omega_i) \prod_{j=1, j \neq i}^n (1 - m_j^{\Omega^*} \{Y_l\} (\omega_j)) \quad (3.11)$$

$$m_{\cap}^{\Omega^*} \{Y_l\} (\omega^*) = \prod_{j=1}^n m_j^{\Omega^*} \{Y_l\} (\bar{\omega}_j)$$

Masse de croyance résultante sur l'ensemble vide

$$m_{\cap}^{\Omega^*} \{Y_l\} (\emptyset) = 1 - \left[\prod_{j=1}^n (1 - m_j^{\Omega^*} \{Y_l\} (\omega_j)) + \sum_{j=1}^n m_j^{\Omega^*} \{Y_l\} (\omega_j) \prod_{k=1, k \neq j}^n (1 - m_k^{\Omega^*} \{Y_l\} (\omega_k)) \right]$$

*Masse de croyance résultante sur l'union d'une hypothèse et de l'hypothèse ω^**

$$m_{\cap}^{\Omega^*} \{Y_l\} (\omega_i \cup \omega^*) = m_i^{\Omega^*} \{Y_l\} (\Omega_X^*) \times \prod_{j=1, j \neq i}^n m_j^{\Omega^*} \{Y_l\} (\bar{\omega}_j)$$

*Masse de croyance résultante sur toutes les unions de 2 à $n-1$ hypothèses ne comprenant pas ω^**

$$m_{\cap}^{\Omega_X^*} \{Y_l\} (\omega_i \cup \dots \cup \omega_k \cup \omega^*) = m_i^{\Omega_X^*} \{Y_l\} (\Omega_X^*) \dots m_k^{\Omega_X^*} \{Y_l\} (\Omega_X^*) \\ \times \prod_{j=1, j \neq i, \dots, j \neq k}^n m_j^{\Omega_X^*} \{Y_l\} (\bar{\omega}_j)$$

*Masse de croyance résultante sur l'union de toutes les hypothèses de Ω_X^**

$$m_{\cap}^{\Omega_X^*} \{Y_l\} (\Omega_X^*) = \prod_{i=1}^n m_i^{\Omega_X^*} \{Y_l\} (\Omega_X^*)$$

3.2.1.2 Hypothèse du monde ouvert avec gestion partielle du conflit

Les équations suivantes expriment de manière générale les masses de croyance résultantes de la combinaison de n fonctions de croyance initiales dans le cas de notre méthode que nous désignerons par **NM** dans la suite. Les équations permettent d'obtenir le résultat de la combinaison des fonctions de croyance dans le monde ouvert après l'étape de redistribution sur Ω_X^* . Tout comme dans la section précédente, nous présentons les équations selon les hypothèses ou les disjonctions d'hypothèses sur lesquelles les masses de croyance sont placées.

Masse de croyance résultante sur les singletons

$$m_{\cap}^{\Omega_X^*} \{Y_l\} (\omega_i) = m_i^{\Omega_X^*} \{Y_l\} (\omega_i) \prod_{j=1, j \neq i}^n (1 - m_j^{\Omega_X^*} \{Y_l\} (\omega_j)) \\ + m_i^{\Omega_X^*} \{Y_l\} (\Omega_X^*) \prod_{j=1, j \neq i}^n m_j^{\Omega_X^*} \{Y_l\} (\bar{\omega}_j)$$

*Masse de croyance résultante sur l'hypothèse ω^**

$$m_{\cap}^{\Omega_X^*} \{Y_l\} (\omega^*) = \prod_{j=1}^n m_j^{\Omega_X^*} \{Y_l\} (\bar{\omega}_j)$$

Masse de croyance résultante sur l'ensemble vide

$$m_{\cap}^{\Omega_X^*} \{Y_l\} (\emptyset) = 1 - \\ \left[\prod_{j=1}^n (1 - m_j^{\Omega_X^*} \{Y_l\} (\omega_j)) + \sum_{j=1}^n m_j^{\Omega_X^*} \{Y_l\} (\omega_j) \prod_{k=1, k \neq j}^n (1 - m_k^{\Omega_X^*} \{Y_l\} (\omega_k)) \right]$$

Masse de croyance résultante sur l'union de 2 à $n - 1$ hypothèses

$$m_{\cap}^{\Omega_X^*} \{Y_l\} (\omega_i \cup \dots \cup \omega_k) = m_i^{\Omega_X} \{Y_l\} (\Omega_X) \dots m_k^{\Omega_X} \{Y_l\} (\Omega_X) \\ \times \prod_{j=1, j \neq i, \dots, j \neq k}^n m_j^{\Omega_X} \{Y_l\} (\bar{\omega}_i)$$

Masse de croyance résultante sur l'union de toutes les hypothèses de Ω_X

$$m_{\cap}^{\Omega_X} \{Y_l\} (\Omega_X) = \prod_{i=1}^n m_i^{\Omega_X} \{Y_l\} (\Omega_X)$$

Les différences entre toutes ces équations résident dans l'utilisation ou non de l'hypothèse de fermeture w^* . Ces différences sont les suivantes :

- Les masses de croyance affectées aux hypothèses $w_i \cup w^*$ dans le MOE sont affectées aux hypothèses singletons w_i avec NM.
- Les masses affectées aux unions d'hypothèses dans les deux méthodes sont les mêmes, ces dernières regroupent l'hypothèse de fermeture w^* dans le MOE contrairement à NM. Ceci est également vrai pour la masse affectée à l'union de toutes les hypothèses du cadre de discernement, représentant le cadre de discernement lui-même (Ω_X^* pour le MOE et Ω_X pour NM).

3.2.2 Les fonctions de croyance initiales

Dans la suite de ce chapitre, nous présentons une étude comparative des résultats obtenus par une méthode opérant sous l'hypothèse du monde ouvert étendu et ceux donnés par notre solution.

Sur le même exemple de la section 3.1.3.2 où le cadre de discernement Ω_X est composé de deux hypothèses, nous définissons les deux fonctions de croyance initiales $m_i^{\Omega_X} \{Y_l\} (A)$ sur la base de deux méthodes de modélisation fondées sur une distance. Ces deux fonctions de croyance proposent des distributions de masses de croyance sur les trois parties suivantes : $\{\omega_i\}$, $\{\bar{\omega}_i\} = \{\omega_j : j \in [1, \dots, n] \text{ et } j \neq i\}$ et $\{\Omega_X\}$. Nous utilisons les fonctions de croyance initiales proposées par Rombaut [C65] et Gruyer [T3]. Le jeu de masses proposé par Gruyer diffère de celui proposé par Rombaut en ce qu'il interdit d'avoir simultanément une masse de croyance initiale non nulle sur $\{\omega_i\}$ et $\{\bar{\omega}_i\}$.

Dans [C65], Rombaut propose de construire les fonctions de croyance comportant trois

éléments focaux de la manière suivante :

$$\begin{cases} m_i^{\Omega_X} \{Y_l\} (\omega_i) &= \alpha\beta \\ m_i^{\Omega_X} \{Y_l\} (\bar{\omega}_i) &= \alpha(1 - \beta) \\ m_i^{\Omega_X} \{Y_l\} (\Omega_X) &= 1 - \alpha \end{cases} \quad (3.12)$$

Dans [T3], Gruyer propose de construire les fonctions de croyance comportant deux éléments focaux de la manière suivante :

$$\begin{cases} m_i^{\Omega_X} \{Y_l\} (\omega_i) &= \alpha\beta \\ m_i^{\Omega_X} \{Y_l\} (\Omega_X) &= 1 - \alpha\beta \\ m_i^{\Omega_X} \{Y_l\} (\bar{\omega}_i) &= 0 \end{cases} \quad \text{ou} \quad \begin{cases} m_i^{\Omega_X} \{Y_l\} (\bar{\omega}_i) &= \alpha\beta \\ m_i^{\Omega_X} \{Y_l\} (\Omega_X) &= 1 - \alpha\beta \\ m_i^{\Omega_X} \{Y_l\} (\omega_i) &= 0 \end{cases} \quad (3.13)$$

Dans ces deux jeux, le coefficient α quantifie la fiabilité de chaque source d'informations et β est une mesure de distance entre la mesure Y_l et le vecteur X_i représentant l'objet connu O_i (cf.2.17, 2.18 et 2.19). Ces deux méthodes sont des approches "orientées distance" que nous utiliserons dans le chapitre suivant.

Pour la phase de comparaison, nous utilisons des jeux de masses dont les valeurs sont identiques aussi bien dans le contexte du monde ouvert étendu que dans celui du monde ouvert non exhaustif. Toutefois, la différence réside dans la définition des éléments focaux de la fonction de croyance résultante de la combinaison.

3.2.3 Cas 1 : Contexte de décision identique pour les deux méthodes

Les deux jeux de masses de croyance utilisées sont les suivantes :

Jeu de masses obtenu par le générateur de Rombaut

$$\begin{cases} m_1^{\Omega_X} \{Y_1\} (\omega_1) &= m_1^{\Omega_X^*} \{Y_1\} (\omega_1) &= 0.8141 \\ m_1^{\Omega_X} \{Y_1\} (\bar{\omega}_1) &= m_1^{\Omega_X^*} \{Y_1\} (\bar{\omega}_1) &= 0.0859 \\ m_1^{\Omega_X} \{Y_1\} (\Omega_X) &= m_1^{\Omega_X^*} \{Y_1\} (\Omega_X^*) &= 0.1 \end{cases}$$

et

$$\begin{cases} m_2^{\Omega_X} \{Y_1\} (\omega_2) &= m_2^{\Omega_X^*} \{Y_1\} (\omega_2) &= 0.049 \\ m_2^{\Omega_X} \{Y_1\} (\bar{\omega}_2) &= m_2^{\Omega_X^*} \{Y_1\} (\bar{\omega}_2) &= 0.851 \\ m_2^{\Omega_X} \{Y_1\} (\Omega_X) &= m_2^{\Omega_X^*} \{Y_1\} (\Omega_X^*) &= 0.1 \end{cases}$$

Jeu de masses obtenu par le générateur de Gruyer

$$\begin{cases} m_1^{\Omega_X} \{Y_1\} (\omega_1) = m_1^{\Omega_X^*} \{Y_1\} (\omega_1) = 0.5891 \\ m_1^{\Omega_X} \{Y_1\} (\Omega_X) = m_1^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 0.4109 \end{cases}$$

et

$$\begin{cases} m_2^{\Omega_X} \{Y_1\} (\bar{\omega}_2) = m_2^{\Omega_X^*} \{Y_1\} (\bar{\omega}_2) = 0.7145 \\ m_2^{\Omega_X} \{Y_1\} (\Omega_X) = m_2^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 0.2855 \end{cases}$$

Ces deux jeux de masses soutiennent l'hypothèse ω_1 .

3.2.3.1 Combinaison des jeux de masses

Les tableaux 3.3(a,b) et 3.4(a,b) représentent les matrices de croyance obtenues après combinaison des deux jeux de masses initiaux respectivement dans le contexte du monde ouvert étendu et dans celui de notre méthode. Les graphes des figures 3.1 et 3.2 illustrent les probabilités pignistiques obtenues à partir de ces matrices de croyances. Les probabilités pignistiques fournies par notre méthode ont été obtenues après avoir redistribué les masses de croyance sur le cadre de discernement Ω_X^* .

$m_{1 \cap 2}^{\Omega_X^*}$		$m_2^{\Omega_X^*}$		
		ω_2	Ω_X^*	$\bar{\omega}_2$
$m_1^{\Omega_X^*}$	ω_1	0	0.1682	0.4209
	Ω_X^*	0	0.1173	0.2936
	$\bar{\omega}_1$	0	0	0

(a)

$m_{1 \cap 2}^{\Omega_X^*}$		$m_2^{\Omega_X^*}$		
		ω_2	Ω_X^*	$\bar{\omega}_2$
$m_1^{\Omega_X^*}$	ω_1	0.04	0.0814	0,6928
	Ω_X^*	0.0049	0.01	0.085
	$\bar{\omega}_1$	0.0042	0.0086	0.0731

(b)

TAB. 3.3 : Masses de croyance résultantes obtenues dans le contexte du monde ouvert étendu dans le cas du jeu de masses de (a) Gruyer et (b) Rombaut.

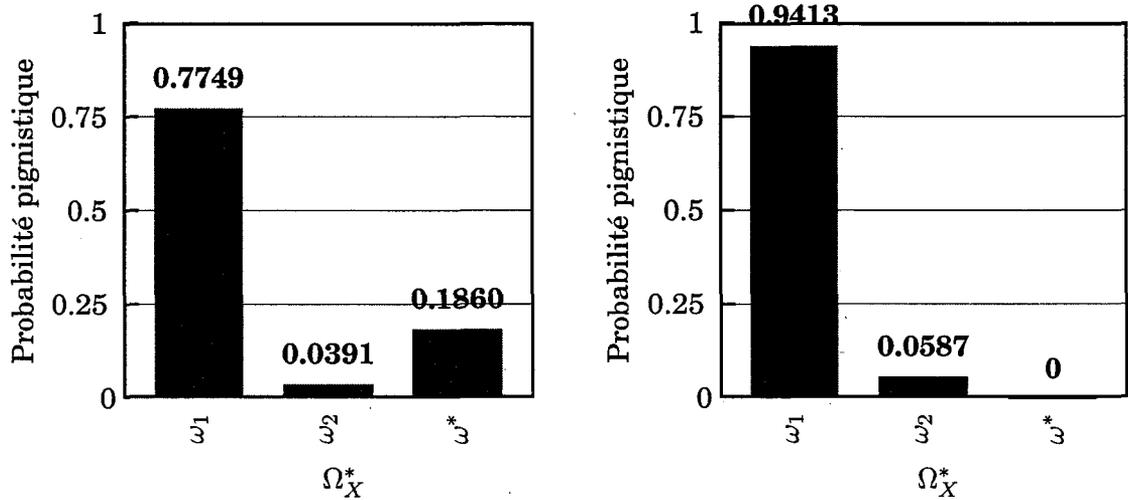
$m_{1 \cap 2}^{\Omega_X}$		$m_2^{\Omega_X}$		
		ω_2	Ω_X	$\bar{\omega}_2$
$m_1^{\Omega_X}$	ω_1	0	0.1682	0.4209
	Ω_X	0	0.1173	0.2936
	$\bar{\omega}_1$	0	0	0

(a)

$m_{1 \cap 2}^{\Omega_X}$		$m_2^{\Omega_X}$		
		ω_2	Ω_X	$\bar{\omega}_2$
$m_1^{\Omega_X}$	ω_1	0.04	0.0814	0,6928
	Ω_X	0.0049	0.01	0.085
	$\bar{\omega}_1$	0.0042	0.0086	0.0731

(b)

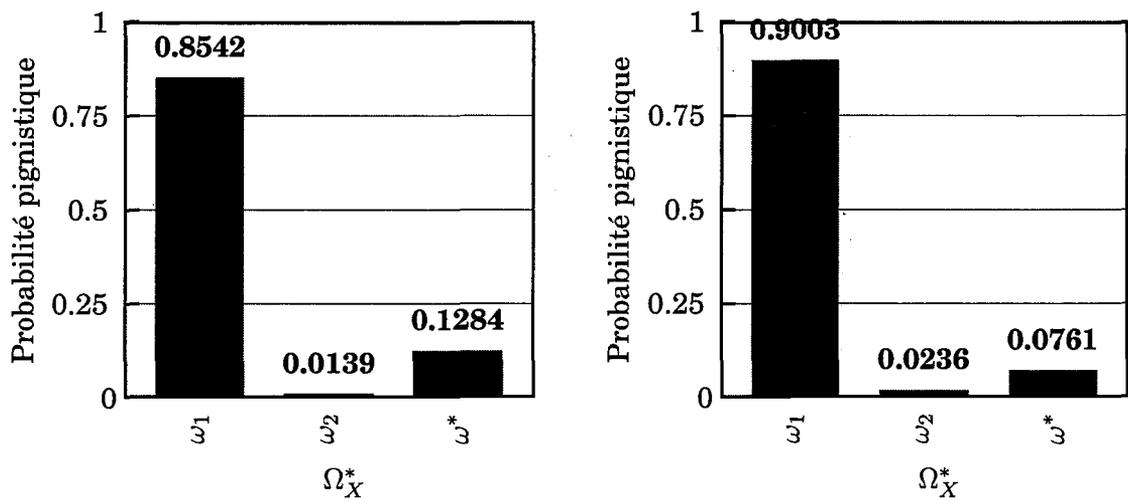
TAB. 3.4 : Masses de croyance résultantes obtenues dans le contexte de notre méthode dans le cas du jeu de masses de (a) Gruyer et (b) Rombaut.



(a) Contexte du monde ouvert étendu

(b) Contexte de notre méthode

FIG. 3.1 : Probabilités pignistiques obtenues sur le jeu de masses de Gruyer.



(a) Contexte du monde ouvert étendu

(b) Contexte de notre méthode

FIG. 3.2 : Probabilités pignistiques obtenues sur le jeu de masses de Rombaut.

3.2.3.2 Discussion

La différence entre le monde ouvert étendu et le monde ouvert non exhaustif (cf. les tableaux 3.1 et 3.2) réside dans la définition des éléments focaux dans lesquels intervient ω^* (cas du monde ouvert étendu). L'expression des masses de croyance initiales sur Ω_X^* engendre l'apparition de l'hypothèse ω^* dans la composition des sous-ensembles $\overline{\omega_1}$ et $\overline{\omega_2}$. Le résultat de la combinaison est donc moins précis. Cette imprécision est typique du monde ouvert étendu et n'apparaît pas dans le cadre du monde ouvert.

Lors de la transformation pignistique, l'imprécision a pour conséquence de transférer une partie de la masse de croyance sur l'élément ω^* . Ainsi, la probabilité que l'hypothèse ω^* soit vraie, c'est à dire que la mesure Y_1 soit un nouvel objet, est augmentée et la probabilité que les autres hypothèses soient vraies diminuent. Selon notre méthode, seule la masse de croyance affectée à \emptyset^* est prise en compte dans le calcul de la probabilité pignistique de ω^* . Comme nous pouvons le vérifier en comparant deux à deux les graphes des figures 3.1 et 3.2, la probabilité pignistique sur ω^* calculée sous l'hypothèse du monde ouvert étendu est inférieure à la probabilité pignistique obtenue avec notre méthode. Ces mêmes graphes mettent en évidence la tendance inverse en ce qui concerne la probabilité pignistique sur les autres hypothèses du cadre de discernement.

Dans ces deux exemples, une décision fondée sur le maximum de probabilité pignistique mène au même résultat quel que soit le jeu de masse de croyance initiale considéré : ω_1 est retenue comme l'hypothèse vraie et la mesure Y_1 est associée à l'objet O_1 . On vérifie aisément que la probabilité sur cette hypothèse est plus importante avec notre méthode ce qui augmente la fiabilité de la décision par rapport à la méthode du monde ouvert étendu.

3.2.4 Cas 2 : Les deux méthodes présentent des décisions différentes

Les cas présentés ci-après mettent en évidence des décisions différentes sur un même jeu de masses de croyance selon l'utilisation de l'une ou de l'autre méthode.

Les deux jeux de masses de croyance utilisés sont les suivants :

Jeu de masses No 1 obtenu par le générateur de Rombaut

$$\begin{cases} m_1^{\Omega_X} \{Y_1\} (\omega_1) = m_1^{\Omega_X^*} \{Y_1\} (\omega_1) = 0 \\ m_1^{\Omega_X} \{Y_1\} (\overline{\omega_1}) = m_1^{\Omega_X^*} \{Y_1\} (\overline{\omega_1}) = 0 \\ m_1^{\Omega_X} \{Y_1\} (\Omega_X) = m_1^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 1.0 \end{cases} \quad (3.14)$$

et

$$\begin{cases} m_2^{\Omega_X} \{Y_1\} (\omega_2) = m_2^{\Omega_X^*} \{Y_1\} (\omega_2) = 0.25 \\ m_2^{\Omega_X} \{Y_1\} (\bar{\omega}_2) = m_2^{\Omega_X^*} \{Y_1\} (\bar{\omega}_2) = 0.25 \\ m_2^{\Omega_X} \{Y_1\} (\Omega_X) = m_2^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 0.5 \end{cases} \quad (3.15)$$

Jeu de masses No 2 obtenu par le générateur de Rombaut

$$\begin{cases} m_1^{\Omega_X} \{Y_1\} (\omega_1) = m_1^{\Omega_X^*} \{Y_1\} (\omega_1) = 0 \\ m_1^{\Omega_X} \{Y_1\} (\bar{\omega}_1) = m_1^{\Omega_X^*} \{Y_1\} (\bar{\omega}_1) = 1.0 \\ m_1^{\Omega_X} \{Y_1\} (\Omega_X) = m_1^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 0 \end{cases}$$

et

$$\begin{cases} m_2^{\Omega_X} \{Y_1\} (\omega_2) = m_2^{\Omega_X^*} \{Y_1\} (\omega_2) = 0.2 \\ m_2^{\Omega_X} \{Y_1\} (\bar{\omega}_2) = m_2^{\Omega_X^*} \{Y_1\} (\bar{\omega}_2) = 0.3 \\ m_2^{\Omega_X} \{Y_1\} (\Omega_X) = m_2^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 0.5 \end{cases}$$

Jeu de masses No 1 obtenu par le générateur de Gruyer

$$\begin{cases} m_1^{\Omega_X} \{Y_1\} (\bar{\omega}_1) = m_1^{\Omega_X^*} \{Y_1\} (\bar{\omega}_1) = 1 \\ m_1^{\Omega_X} \{Y_1\} (\Omega_X) = m_1^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 0 \end{cases}$$

et

$$\begin{cases} m_2^{\Omega_X} \{Y_1\} (\bar{\omega}_2) = m_2^{\Omega_X^*} \{Y_1\} (\bar{\omega}_2) = 0.25 \\ m_2^{\Omega_X} \{Y_1\} (\Omega_X) = m_2^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 0.75 \end{cases}$$

Jeu de masses No 2 obtenu par le générateur de Gruyer

$$\begin{cases} m_1^{\Omega_X} \{Y_1\} (\bar{\omega}_1) = m_1^{\Omega_X^*} \{Y_1\} (\bar{\omega}_1) = 0 \\ m_1^{\Omega_X} \{Y_1\} (\Omega_X) = m_1^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 1.0 \end{cases}$$

et

$$\begin{cases} m_2^{\Omega_X} \{Y_1\} (\bar{\omega}_2) = m_2^{\Omega_X^*} \{Y_1\} (\bar{\omega}_2) = 0.25 \\ m_2^{\Omega_X} \{Y_1\} (\Omega_X) = m_2^{\Omega_X^*} \{Y_1\} (\Omega_X^*) = 0.75 \end{cases}$$

Dans les deux cas présentés dans la section 3.2.3, il est relativement aisé à un expert de décider de la véracité de l'une ou de l'autre des deux hypothèses. Or la définition d'une vérité terrain est indispensable si nous désirons comparer le comportement des deux méthodes et déterminer leur point de fonctionnement et de dysfonctionnement. Dans les quatre cas présentés ci-dessus, le choix est rendu plus complexe par la présence d'au minimum une fonction de croyance soutenant de manière importante l'ignorance. L'étude dé-

critère dans la suite de cette partie montre comment l'ignorance portée par une ou plusieurs sources d'informations modifie la décision proposée par les deux méthodes et vérifie leur équivalence.

3.2.4.1 Combinaison des jeux de masses

Les tableaux 3.5 représentent les matrices de croyance obtenues après combinaison deux à deux des huit jeux de masses initiaux dans le contexte du monde ouvert étendu. Ces quatre tableaux restent valables pour la combinaison dans le contexte de notre méthode simplement en remplaçant Ω_X^* par Ω_X . Les graphes en barres des figures 3.3 et 3.4 illustrent les probabilités pignistiques obtenues à partir des précédentes matrices de croyances. Les probabilités pignistiques obtenues avec notre méthode ont été obtenues après avoir redistribué les masses de croyance sur le cadre de discernement Ω_X^* .

$m_{1 \cap 2}^{\Omega_X^*}$		$m_2^{\Omega_X^*}$		
		ω_2	Ω_X^*	$\bar{\omega}_2$
$m_1^{\Omega_X^*}$	ω_1	0	0	0
	Ω_X^*	0.25	0.5	0.25
	$\bar{\omega}_1$	0	0	0

(a)

$m_{1 \cap 2}^{\Omega_X^*}$		$m_2^{\Omega_X^*}$		
		ω_2	Ω_X^*	$\bar{\omega}_2$
$m_1^{\Omega_X^*}$	ω_1	0	0	0
	ω_X^*	0	0	0
	$\bar{\omega}_1$	0.2	0.5	0.3

(b)

$m_{1 \cap 2}^{\Omega_X^*}$		$m_2^{\Omega_X^*}$		
		ω_2	Ω_X^*	$\bar{\omega}_2$
$m_1^{\Omega_X^*}$	ω_1	0	0	0
	Ω_X^*	0	0	0
	$\bar{\omega}_1$	0	0.75	0.25

(c)

$m_{1 \cap 2}^{\Omega_X^*}$		$m_2^{\Omega_X^*}$		
		ω_2	Ω_X^*	$\bar{\omega}_2$
$m_1^{\Omega_X^*}$	ω_1	0	0	0
	ω_X^*	0	0.75	0.25
	$\bar{\omega}_1$	0	0	0

(d)

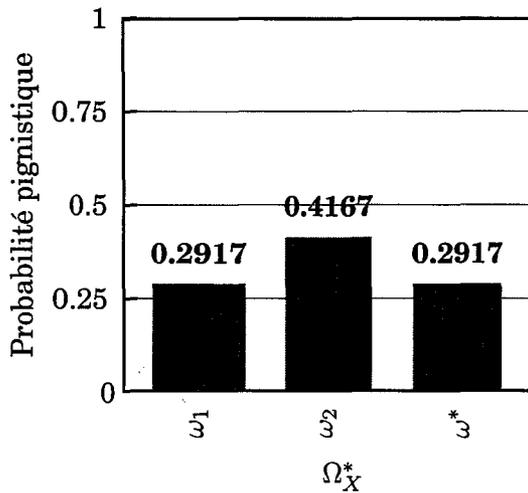
TAB. 3.5 : Masses de croyance résultantes pour le monde ouvert étendu dans le cas du (a) jeu de masses No 1 de Rombaut, (b) jeu de masses No 2 de Rombaut, (c) jeu de masses No 1 de Gruyer et (d) jeu de masses No 2 de Gruyer.

3.2.4.2 Discussion

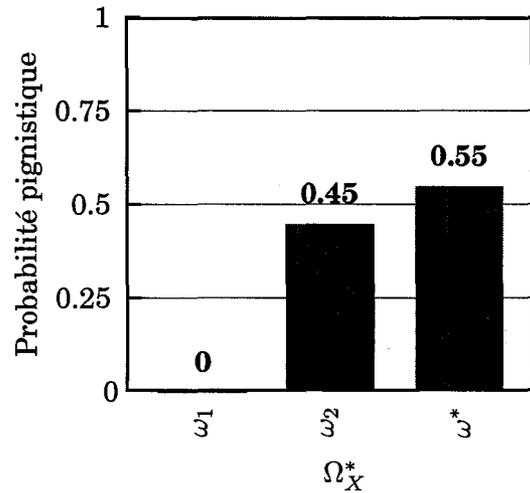
Jeu No 1 de Rombaut

Ce jeu de masses nous indique que deux sources se prononcent sur une situation avec une ignorance importante. Par conséquent, il paraît peu judicieux de prendre une décision en faveur de l'une ou l'autre des hypothèses du cadre de discernement.

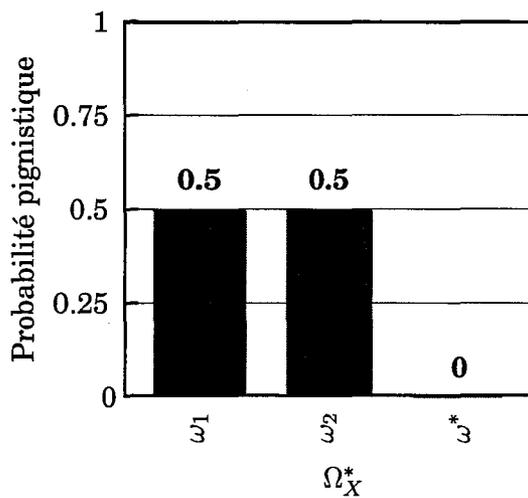
La méthode du MOE (graphe 3.3(a)) indique que la solution est l'hypothèse ω_2 . Nous constatons que la croyance sur ω^* provient de la masse portée par l'ignorance et par l'hypo-



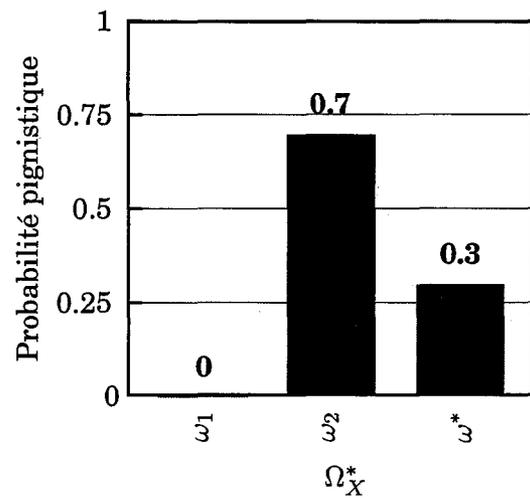
(a) Jeu de masse No 1 et contexte du monde ouvert étendu



(b) Jeu de masse No 2 et contexte du monde ouvert étendu

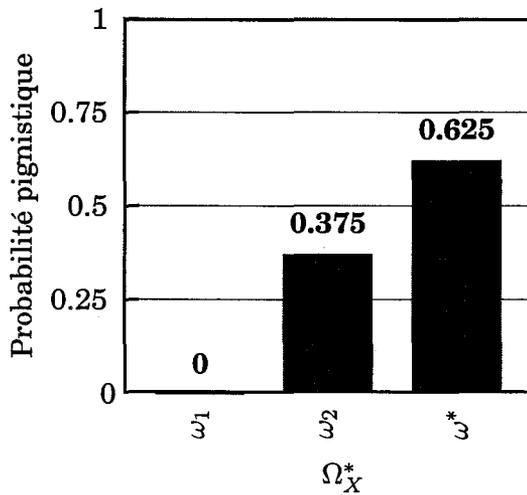


(c) Jeu de masse No 1 et contexte de notre méthode

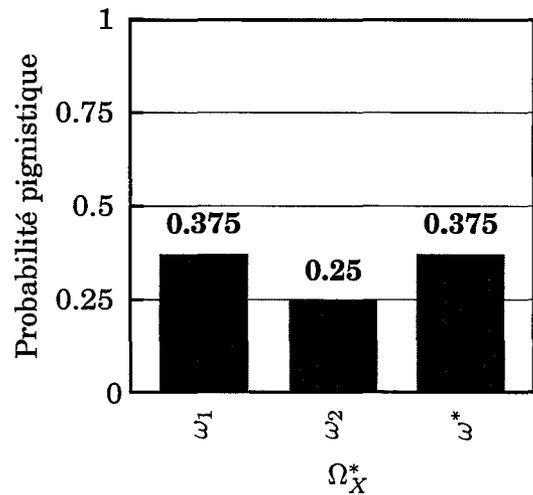


(d) Jeu de masse No 2 et contexte de notre méthode

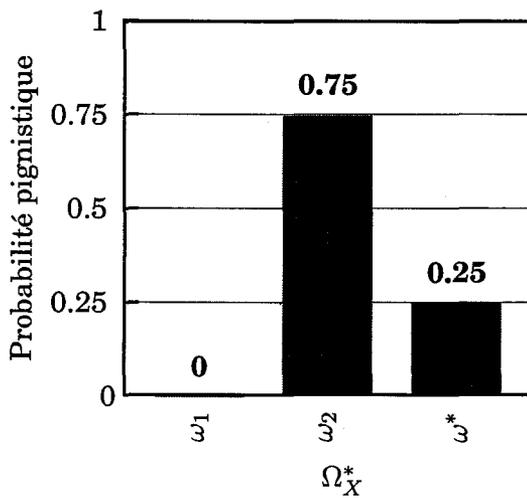
FIG. 3.3 : Probabilités pignistiques obtenues sur le jeu de masses de Rombaut.



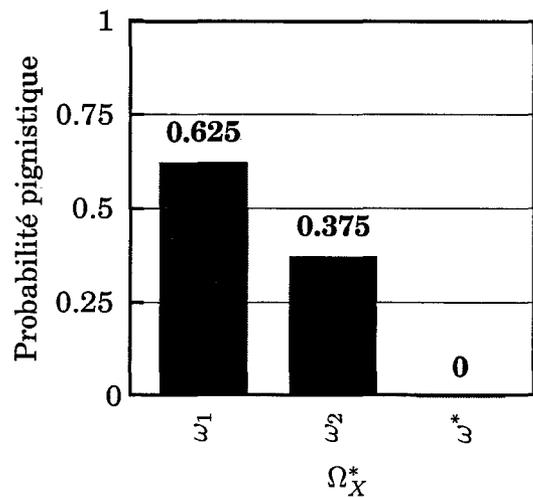
(a) Jeu de masse No 1 et contexte du monde ouvert étendu



(b) Jeu de masse No 2 et contexte du monde ouvert étendu



(c) Jeu de masse No 1 et contexte de notre méthode



(d) Jeu de masse No 2 et contexte de notre méthode

FIG. 3.4 : Probabilités pignistiques obtenues sur le jeu de masses de Gruyer.

thèse $\overline{\omega_2}$. Dans ce cas de figure, la NM reste indécise (graphe 3.3(c)) et tant que la fonction de croyance m_1 sera nulle sur $\overline{\omega_1}$, l'hypothèse ω^* ne peut pas être une solution. La probabilité pignistique sur ω^* n'a que peu d'importance lors de la phase de décision. En effet, tant que la masse de croyance sur $\overline{\omega_1}$ est nulle, le maximum de probabilité ne peut jamais lui être attribué. En présence d'une telle distribution de croyances, la NM préconise l'indécision, ce qui semble plus logique.

Jeu No 2 de Rombaut

Nous sommes en présence d'une source qui soutient sans détour l'hypothèse $\overline{\omega_1}$, c'est-à-dire la proposition ω_2 . La seconde source fait preuve d'une ignorance importante bien qu'elle soutienne davantage ω_1 . Nous sommes dans une situation particulièrement délicate. En considérant qu'à coup sûr ω_1 ne peut pas être la solution, deux solutions restent alors possibles : une décision en faveur de la création d'une nouvelle cible ou en faveur de l'hypothèse ω_2 .

La méthode du MOE (graphe 3.3(b)) soutient ω^* . La NM indique que la solution est l'hypothèse ω_2 . Cette solution peut paraître incorrecte : en effet, puisque m_1 soutient l'hypothèse $\overline{\omega_1}$ et m_2 soutient plus $\overline{\omega_2}$ qu'elle ne soutient ω_2 , la solution devrait être ω^* . En ce sens, la méthode du MOE semble être la meilleure. Toutefois, il apparaît une croyance non négligeable sur l'ignorance. Il faut remarquer que la NM parvient à une conclusion identique à celle de la méthode MOE pour $m_2^{\Omega^*}(\Omega_X^*) \leq 0.3$ et $m_2^{\Omega^*}(\overline{\omega_2}) \geq 0.5$. La méthode MOE est plus favorable à la création d'une nouvelle cible que la méthode NM, cette dernière favorisant une hypothèse déjà existante.

Jeu No 1 de Gruyer

Ce jeu de masse a les mêmes propriétés que le jeu de masse No2 de Rombaut. L'analyse des graphes 3.1(a) et 3.1(c) permet d'arriver à la même conclusion : la méthode du MOE est plus favorable à la création d'une nouvelle cible que la méthode NM qui favorise une hypothèse déjà existante. Grâce à cet exemple, nous pouvons ajouter que la distribution des probabilités pignistiques fournies par la méthode NM reste la même quelle que soit la répartition des masses de croyance de la source 2 entre ω_2 et Ω_X^* .

Jeu No 2 de Gruyer

Ce jeu de masses met en évidence deux sources dont l'ignorance est très importante. Le peu d'information disponible est placée sur l'hypothèse ω_1 par la source 2. Par conséquent, si une décision doit être prise, elle devrait l'être en faveur de cette hypothèse mais en tenant compte du contexte d'ignorance dans lequel sont placées les deux sources.

La méthode du MOE (graphe 3.4(b)) reste neutre. La méthode de NM propose une décision en faveur de l'hypothèse ω_1 , ce qui semble être la meilleure solution au regard de la remarque faite au paragraphe précédent. Dans la méthode du MOE, le contexte d'ignorance se traduit par une probabilité non nulle sur l'hypothèse ω^* : il s'agit d'une conséquence directe du passage au niveau pignistique dans le contexte du monde ouvert étendu. Notre méthode impose que l'hypothèse ω^* ne soit soutenue que lorsque toutes les sources se prononcent sur le contraire de l'hypothèse pour laquelle elles sont spécialisées.

3.3 Conclusions

Comme nous l'avons décrit dans le chapitre 2, le suivi multi-objet est exprimé selon un formalisme crédibiliste. La méthode que nous proposons diffère de celle proposée par Rombaut, Gruyer et Royère en ce qu'elle ne fait pas l'hypothèse d'un monde ouvert étendu.

Comme le préconise Smets, notre méthode adopte l'hypothèse d'un monde ouvert dont le cadre de discernement n'est pas exhaustif. Nous avons supposé que les sources d'informations peuvent entrer en conflit mais toujours en considérant que la masse conflictuelle est engendrée majoritairement par des hypothèses manquantes. Nous avons recours à une procédure de redistribution d'une partie de la masse conflictuelle afin de diminuer la masse conflictuelle finale et de réduire les méfaits de la normalisation lors du passage au niveau pignistique.

Nous avons comparé notre méthode avec celle développée dans le contexte du monde ouvert étendu. Les deux méthodes présentent des comportements différents notamment lorsque les sources opèrent dans un contexte d'ignorance important. En plaçant la décision au niveau pignistique, la méthode du MOE tend à soutenir davantage l'hypothèse ω^* . Dans notre contexte applicatif, cela s'exprime par une corrélation entre la part d'ignorance et la création d'une nouvelle classe. Dans notre méthode, la part d'ignorance portée par les sources n'a aucune influence sur la décision de créer ou non une classe supplémentaire. Seule l'existence d'une part de conflit spécifique permet une telle décision. Contrairement à la méthode du monde ouvert étendu, notre méthode permet de gérer de manière indépendante la masse de croyance attribuée à l'ignorance.

Chapitre 4

Application au suivi audio et vidéo des individus

Dans ce chapitre, nous présentons une étude menée afin de valider le processus de gestion de conflit sous l'hypothèse du monde ouvert présenté dans le chapitre 3. Les résultats présentés sont obtenus dans le cadre d'une application d'audio et de vidéo surveillance. De telles applications s'insèrent traditionnellement dans le domaine de l'interaction Homme/Machine et ont été abondamment étudiées dans le contexte de la communication interactive. Les avancées ont été atteintes grâce à la multiplicité et aux performances (sensibilité, miniaturisation) des capteurs et grâce à la puissance toujours grandissante des calculateurs.

L'utilisation conjointe de capteurs audio et vidéo a pour objectif d'obtenir une description plus fidèle et plus robuste de l'environnement dans lequel des objets évoluent. Cet enrichissement est rendu possible car les deux signaux sont le résultat de la perception de deux phénomènes physiques différents dont il est possible d'extraire des informations complémentaires et redondantes. En effet, un objet occulté, c'est-à-dire non visible par une caméra, peut toujours être détecté par un microphone s'il émet un signal sonore. Inversement, un objet qui n'émet aucun signal sonore ou dont le signal n'est pas perceptible, peut être détecté par une caméra. Ainsi, à chaque instant, toute entité présente dans la scène peut être décrite par des caractéristiques audio et/ou vidéo, à la condition qu'elle soit perçue grâce à l'un ou l'autre des deux modes.

A ce jour, très peu d'études ont été menées sur l'utilisation conjointe des capteurs sonores et visuels pour améliorer la sécurité dans les enceintes de transport. L'analyse des signaux audio issus de ces environnements apparaît pour la première fois dans le projet PRISMATICA avec pour objectif de détecter de manière automatique des cris de détresse tels que "au secours", "help", et "stop thief" [C47]. Plus récemment, des recherches menées dans les projets TESS ("Transport, Espace et Société"), EVAS ("Etude de système de Vidéo et Audio Surveillance Sans fil") et SAMSIT ("Système d'Analyse de Médias pour une

Sécurité Intelligente dans les Transports publics") proposent de détecter des situations critiques à partir de l'analyse automatique de l'ambiance sonore de l'enceinte d'un véhicule de transport [C66, C85]. Modéliser et détecter une situation critique à partir de l'analyse des images fournies par un système de surveillance est un domaine de recherche qui a connu et connaît encore une activité importante. Les environnements rencontrés dans le domaine du transport sont particulièrement complexes : présence de nombreux objets, richesse du signal sonore, occultations etc. Implémenter des modèles audio et vidéo de chaque situation à détecter est une activité très récente et il est encore difficile d'imaginer de quelle manière l'analyse conjointe des deux signaux peut être élaborée.

Nous décrivons dans ce chapitre la tâche qui consiste à suivre un ensemble d'individus en déplacement dans un environnement surveillé par un ensemble de caméras et de microphones. C'est une tâche essentielle dans un processus de reconnaissance de situations potentiellement critiques car il permet de décrire l'interaction des individus entre eux et avec leur environnement. Ce chapitre se décompose en trois sections principales.

Dans la première section, nous positionnons par rapport à l'existant nos travaux relatifs à l'analyse conjointe des signaux vidéo et des signaux audio pour la localisation et le suivi des individus. Cette présentation ne se veut pas exhaustive. Elle cherche à préciser les éléments clés liés à notre application. La deuxième section de ce chapitre est dédiée à la description du système proposé et plus spécifiquement des capteurs logiques vidéo et audio et du module de fusion en charge de la tâche de suivi/association. La troisième section est consacrée à la présentation des résultats.

4.1 Le suivi audio et vidéo : un état de l'art

De nombreux travaux sur l'utilisation jointe des données audio et vidéo pour la localisation et le suivi d'individus ont été menés dans le domaine de la téléconférence interactive. Ces travaux peuvent être classés en deux catégories. La première catégorie regroupe les méthodes qui exploitent séquentiellement les deux modalités. Ces méthodes procèdent en deux temps. Dans un premier temps, elles utilisent une de ces deux modalités pour détecter d'éventuels individus présents dans la scène. Dans un second temps, l'autre modalité est introduite pour confirmer et accomplir le suivi des individus détectés [A67, C88, C86]. Les méthodes de la seconde catégorie utilisent simultanément les deux modalités afin de localiser et suivre les individus [A77, C56, C2, A4, A57, A25, C1, L19, C67, C16, A7].

Afin de localiser et de suivre des individus, ces méthodes exploitent uniquement l'infor-

mation de position estimée séparément à partir des signaux audio et vidéo. Les caractéristiques acoustiques et visuelles qui peuvent être extraites de ces deux modalités ne sont pas exploitées. Or, dans les situations où les individus sont proches les uns des autres, l'information de position ne suffit pas à assurer la distinction de la trajectoire de chacun d'entre eux. Ces méthodes présentent donc une robustesse insuffisante dans ces situations.

Dans les travaux de [C4, C98], les auteurs captent des informations de position et des caractéristiques modales afin de suivre et d'identifier les locuteurs parmi les participants à une session de vision conférence. L'ensemble des informations sont fusionnées selon le formalisme probabiliste. Dans les travaux de Neti et al. [C57, C39] et de Hershey et al. [C36], les auteurs souhaitent discriminer les différents locuteurs à partir de la synchronisation temporelle entre les signaux audio et vidéo. Dans [L10, C21, C26], Fisher exploite les statistiques jointes des deux modalités.

Dans les travaux précédemment cités, la tâche de suivi/association des objets est formulée dans un cadre probabiliste ou est assurée grâce des outils de filtrage particulière. Dans [C38, C37], l'application se situe toujours dans le contexte de la vidéoconférence. Son objectif est de déterminer l'identité et la position du locuteur à partir de la direction du regard des participants situés autour de la table et de la position angulaire des sources acoustiques en présence. Dans ce travail, les auteurs formulent ce problème en adoptant un formalisme crédibiliste. Toutefois, le problème du suivi/Association des mesures aux individus n'est pas posé puisque ces derniers sont assis autour d'une table et ne sont pas amenés à se déplacer dans la scène.

Notre application se positionne dans le contexte du suivi d'un ensemble d'individus observés et écoutés par un ensemble de caméras et microphones. Contrairement à une grande majorité des travaux présentés dans la littérature, lors de leur déplacement, les individus sont susceptibles :

- d'être occultés par d'autres individus ou par les objets présents dans la scène et
- d'entrer ou de sortir de la scène à tout moment.

Nous caractérisons chaque individu par sa position et par un ensemble d'attributs modaux. Ces caractéristiques sont extraites à chaque instant des images et du signal sonore. L'association des caractéristiques décrivant chaque individu avec celles de l'ensemble des cibles indique les individus qui sont encore présents dans la scène et ceux qui sont sortis et ceux qui sont entrés dans la scène. Par ailleurs, cette association conforte la présence de ceux qui sont restés dans la scène observée. Nous avons adopté le formalisme crédibiliste décrit

dans le chapitre précédent afin de le valider sur chacune de ces trois situations.

4.2 Le système proposé

Comme nous l'avons décrit dans le chapitre 2, des recherches ont été menées afin de décrire le suivi multi-objet selon un formalisme crédibiliste. Ces recherches s'insèrent dans le cadre d'une application de suivi de mobiles visant à reconstruire l'environnement proche d'un véhicule. La méthode que nous proposons diffère de celle proposées par Rombaut, Gruyer et Royère en ce qu'elle ne fait pas l'hypothèse d'un monde ouvert étendu. Toutefois, la démarche que nous adoptons reste similaire à certains égards.

La procédure de suivi que nous proposons est fondée sur l'architecture présentée à la figure 4.1. L'environnement est observé par deux capteurs logiques audio et vidéo. Chaque capteur fournit un ensemble d'objets qui sont caractérisés par un vecteur d'attributs modaux et un vecteur de position. Soit $Y^a(k) = (Y_l^a(k))_{l=1,\dots,p_k}$ l'ensemble des p_k couples des vecteurs $\{P_{l,k}^a, A_{l,k}^a\}$ où $P_{l,k}^a$ et $A_{l,k}^a$ représentent respectivement le vecteur de position et le vecteur d'attributs acoustiques du $l^{i\text{eme}}$ objet perçu. Soit $Y^v(k) = (Y_l^v(k))_{l=1,\dots,r_k}$ l'ensemble des r_k couples des vecteurs $\{P_{l,k}^v, A_{l,k}^v\}$ où $P_{l,k}^v$ et $A_{l,k}^v$ représentent respectivement le vecteur de position et le vecteur d'attributs vidéo du $l^{i\text{eme}}$ objet perçu.

Soient n_k le nombre d'objets connus à l'instant k et $X(k) = (X_l(k))_{l=1,\dots,n_k}$ l'ensemble des triplets $\{P_{l,k}^X, A_{l,k}^{X,a}, A_{l,k}^{X,v}\}$ où $P_{l,k}^X, A_{l,k}^{X,a}$ et $A_{l,k}^{X,v}$ représentent respectivement le vecteur position, le vecteur d'attributs vidéo et le vecteur d'attributs audio du $l^{i\text{eme}}$ objet connu.

L'objectif de la phase d'association est de relier chaque couple composant les deux ensembles de vecteurs d'observation vidéo et audio à l'un triplet caractérisant un des objets connus. Afin de tenir compte des propriétés dynamiques de l'état de chaque objet connu (changement de son apparence, changement de sa position, changement de son intensité acoustique etc.), ce dernier est prédit pour l'instant d'association. Nous notons $X(k-1, k) = (X_l(k-1, k))_{l=1,\dots,n_k}$ cet ensemble des prédictions. Ainsi, le système initie simultanément deux procédures d'association "observations/cibles prédites" : la première entre $X(k-1, k) = (X_l(k-1, k))_{l=1,\dots,n_k}$ et $(Y_l^a(k))_{l=1,\dots,p_k}$ et la seconde entre $X(k-1, k) = (X_l(k-1, k))_{l=1,\dots,n_k}$ et $(Y_l^v(k))_{l=1,\dots,r_k}$. $X_l(k-1, k)$ représente le triplet constitué par le vecteur de position et les deux vecteurs d'attributs modaux du $l^{i\text{eme}}$ objet connu dont l'état a été prédit pour l'instant k à partir des données connues à l'instant $k-1$.

Les résultats des associations sont alors interprétés lors de la phase de maintenance des cibles afin de définir, à l'instant k , les nouveaux objets, les objets disparus et les objets dont

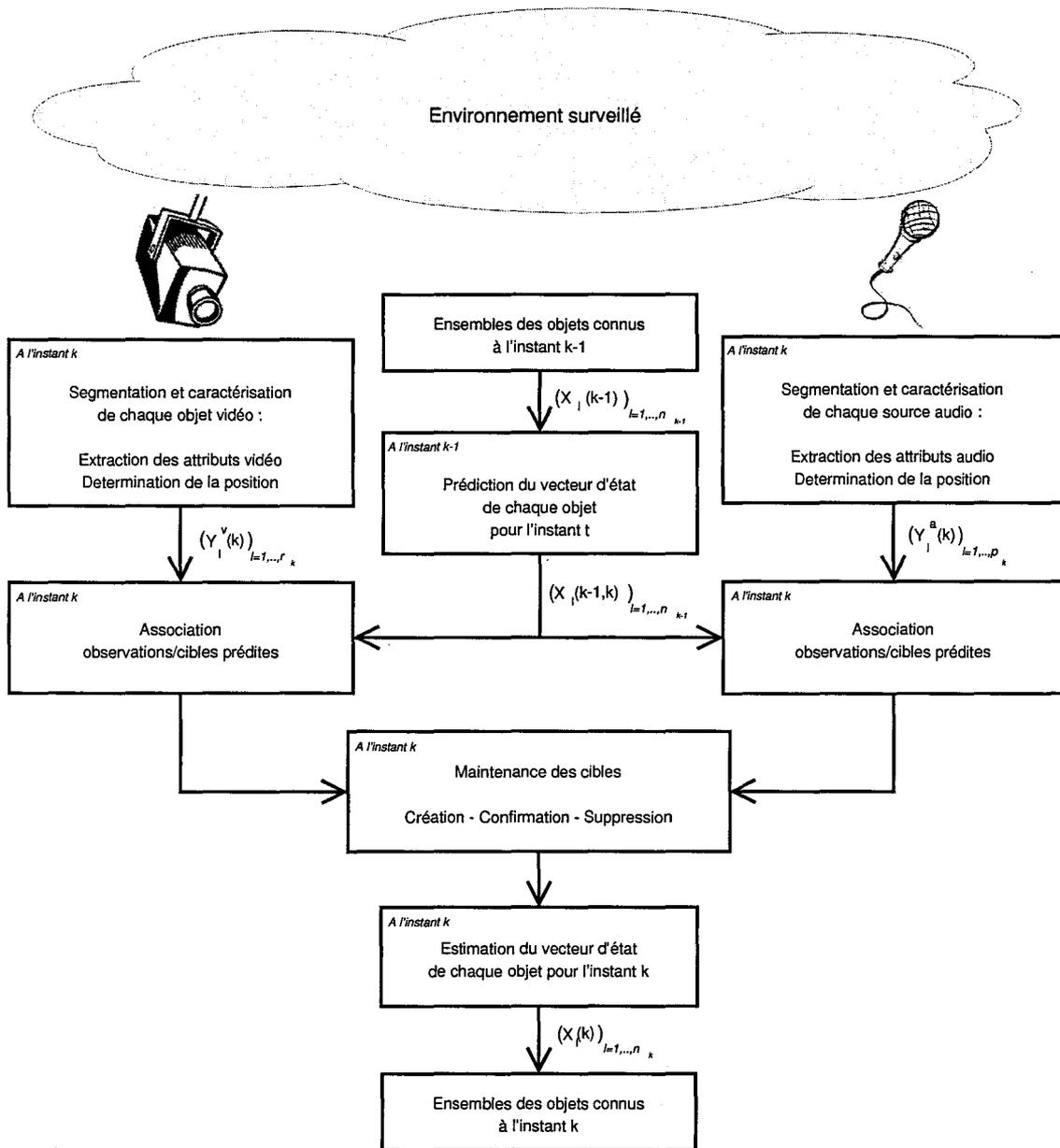


FIG. 4.1 : Architecture proposée du système de suivi/association audio-vidéo.

l'existence est confirmée. A ce même instant, une estimation du triplet $X_i(k)$ de vecteurs caractérisant chaque objet vient clore l'itération correspondant à une série de mesures. Remarquons que suite à l'étape de maintenance des cibles, le nombre n_k d'objets connus peut être différent du nombre n_{k-1} d'objets connus à l'instant précédent.

Dans la suite du chapitre, nous considérerons que la position de chaque objet évolue de manière linéaire et que son processus d'évolution peut être modélisé par un filtre de Kalman classique. Nous supposerons que les paramètres modaux sont invariants dans le temps. Cette hypothèse représente une contrainte forte sur les performances du système

de suivi-association dont il faudra tenir compte lors de la phase d'évaluation. Toutefois, la méthode proposée peut être adaptée afin d'intégrer des modèles plus complexes qui tiennent compte par exemple du changement d'apparence des objets ou encore de la variabilité du contenu fréquentiel du signal acoustique. Par ailleurs, afin d'alléger la notation, nous omettrons l'indice k sur les couples vecteurs $\{P_l^a, A_l^a\}$ et $\{P_l^v, A_l^v\}$ et sur les triplets $\{P_l^X, A_l^{X,a}, A_l^{X,v}\}$.

Dans la suite de cette section nous décrivons les fonctionnalités des deux capteurs logiques. Nous présentons tout particulièrement les attributs modaux que nous avons choisi d'extraire de chaque signal afin de caractériser chaque objet détecté et la manière dont l'information de position peut être extraite du signal acoustique. Dans un second temps, nous détaillons les éléments fonctionnels des deux modules d'association observation/cibles prédites. Nous mettons davantage l'accent sur la manière dont sont générées les fonctions de croyance initiales, étape essentielle dans l'application de la théorie des fonctions de croyance.

4.2.1 Le capteur logique vidéo

Le capteur logique vidéo est vu comme un ensemble de caméras dont les images sont analysées par un ensemble d'algorithmes dont les objectifs sont de détecter et de localiser chaque individu et de le distinguer des autres. Comme l'illustre la figure 4.2, ce capteur fournit l'ensemble des individus détectés, chaque individu l étant associé à un couple de vecteurs $\{P_l^v, A_l^v\}$.

La résolution des problèmes de détection, de localisation et de modélisation d'un individu dans une séquence vidéo dépasse l'objectif des travaux de cette thèse. Depuis de plusieurs années, des travaux sont menés sur ces problématiques. Bien que des solutions existent, elles montrent encore peu de robustesse vis-à-vis des contraintes fixées par le contexte d'une application réelle.

Dans la littérature, les problèmes de détection et de localisation d'un individu dans une séquence d'images sont traditionnellement vus sous l'angle de la détection et du suivi de son visage [A86, A89, A31] et de celui de la détection et du suivi de sa silhouette [A27, C24, C59]. Le choix de l'une ou l'autre de ces deux approches est guidé par les objectifs de l'application, la complexité de l'environnement observé et les caractéristiques ou la configuration d'installation des caméras vidéo du système.

Dans notre application, la tâche de détection d'un visage semble plus en adéquation

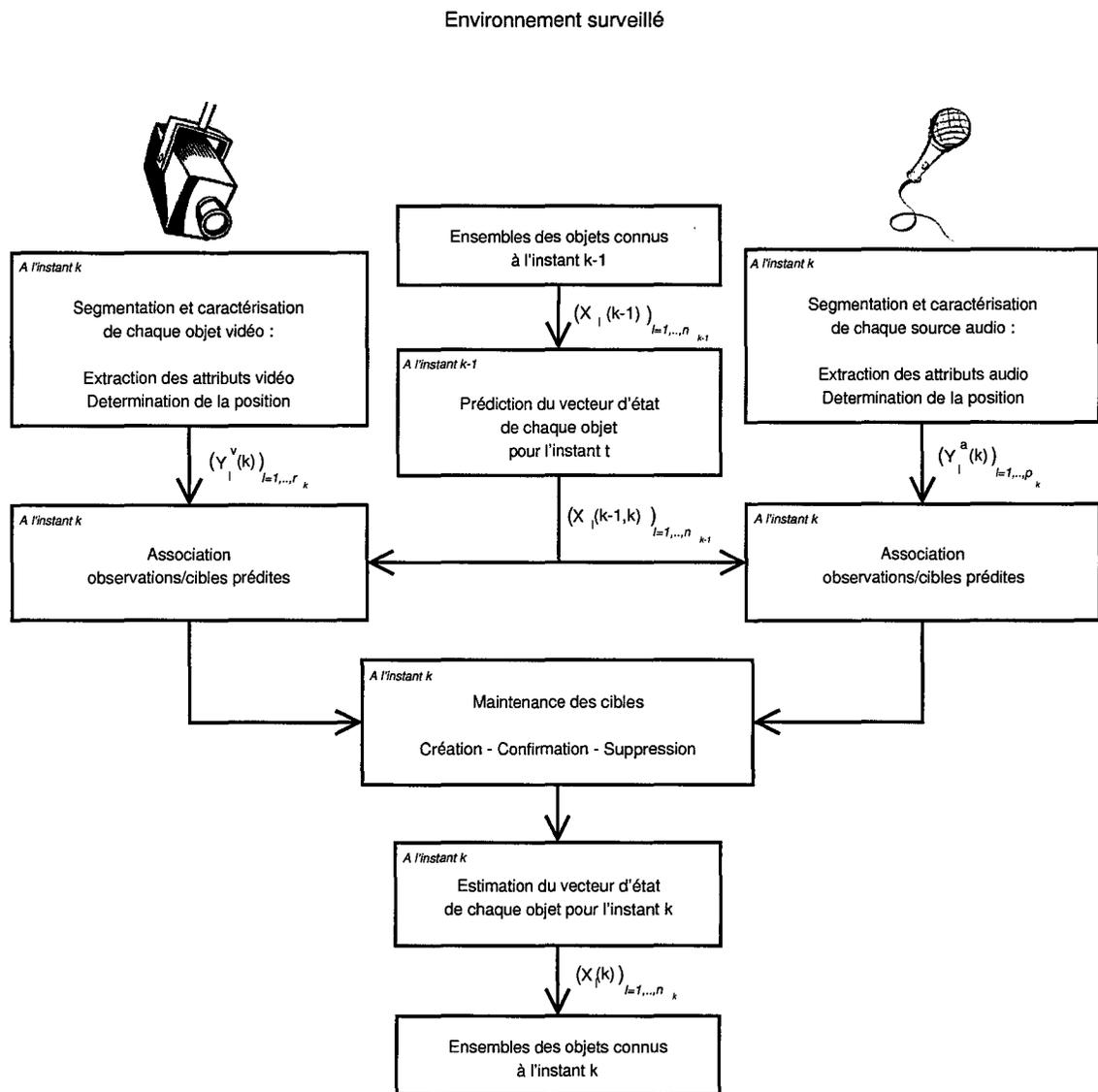


FIG. 4.1 : Architecture proposée du système de suivi/association audio-vidéo.

l'existence est confirmée. A ce même instant, une estimation du triplet $X_l(k)$ de vecteurs caractérisant chaque objet vient clore l'itération correspondant à une série de mesures. Remarquons que suite à l'étape de maintenance des cibles, le nombre n_k d'objets connus peut être différent du nombre n_{k-1} d'objets connus à l'instant précédent.

Dans la suite du chapitre, nous considérerons que la position de chaque objet évolue de manière linéaire et que son processus d'évolution peut être modélisé par un filtre de Kalman classique. Nous supposerons que les paramètres modaux sont invariants dans le temps. Cette hypothèse représente une contrainte forte sur les performances du système

ces méthodes diffèrent fortement selon que le fond des images est fortement structuré ou non.

Approches fondées sur apprentissage et reconnaissance des caractéristiques invariantes du visage

Ces approches sont fondées sur l'extraction d'éléments caractéristiques du visage. Elles opèrent en deux temps. Dans un premier temps, les éléments caractéristiques du visage tels que la bouche, les yeux, les sourcils et le nez sont localisés. Le visage est alors caractérisé par les relations géométriques qui relient ces différents éléments [C69, L21, A18, A51].

Certaines méthodes proposent d'utiliser des caractéristiques robustes vis-à-vis d'une variation de l'orientation du visage et du fond des images, telles que la couleur et la texture [C3, C93, C91, A75, C43, C64, C63, C46, C13, A80, A76, C79, A88]. Certains auteurs proposent de combiner simultanément plusieurs caractéristiques telles que la taille, la forme et la couleur de la peau afin d'augmenter la fiabilité de la détection [A76, C79, C46, A78, C91, C23].

Parmi ces méthodes, certaines font appel à une phase d'apprentissage afin de concevoir un modèle que l'utilisateur cherchera à retrouver dans les images. Ce modèle est défini à partir de l'analyse d'un ensemble d'images qui présentent une variété de visages à détecter ou à identifier. La phase d'apprentissage peut être réalisée grâce à un réseau de neurones [C42, L15, L14], à une analyse en composante principale [A41, A92, C84] ou à des méthodes à noyaux (Support Vector Machine) [A33].

Approches fondées sur la mise en correspondance modèle / images

Puisque la forme globale du visage et de ses éléments caractéristiques sont connues a priori, certaines méthodes proposent d'utiliser un modèle 2D prédéfini du visage afin de faire la détection [A58, A62, C83]. L'image est alors mise en correspondance avec le modèle retenu. Le terme de corrélation est calculé et minimisé afin de réduire l'écart entre les deux images. L'optimisation est réalisée sur les paramètres régissant la physionomie et la géométrie du visage. Les modèles peuvent être simples (modèle elliptique) ou plus complexe (modèle 2D à facettes).

Les modèles actifs permettent de mener simultanément une segmentation et un suivi des visages [A43, C97, C96, C92]. Dans [C52, C53] nous avons proposé une méthode d'extraction et de suivi de visages dans des images en niveau de gris en combinant les méthodes d'optimisation génétiques et les contours actifs. Les algorithmes génétiques sont introduits afin de pallier le problème d'initialisation des contours et d'obtenir une méthode d'extrac-

tion complètement automatique.

4.2.1.2 L'attribut visuel

Nous avons choisi de caractériser chaque visage par l'histogramme couleur. De nombreuses études ont montré son efficacité [C93, C95, C30, C29]. Dans certaines limites d'utilisation, l'histogramme couleur de plusieurs visages d'un même individu varie faiblement. Il présente cependant des différences importantes entre deux individus distincts. Il est robuste aux occultations partielles, invariant à certaines rotations, insensible aux faibles variations d'échelles et est peu coûteux en temps de calcul [A68, C95, A19].

L'histogramme couleur d'une région de l'image représente la distribution des couleurs des pixels contenus dans cette région. La couleur d'un pixel peut être définie dans plusieurs espaces de représentation. Dans la littérature, de nombreux travaux de recherche ont proposé de modéliser un visage grâce à un histogramme couleur déterminé sur différents espaces de représentation de la couleur : l'espace RGB dans [A75, C43], l'espace RGB normalisé dans [A82, C64, C63, C46], l'espace YCrCb dans [C13, A80], l'espace HSV (HSI) dans [A76, C79] et l'espace YIQ dans [A88]. Le choix de l'espace de représentation dépend des conditions d'acquisition des séquences d'images vidéo, en particulier des conditions d'éclairage de la scène.

Quel que soit l'espace couleur utilisé, il est nécessaire de faire précéder l'étape de calcul de l'histogramme par une procédure de quantification de la couleur. Celle-ci a pour rôle de déterminer, dans l'espace de représentation considéré, les couleurs les plus représentatives de l'image. Cette étape permet de réduire le nombre de couleurs nécessaires à la représentation de l'information sans générer de dégradations importantes.

Pour une région de l'image composée de N pixels et quantifiée par k couleurs (c_1, c_2, \dots, c_k) , l'histogramme associé H est un vecteur à k composantes $(h_{c_1}, h_{c_2}, \dots, h_{c_k})$. La valeur de chaque composante h_{c_i} représente le nombre de pixels de la région de l'image dont la couleur est c_i . Cet histogramme vérifie l'égalité $\sum_{i=1}^k h_{c_i} = N$.

Dans notre application nous avons retenu l'espace colorimétrique RVB qui a donné de bons résultats durant l'étape de segmentation des images acquises dans notre laboratoire. La base de données images a été créée à partir d'une caméra couleur tri-ccd. La phase de quantification suit une démarche similaire à celle proposée par Heckbert [C34] dans laquelle l'auteur supprime les 3 bits de poids faible de chaque composante couleur.

4.2.2 Le capteur logique audio

Le capteur logique audio est vu comme une matrice de microphones dont les signaux sont analysés par un ensemble d'algorithmes dont les objectifs sont de détecter et de localiser chaque individu et de le distinguer des autres. Ici chaque individu est considéré comme une source acoustique. Un schéma similaire à celui de la figure 4.2 est présenté à la figure 4.3. Ce capteur fournit l'ensemble des sources détectées, chacune étant associée à un couple de vecteurs $\{P_l^a, A_l^a\}$.

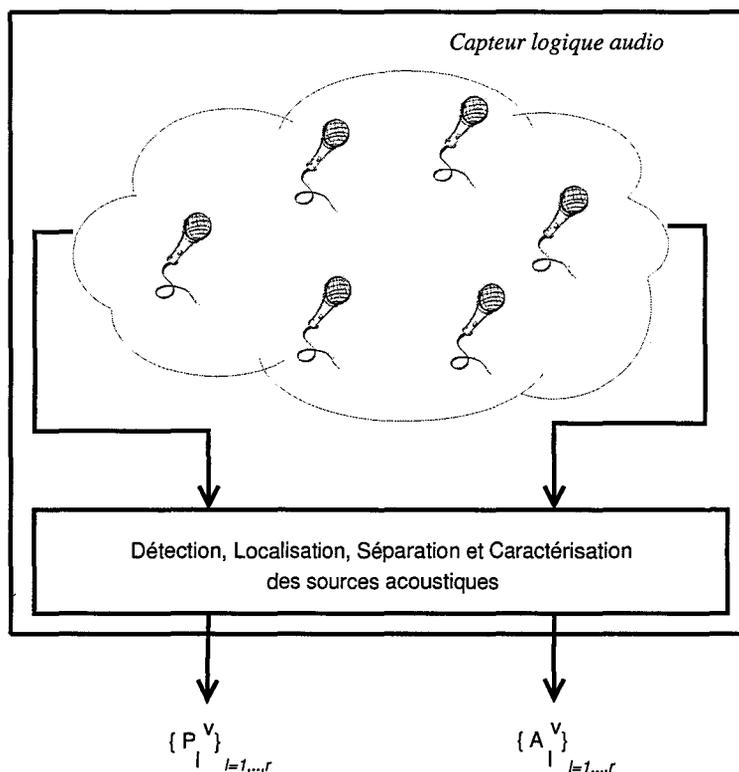


FIG. 4.3 : Capteur logique Audio.

Le signal perçu par un microphone est le mélange de toutes les sources actives au moment de l'acquisition, c'est-à-dire la contribution de tous les individus parlant à cet instant. Ainsi, obtenir l'ensemble des couples de vecteurs $\{\{P_l^a, A_l^a\}\}_{l=1,\dots,p}$ nécessite de résoudre les problèmes de détection, de localisation, de séparation et de modélisation de chaque source en présence. Comme pour le capteur logique vidéo, offrir une solution à chacun de ces problèmes dépasse l'objectif des travaux de cette thèse. Des travaux sont d'ores et déjà menés sur ces sujets depuis plus ou moins longtemps. Des solutions sont décrites dans la littérature mais elles montrent encore très peu de robustesse vis-à-vis des contraintes fixées par le contexte d'une application réelle.

Lors de la phase de validation, nous considèrerons que les tâches de localisation et de séparation des différentes sources sont réalisées. Nous nous attacherons uniquement à l'étape de modélisation de chacune d'entre elles. Toutefois, afin d'éclaircir le lecteur sur la finalité de chacune des tâches et les difficultés rencontrées, nous les présentons succinctement dans la suite de la section.

4.2.3 Localisation de sources sonores

L'étape de localisation a pour objectif de positionner chaque source sonore dans l'environnement dans lequel elle évolue. Généralement, la localisation de sources sonores exploite une matrice de microphones et la mesure de différence des temps d'arrivée (TDOA : Time Delays Of Arrival). Le TDOA caractérise le temps que met le son à parvenir à chaque microphone de la matrice. Une description détaillée de ces méthodes est faite dans [L4, A77] où les auteurs les ont classées en deux catégories. D'une part, les méthodes de localisation directes et d'autre part les méthodes indirectes.

Les méthodes de localisation directes sont fondées sur l'analyse de l'énergie du signal obtenu en additionnant tous les signaux perçus par chaque microphone. Cette stratégie est habituellement implantée grâce aux techniques dites de "filtrage directionnel" (steered filter) et de faisceau directionnel (steered beamforming) [A71, C78, C45].

Les méthodes de localisation indirectes procèdent en deux étapes. Dans un premier temps, les TDOA sont calculés en chaque microphone par des méthodes de corrélation. L'estimation de la position de chaque source est alors déterminée à partir de relation spatiale reliant les TDOA et les positions relatives des microphones [A5, A94, A67, C87, C70, A12, C80].

Habituellement les microphones sont placés loin de la source sonore à localiser. La localisation de source se heurte alors au problème causé par les nombreuses réflexions de la source sonore sur les surfaces limitant l'enceinte dans laquelle elle se trouve. Afin de contourner cette difficulté, l'hypothèse de l'existence d'un trajet direct est très souvent retenue, notamment dans le cadre des techniques de localisation directe. Dans notre application, à ce problème s'ajoute celui lié à la présence d'une autre source sonore qui ne serait pas engendrée par un individu et que nous ne voulons pas modéliser.

4.2.4 Séparation des sources

La séparation de sources désigne la tâche visant à reconstruire les signaux provenant de plusieurs sources, à partir de l'analyse des signaux fournis par une matrice de microphones. Cette séparation est dite aveugle (BSS : Blind Source Separation) lorsqu'elle ne fait appel à aucune connaissance a priori, que ce soit sur les sources en présence, ou sur le milieu de propagation.

Le signal fourni par un microphone est équivalent à un mélange des signaux produits par toutes les sources. Les mélanges sont dits convolutifs. Nous distinguons plusieurs types de mélanges selon les propriétés du milieu de propagation : les mélanges linéaires, les mélanges non linéaires, ces dernières intégrant ou non des retards.

Les recherches les plus récentes s'appuient sur les méthodes d'analyse en composantes indépendantes (ACI) [A29, C35, A15, A16]. Elle font les hypothèses suivantes :

- la matrice de microphones doit comporter au minimum autant de capteurs que de sources à séparer,
- les sources doivent être indépendantes,
- une seule source doit être gaussienne,
- le nombre de sources à séparer est connu.

Dans la pratique, le problème de séparation de sources est rendu difficile, notamment à causes des situations suivantes :

- les sources sont corrélées dans le temps, donc non indépendantes,
- les sources sont en mouvement (ex. des locuteurs en déplacement),
- le nombre de sources est inconnu et variable [A93, A17],
- le nombre de capteurs est inférieur au nombre de sources à séparer [A48, C8].

Dans le cas particulier de la séparation de plusieurs locuteurs, les performances atteintes par les méthodes sont loin d'être satisfaisantes pour envisager leur implantation dans une application réelle, en particulier dans un milieu transport [A90, C89, C49, C61, A60, A59].

4.2.5 Modélisation des sources

Il s'agit de caractériser chaque source sonore séparée afin de les distinguer les unes des autres. Cette fonction est fondamentale dans les applications d'identification et de vérification du locuteur :

- L'identification du locuteur consiste à reconnaître un locuteur parmi un ensemble fini

de locuteurs.

- La vérification du locuteur consiste à vérifier si un locuteur est bien celui qu'il prétend être.

L'identification et la vérification peuvent alors être "dépendantes du texte" (effectuées sur la base d'un texte imposé) ou "indépendantes du texte" (indépendamment du contenu des phrases prononcées). Notre application se place dans le contexte de l'identification de locuteur car nous cherchons à associer un ensemble de sources détectées à un ensemble de sources connues. Par ailleurs, l'association doit être indépendante du contenu lexical de la parole produite par chaque individu.

La figure 4.4 présente la structure classique d'un système d'identification du locuteur. Un tel système comporte trois étapes : une étape de paramétrisation, une étape de classification et une étape de décision. L'étape de paramétrisation consiste à extraire un ensemble de paramètres acoustiques du signal vocal à identifier. Lors de l'étape de classification, cet ensemble de paramètres est comparé au modèle de référence lié à chaque source à identifier, chaque modèle étant déterminé sur la base du même lot de paramètres. Un ensemble d'indices de similarité est calculé puis analysé au sein du module de décision afin de trouver l'identité du locuteur.

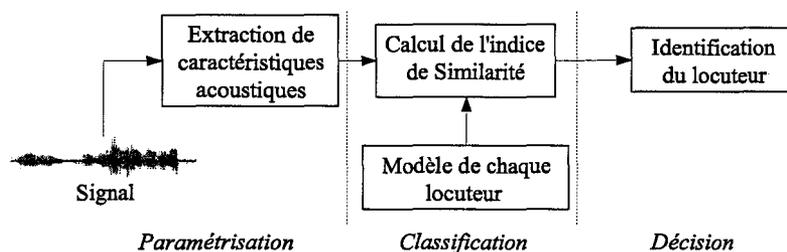


FIG. 4.4 : Structure d'un système d'identification du locuteur.

L'étape de décision est intégrée au module d'association observations/cibles qui est décrit dans la section 4.2.6. Les deux sous-sections suivantes présentent les étapes de paramétrisation et de classification.

4.2.5.1 Etape de paramétrisation

Cette étape a pour objectif d'extraire un ensemble de paramètres du signal à analyser afin de réduire la quantité d'informations à traiter lors de la phase de classification. Le signal est découpé en fenêtres d'analyse sur laquelle le signal est considéré comme stationnaire. Cette fenêtre est appelée trame. Elle est définie comme une succession d'échantillons

et sa durée varie de 10 à 40ms. A l'issue de cette étape chaque trame est caractérisée par un vecteur de paramètres.

Les paramètres acoustiques décrits dans la littérature [A81, A49, A50, L1, L13, L18] peuvent être classés en deux classes :

- Les paramètres temporels et fréquentiels.
- Les paramètres issus de la modélisation du système auditif humain et du mécanisme de production de la parole.

La première classe regroupe les termes suivants : l'énergie, le taux de passage par zéro, la fréquence centroïde (Frequency Centroid), la largeur de bande (Bandwidth), la fréquence de roll-off (roll-off frequency) et le flux spectral (spectral flux) et le pitch.

Parmi les paramètres issus de la modélisation du système auditif humain nous pouvons citer : les coefficients PLP (Prédiction Linéaire Perceptive ou Perceptual linear prediction), les coefficients RASTA-PLP, (RelAtive SpecTrAl Technique), les coefficients MFC (Mel Frequency Cepstral). Parmi les paramètres inspirés du mécanisme de production de la parole, nous pouvons citer les coefficients LPC (Codage Prédicatif Linéaire ou Linear Predictive Coding).

Cette liste de paramètres acoustiques n'est pas exhaustive. Le lecteur trouvera des informations détaillées sur les différents paramètres dans [L1, L13].

Dans notre travail, notre choix s'est porté sur les coefficients LPC auxquels nous avons associé un terme d'énergie. Certains auteurs ont montré l'efficacité d'une telle combinaison dans des applications d'identification du locuteur [L5, A3, C90, A49, A14]. Les coefficients LPC sont caractéristiques du conduit vocal d'un individu et donc du signal de parole qu'il produit. Ils encodent la fréquence fondamentale et les formants du signal de la parole qui permettent de faire la distinction entre plusieurs voix différentes.

Les coefficients LPC

Soit s le signal d'une source à analyser. Soit s_n la n^{ime} trame de s . Les coefficients LPC sont calculés sur s_n en considérant qu'à chaque instant k de la trame, l'échantillon de signal $s(k)$ peut être approché par la combinaison linéaire de p échantillons précédents :

$$s(k) \approx a_1 s(k-1) + a_2 s(k-2) + \dots + a_p s(k-p) \quad (4.1)$$

où les coefficients a_i sont appelés coefficients de prédiction d'ordre p . L'erreur de prédiction s'exprime alors par $e(k) = s(k) - (a_1 s(k-1) + a_2 s(k-2) + \dots + a_p s(k-p))$. Elle est qualifiée de résidu d'ordre p . Les coefficients de la prédiction linéaire sont estimés de façon à

minimiser la variance σ_e^2 de l'erreur de prédiction $e(k)$. La minimisation est équivalente à la résolution d'un système de p équations [A54, A53, A14, L13, L1]. Lors de notre phase de validation, nous utilisons les coefficients LPC d'ordre 12.

Le coefficient d'énergie

Des études ont montré que l'utilisation d'un terme d'énergie conjointement avec d'autres paramètres améliore considérablement les résultats [L5, L1]. Soit E_n l'énergie de la trame s_n donnée par la formule suivante :

$$E_n = \frac{1}{N} \sum_{i=1}^N s(i)^2 \quad (4.2)$$

où i indique le numéro de l'échantillon du signal dans la trame s_n et N est le nombre d'échantillons contenus dans s_n .

4.2.5.2 Etape de classification

L'étape de paramétrisation est équivalente à la projection du signal sur un espace vectoriel particulier. L'étape de classification a pour objectif d'analyser le contenu d'un vecteur afin de l'affecter à une classe dont le modèle est connu. Une classe est représentée par un ou plusieurs sous-espaces de l'espace vectoriel considéré. Modéliser une classe est donc équivalent à déterminer les frontières de ces sous-espaces vectoriels. Ces modèles sont conçus lors d'une phase d'apprentissage. Afin d'y parvenir, de nombreuses méthodes ont été décrites dans la littérature parmi lesquelles nous pouvons citer les méthodes neuronales [A6, C58], les méthodes à mélanges de gaussiennes (GMM : Gaussian Mixture Model) [A10, A11, T4, L5, T1], les méthodes à noyaux (SVM : Support Vector Machine) [C74, C25].

Nous avons adopté la stratégie proposée par Bimbot et Gish [A11, A28], car elle est simple à implanter et donne de bons résultats en présence de signaux à classer. Cette méthode fait l'hypothèse que les caractéristiques spectrales de la source suivent une loi gaussienne stationnaire au second ordre dans l'espace vectoriel considéré. Une classe est donc caractérisée par un vecteur moyenne μ et une matrice de covariance Σ . Soit $\{A_i\}_{i=1,\dots,m}$ l'ensemble d'apprentissage relatif à une source donnée. Les paramètres statistiques sont

définis par :

$$\mu = \frac{1}{m} \sum_{i=1}^m A_i \quad (4.3)$$

$$\Sigma = \frac{1}{m} \sum_{i=1}^m (\mu - A_i) \cdot (A_i - \mu)^T \quad (4.4)$$

où B^T est la matrice B transposée.

4.2.6 Le module d'association observations/cibles prédites

Ce module a pour objectif d'attribuer, à un instant k , chaque couple de vecteurs des ensembles $Y^a(k)$ et $Y^v(k)$ à un des triplets de l'ensemble $X(k-1, k)$. Rappelons les points suivants :

- $X(k-1, k)$ est l'ensemble des triplets dont les états ont été prédits pour l'instant k à partir des triplets connus à l'instant $k-1$. Ces triplets sont prédits par un filtre de Kalman classique.
- chaque élément l de $Y^a(k)$ est un couple de vecteurs $\{P_{l,k}^a, (\mu_{l,k}^a, \Sigma_{l,k}^a)\}$ où $\mu_{l,k}^a$ et $\Sigma_{l,k}^a$ sont respectivement un vecteur moyenne et une matrice de covariances acoustiques
- chaque élément l de $Y^v(k)$ est un couple de vecteurs $\{P_{l,k}^v, H_{l,k}^v\}$ où $H_{l,k}^v$ est un histogramme
- chaque objet connu l est modélisé par le triplet $X_l = \{P_{l,k}^X, H_{l,k}^X, (\mu_{l,k}^X, \Sigma_{l,k}^X)\}$ où $H_{l,k}^X$ est son histogramme et $(\mu_{l,k}^X, \Sigma_{l,k}^X)$ sont respectivement un vecteur moyenne et une matrice de covariances acoustiques

. Dans la suite nous éliminons l'indice k des notations afin de les alléger.

L'étape principale de ce module est le calcul des fonctions de croyance élémentaires. Ces fonctions sont fondées sur trois indices de similarité que nous présentons par la suite : un indice de similarité acoustique, un indice de similarité vidéo et un indice de similarité de position. Comme dans les précédents travaux de Rombaut, Gruyer et Royère, nous opérons conformément à la notion "des vues partielles du monde" mentionnées par Gruyer. Ainsi le module procède au calcul et à la combinaison des fonctions de croyance initiales en deux temps selon que le cadre de discernement est construit à partir des observations modales ou à partir des objets connus. Comme cela a été dit dans les chapitres précédents, cette démarche assure une représentation complète des relations possibles entre tous les objets manipulés et permet de gérer l'apparition et la disparition d'objets.

La suite de cette section est organisée de la manière suivante. Dans un premier temps, nous présentons les modifications que nous avons apporté à la méthode en complément de la stratégie de gestion du conflit. Dans un second temps, nous décrivons les trois mesures de similarité utilisées pour calculer les fonctions de croyance élémentaires.

4.2.6.1 Les modifications apportées

Les modifications apportées se situent essentiellement au niveau pignistique et plus particulièrement sur la manière de combiner les fonctions de probabilité pignistique. Nous décrivons ces modifications dans le cas particulier du module d'association entre les observations audio et les objets connus. Cette description reste valable dans le cas de l'association entre les observations vidéo et les objets connus.

Soit $\Omega_X = \{\omega_i\}_{i=1,\dots,n}$ le cadre de discernement construit sur la base des objets connus. Soit $m_i^{\Omega_X} \{Y_l^a\}$ la fonction de croyance élémentaire associée à l'objet connu O_i et définie pour chaque mesure Y_l^a et $m_{\cap}^{\Omega_X} \{Y_l^a\}$ la fonction de croyance obtenue après combinaison avec l'opérateur de Dempster pour relier chaque observation Y_l^a à chaque objet connu.

Soit $\Omega_{Y^a} = \{\omega_i^a\}_{i=1,\dots,p}$ le cadre de discernement construit sur la base des observations audio. Soit $m_i^{\Omega_{Y^a}} \{X_l\}$ la fonction de croyance élémentaire associée à l'observation Y_i^a et définie pour chaque objet connu X_l et $m_{\cap}^{\Omega_{Y^a}} \{X_l\}$ la fonction de croyance obtenue après combinaison avec l'opérateur de Dempster pour relier chaque objet X_l à chaque observation Y^a .

Comme proposé au chapitre précédent, les deux fonctions de croyance $m_{\cap}^{\Omega_X} \{Y_l^a\}$ et $m_{\cap}^{\Omega_{Y^a}} \{X_l\}$ sont respectivement distribuées sur les cadres de discernement Ω_X^* et $\Omega_{Y^a}^*$. Les deux fonctions de croyance $m_{\cap}^{\Omega_X^*} \{Y_l^a\}$ et $m_{\cap}^{\Omega_{Y^a}^*} \{X_l\}$ sont transformées afin d'obtenir les deux fonctions de probabilité pignistique $BetP_{\Omega_X^*} \{Y_l^a\}$ et $BetP_{\Omega_{Y^a}^*} \{X_l\}$. La décision est prise en analysant le résultat de la combinaison des ces deux fonctions. Dans [C55], les auteurs réalisent leur moyenne arithmétique. Nous proposons de faire la moyenne géométrique de ces deux fonctions de la manière suivante :

$$\begin{cases} BetP_{\Omega_X^* \times \Omega_{Y^a}^*}(\omega_i, \omega_j^a) = \sqrt{BetP_{\Omega_X^*} \{Y_j^a\}(\omega_i) \cdot BetP_{\Omega_{Y^a}^*} \{X_i\}(\omega_j^a)} \\ \forall (\omega_i, \omega_j^a) \in (\Omega_X^* - \{\omega^*\}) \times (\Omega_{Y^a}^* - \{\omega^*\}) \end{cases} \quad (4.5)$$

$$\begin{cases} BetP_{\Omega_X^* \times \Omega_{Y^a}^*}(\omega^*, \omega_j^a) = BetP_{\Omega_X^*} \{Y_j^a\}(\omega^*) \\ \forall \omega_j^a \in (\Omega_{Y^a}^* - \{\omega^*\}) \end{cases} \quad (4.6)$$

$$\begin{cases} \text{Bet}P_{\Omega_X^* \times \Omega_{Y^a}^*}(\omega_i, \omega^*) = \text{Bet}P_{\Omega_{Y^a}^*} \{X_i\}(\omega^*) \\ \forall \omega_i \in (\Omega_X^* - \{\omega^*\}) \end{cases} \quad (4.7)$$

$$\text{Bet}P_{\Omega_X^* \times \Omega_{Y^a}^*}(\omega^*, \omega^*) = 0 \quad (4.8)$$

Cette variante permet de garantir les deux contraintes suivantes :

1. $\text{Bet}P_{\Omega_X^*} \{Y_j^a\}(\omega_i) \leq \text{Bet}P_{\Omega_X^* \times \Omega_{Y^a}^*}(\omega_i, \omega_j^a) \leq \text{Bet}P_{\Omega_{Y^a}^*} \{X_i\}(\omega_j^a)$
si $\text{Bet}P_{\Omega_X^*} \{Y_j^a\}(\omega_i) \leq \text{Bet}P_{\Omega_{Y^a}^*} \{X_i\}(\omega_j^a)$ et
2. $\text{Bet}P_{\Omega_X^* \times \Omega_{Y^a}^*}(\omega_i, \omega_j^a) = 0$
si $\text{Bet}P_{\Omega_X^*} \{Y_j^a\}(\omega_i) = 0$ ou $\text{Bet}P_{\Omega_{Y^a}^*} \{X_i\}(\omega_j^a) = 0$

La première condition permet de borner la probabilité résultant de la combinaison. La seconde condition, une probabilité nulle est totalement absorbante.

Après cette combinaison, nous obtenons une matrice de probabilité pignistique. Nous la normalisons grâce à l'algorithme de Sinkhorn [A70] afin d'obtenir une matrice doublement stochastique $M_{combine}^{normalise}$, c'est-à-dire vérifiant les deux égalités suivantes :

$$\begin{aligned} \sum_{\omega_j^a \in \Omega_{Y^a}^*} \text{Bet}P_{\Omega_X^* \times \Omega_{Y^a}^*}(\omega_i, \omega_j^a) &= 1 \\ \sum_{\omega_i \in \Omega_X^*} \text{Bet}P_{\Omega_X^* \times \Omega_{Y^a}^*}(\omega_i, \omega_j^a) &= 1 \end{aligned}$$

Nous appliquons ensuite sur cette matrice normalisée l'algorithme d'affectation Hongrois étendu [A8] afin d'obtenir la matrice $M_{combine}^{affecte}$ maximisant la somme des probabilités pignistiques. Cet algorithme d'affectation permet de manipuler directement la matrice combinée sans changer sa dimension contrairement à l'algorithme d'affectation utilisé par Gruyer et Royère qui nécessite une matrice carrée.

4.2.6.2 Mesure de similarité acoustique

De nombreuses mesures de similarité utilisées dans le domaine de l'identification du locuteur ont été proposées dans la littérature. Le choix d'une mesure dépend fortement du modèle acoustique utilisé pour caractériser les sources sonores. Une description précise et exhaustive de l'ensemble de ces mesures est faite dans [T4].

Dans notre travail, puisque nous utilisons une modélisation gaussienne des segments de parole à comparer, nous nous intéressons aux mesures de similarité fondées sur les

caractéristiques statistiques. Parmi les mesures les plus utilisées, nous pouvons citer la mesure de similarité AHS (Arithmetic Harmonic Sphericity) et la distance de Bhattacharyya. Nous avons choisi de travailler avec la distance de Bhattacharyya car elle présente un bon compromis entre complexité calculatoire et performances [T4].

Soient $g_1 = N(\mu_1, \Sigma_1)$ et $g_2 = N(\mu_2, \Sigma_2)$ deux gaussiennes modélisant deux sources sonores. La ressemblance entre les deux distributions selon la métrique de Bhattacharyya est définie par :

$$D_{bh} = \frac{1}{8}(\mu_1 - \mu_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\mu_1 - \mu_2) + \frac{1}{2} \ln \left(\frac{\Sigma_1 + \Sigma_2}{2\sqrt{|\Sigma_1||\Sigma_2|}} \right) \quad (4.9)$$

Une grande valeur de D_{bh} indique que les deux distributions modélisent deux sources sonores différentes. La distance de Bhattacharyya définie par l'équation 4.9 est non normalisée. Afin d'obtenir une mesure de similarité acoustique appartenant à l'intervalle $[0, 1]$, nous lui appliquons une fonction sigmoïde [C12]. Ainsi la mesure de similarité entre les deux sources gaussiennes est finalement donnée par :

$$\mathfrak{S}_A(g_1, g_2) = \begin{cases} \frac{1}{1 + \exp \left[-2 \left(\frac{D_{bh} - D_{bh}^{seuil}}{r_1} \right) \right]} & \forall D_{bh} < D_{bh}^{seuil} \\ \frac{1}{1 + \exp \left[-2 \left(\frac{D_{bh} - D_{bh}^{seuil}}{r_2} \right) \right]} & \text{sinon} \end{cases} \quad (4.10)$$

Dans cette équation, D_{bh}^{seuil} est le seuil d'acceptation utilisé pour valider l'identification du locuteur. r_1 et r_2 sont deux constantes qui délimitent la zone linéaire de la fonction sigmoïde comme le montre la figure 4.5.

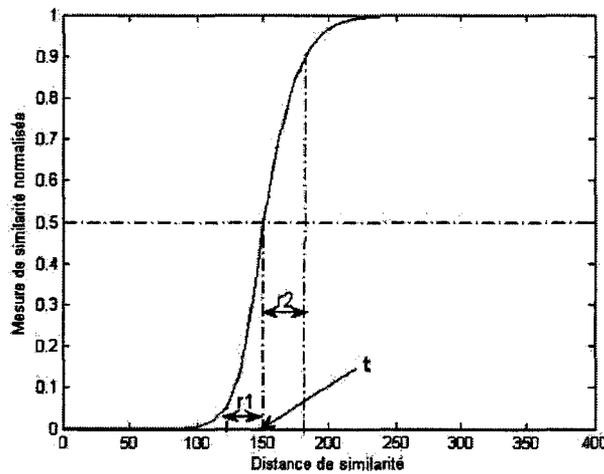


FIG. 4.5 : Fonction sigmoïde de normalisation : $D_{bh}^{seuil} = 150$, $r_1 = 20$ et $r_2 = 30$.

4.2.6.3 Mesure de similarité visuelle

Dans notre travail, chaque individu est caractérisé par son histogramme couleur. Il s'agit donc de définir une métrique capable de quantifier les différences entre deux histogrammes et ainsi d'être capable de se prononcer sur la ressemblance entre deux individus.

La définition d'une mesure de similarité ou de distance entre deux histogrammes est un problème fondamental dans les applications de classification [L8, A46, A56], de fouille de données (retrieval) [A26, A64, C60, C68, A68] et de segmentation des images [A32, C54, A63]. De nombreuses méthodes ont été proposées dans la littérature et nous pouvons les séparer en deux grandes classes : les techniques géométriques et les méthodes probabilistes.

Les approches géométriques supposent que l'histogramme est un vecteur de dimension fixe. Elles exploitent les normes standards des vecteurs telles que la distance de Manhattan et les distances euclidiennes [A26, C62, C75, A68] ou encore le principe de l'intersection d'histogrammes [A68, A69].

Les approches probabilistes considèrent l'histogramme d'une région de l'image comme une distribution de probabilité des couleurs composant cette région. Dans ce contexte, le calcul de la distance entre deux histogrammes est équivalent à une mesure de divergence entre deux distributions de probabilité. Dans la littérature, de nombreuses mesures de distance entre distributions de probabilité sont proposées. Nous trouvons par exemple la distance de Bhattacharyya [A37], la distance de Matusita [A55] et la distance de Kullback et Leibler [A40]. Pour plus de détails concernant les mesures de similarité, le lecteur pourra consulter les publications [A19, A66] dans lesquelles une étude comparative des mesures de similarité est effectuée.

Dans notre travail, nous avons choisi de calculer la mesure de similarité entre deux histogrammes couleur sur la base de leur intersection telle que la définit Smith dans [A69]. Cette mesure de similarité présente l'avantage d'être symétrique et normalisée. Soit $H^1 = \{h_{c_1}^1, \dots, h_{c_k}^1\}$ et $H^2 = \{h_{c_1}^2, \dots, h_{c_k}^2\}$ deux histogrammes couleur définis sur l'espace de couleur quantifié en k tons $\{c_1, \dots, c_k\}$. La mesure de similarité entre ces deux histogrammes est définie de la manière suivante :

$$\mathfrak{S}_V(H^1, H^2) = 1 - \frac{\sum_{i=1}^k \min(h_{c_i}^1, h_{c_i}^2)}{\min(\sum_{i=1}^k (h_{c_i}^1), \sum_{i=1}^k (h_{c_i}^2))} \quad (4.11)$$

Si les deux histogrammes sont identiques, $\mathfrak{S}_V(H^1, H^2) = 1$. Une valeur faible de cet indice

de similarité signifie que les deux histogrammes modélisent deux objets différents.

4.2.6.4 Mesure de similarité de position

Nous avons choisi d'utiliser l'indice proposé par Gruyer dans [T3]. L'indice de similarité entre deux positions s'exprime sur la base d'une modélisation floue de chaque position. Cette modélisation permet d'exprimer l'incertitude d'une mesure en fonction de son imprécision. Nous avons choisi de représenter la fonction d'appartenance de chaque mesure par une forme triangulaire.

Soit P_1 et P_2 deux positions. La mesure de similarité entre deux positions est une fonction de l'intersection de la représentation géométrique de leur fonction d'appartenance. Cette mesure est donnée par la formule suivante :

$$\mathfrak{S}_P(P_1, P_2) = \frac{S^{P_1} \cap S^{P_2}}{\min(S^{P_1}, S^{P_2})} \quad (4.12)$$

où S^{P_1} et S^{P_2} correspondent respectivement aux surfaces des triangles représentant les fonctions d'appartenance des positions P_1 et P_2 .

Sur la figure 4.6, nous schématisons cette mesure de similarité. Dans ce schéma, σ_1 est l'imprécision liée à P_1 et σ_2 est l'imprécision liée à P_2 . Notons que dans notre travail l'imprécision sur la position d'une cible prédite est l'erreur de prédiction représentée par la covariance de l'état obtenue par le filtre de Kalman.

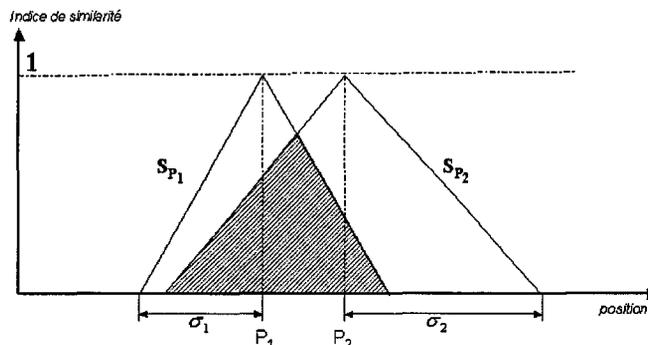


FIG. 4.6 : Mesure de similarité de position entre P_1 et P_2 .

4.2.6.5 Calcul des fonctions de croyance élémentaires

Pour chaque instanciation du module d'association observations/cibles prédites, les fonctions de croyance élémentaires sont générées à partir des mesures de similarité précédentes.

ment décrites. Les évaluations ont été réalisées avec le générateur de Rombaut.

Plaçons nous, comme nous l'avons fait la section 4.2.6.1, dans le contexte du module d'association observations audio/cibles prédites. Chaque fonction de croyance élémentaire $m_i^{\Omega_X} \{Y_l^a\}$ est définie sur chaque proposition ω_i de Ω_X . Comme nous l'avons indiqué au chapitre 3, la proposition ω_i est la classe associée à l'objet connu O_i . Dans ce contexte, $\mathfrak{S}_P(P_l^a, P_l^X)$ et $\mathfrak{S}_A((\mu_l^a, \Sigma_l^a), (\mu_l^X, \Sigma_l^X))$ représentent respectivement la mesure de similarité de position et la mesure de similarité acoustique entre le couple de vecteurs d'observation Y_l^a et la cible connue X_i .

La fonction $m_i^{\Omega_X} \{Y_l^a\}$ est donnée par les équations suivantes :

$$\begin{cases} m_i^{\Omega_X} \{Y_l^a\} (\omega_i) &= \alpha e^{-\mathfrak{S}_{PA}(Y_l^a, X_i)} \\ m_i^{\Omega_X} \{Y_l^a\} (\bar{\omega}_i) &= \alpha (1 - e^{\mathfrak{S}_{PA}(Y_l^a, X_i)}) \\ m_i^{\Omega_X} \{Y_l^a\} (\Omega_X) &= 1 - \alpha \end{cases} \quad (4.13)$$

où α quantifie la fiabilité de la source d'information et $\mathfrak{S}_{PA}(Y_l^a, X_i)$ est la combinaison suivante :

$$\mathfrak{S}_{PA}(Y_l^a, X_i) = \sqrt{\mathfrak{S}_P(P_l^a, P_l^X) \times \mathfrak{S}_A((\mu_l^a, \Sigma_l^a), (\mu_l^X, \Sigma_l^X))}$$

Les fonctions de croyance élémentaires du module d'association observations vidéo/cibles prédites sont définies en adaptant cette méthode à l'indice de similarité vidéo.

4.2.7 Conclusion

Dans cette section nous avons présenté le contexte général de notre application qui consiste à suivre des individus par un système de perception audio et vidéo. Il s'agit plus spécifiquement de suivre des visages parlants.

Après avoir positionné notre application par rapport à l'état de l'art, nous avons défini chacun des capteurs logiques utilisés lors de la phase de validation. Le capteur logique vidéo assure les fonctions de détection, de localisation et de caractérisation des visages présents dans les images. Le visage de chaque individu est caractérisé par son histogramme couleur. Le capteur logique audio assure les fonctions de détection, de localisation et caractérisation des sources sonores. La voix de chaque individu est caractérisée par un modèle gaussien des paramètres LPC extraits du signal.

La méthode d'association observations/cibles prédites est fondée sur les méthodes antérieures présentées par Rombaut, Gruyer et Royère. Nous présentons les modifications que nous avons apportées à cette méthode et décrivons les mesures de similarité utilisées afin

de calculer les fonctions de croyance élémentaires.

La suite de ce chapitre est consacrée à la présentation des résultats obtenus lors de la phase de validation.

4.3 Résultats

Dans la dernière section de ce chapitre, nous présentons les résultats obtenus lors de la phase de validation de la méthode d'association et de la technique de gestion de conflit que nous avons proposées. Nous supposons que les individus sont en mouvement dans une salle de forme carrée équipée d'un ensemble de caméras et de microphones. Les deux capteurs logiques sont conformes à ce que nous avons décrit au début de ce chapitre. Toutefois, face à la complexité qu'implique le développement de chaque capteur, l'obtention des données modales (vecteurs de position et vecteurs d'attributs) n'est pas réalisée de manière automatique.

Les données cinématiques

Les trajectoires de tous les individus sont totalement simulées. Elles sont bruitées avec un bruit gaussien de moyenne nulle et d'écart-type de $50cm$.

Afin de déterminer les prédictions des positions des cibles, nous avons utilisé un filtre de Kalman linéaire du premier ordre. Le filtre de Kalman fournit à chaque instant la précision de l'estimation de la position de chaque individu. Ce paramètre intervient lors du calcul de la mesure de similarité de position entre les observations et les cibles prédites.

Les données audio

Comme nous l'avons précisé précédemment, la voix de chaque individu est modélisée par une gaussienne. Le vecteur moyenne et la matrice de covariance de chaque gaussienne ont été calculés sur la base d'un corpus d'apprentissage spécifique à chaque individu susceptible d'être perçu par les capteurs.

La fréquence d'échantillonnage du signal de parole est égale à $44100Hz$. Chaque signal de parole est découpé en segments d'analyse d'une durée de $5secondes$, avec un recouvrement de 50%. Chaque segment est à son tour décomposé en trames de 1024 échantillons (soit $23ms$) avec 512 échantillons de recouvrement (soit $11.6ms$ équivalent à un recouvrement de 50%). Afin de diminuer la distorsion spectrale du signal de parole, une fenêtre de Hamming est appliquée sur chaque trame du signal.

Les 12 LPC et le terme d'énergie sont extraits de chaque trame, créant ainsi un vecteur acoustique à 13 composantes. Ce vecteur caractéristique est extrait uniquement pour les



trames non silencieuses. L'ensemble des vecteurs acoustiques obtenu pour chaque segment est exploité afin de calculer le vecteur moyenne et la matrice de covariance du modèle gaussien recherché.

Les données vidéo

Un histogramme modèle a été déterminé pour chaque individu. Pour cela, nous avons effectué une acquisition de plusieurs séquences vidéo pour chacun d'entre eux, séquences que nous avons segmentées manuellement afin d'extraire leur visage (cf. 4.7).

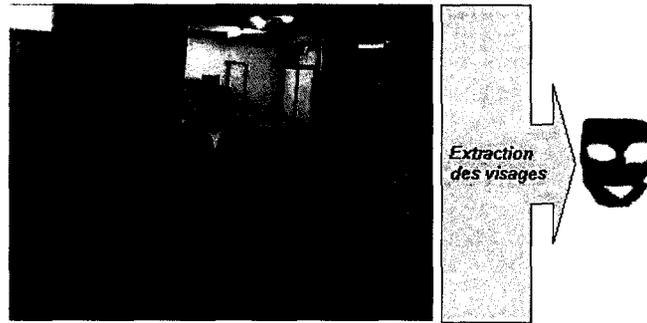


FIG. 4.7 : Extraction manuelle de visage.

La mesure de performance d'association

Afin d'évaluer de manière objective les résultats fournis par notre méthode, nous avons défini une mesure de performance qui quantifie la confiance sur chaque association réalisée. Cette mesure prend en compte les décisions sur les associations et celles relatives aux non-associations. Cette mesure est définie par la relation suivante :

$$C = \frac{\sum_{i=1}^{m+1} \sum_{j=1}^{n+1} M_{combine}^{affecte}(i, j) \cdot M_{combine}^{normalise}(i, j)}{N} \quad (4.14)$$

La valeur N est un facteur de normalisation introduit afin d'obtenir une mesure de performance appartenant à l'intervalle $[0, 1]$. Il est donné par la formule suivante :

$$N = N_c + N_s \quad (4.15)$$

N_c est le nombre de 1 dans la matrice d'affectation $M_{combine}^{affecte}$ en excluant la dernière ligne et colonne, c'est-à-dire le nombre d'associations réalisées. N_s est le nombre de 1 dans la dernière ligne et de la dernière colonne de la matrice d'affectation, c'est-à-dire le nombre de non-associations réalisées. Il indique le nombre de nouveaux objets observés ou le nombre

d'objets connus ayant disparu.

Dans [T3], Gruyer définit une mesure de performance similaire de la manière suivante :

$$C = \frac{\sum_{i=1}^m \sum_{j=1}^n M_{combine}(Y_i, X_j) \cdot x_{i,j}}{\min(n, p)} \quad (4.16)$$

où $x_{i,j}$ représente l'affectation ou non de l'objet connu X_j avec l'observation Y_i :

$$x_{ij} = \begin{cases} 1 & \text{si } M_{combine}(Y_i, X_j) \succ \max(m_{\cap_{partiel}^{\Omega_X} \{Y_i\}}(\omega^*), m_{\cap_{partiel}^{\Omega_Y} \{X_j\}}(\omega^*))_{i \leq m, j \leq n} \\ 0 & \text{sinon} \end{cases} \quad (4.17)$$

Toutefois, à l'instar de notre méthode, seule la décision sur l'association des objets connus aux observations entre en jeu. La masse de croyance portée par les hypothèses de non-associations, quantifiée par $m_{\cap_{partiel}^{\Omega_X} \{Y_i\}}(\omega^*)$ et $m_{\cap_{partiel}^{\Omega_Y} \{X_j\}}(\omega^*)$ constituent pourtant deux parties importantes de l'information fournie par le processus de fusion qui peuvent entrer en jeu dans la phase de décision finale.

Le processus de validation

Afin de valider notre méthode d'association, nous mettons en place plusieurs scénarii montrant chacun une situation critique qui peut se produire dans une application de suivi multi-objet dans un contexte multi-modal.

Dans un premier temps, nous mettons en évidence l'intérêt de notre méthode d'association observations/cibles dans le cas de suivi d'individus très proches. Puis nous évaluons la méthode dans le cas où un individu est détecté par un seul des deux capteurs. Ensuite, nous considérons la situation où de nouveaux individus apparaissent et sont détectés soit grâce à l'une des deux modalités, soit simultanément par les deux capteurs logiques. Enfin, nous présentons les résultats obtenus lors d'un scénario mettant en scène deux individus qui apparaissent et disparaissent au même moment et à la même position.

Avant d'aller plus loin, précisons que dans la suite de ce chapitre, les observations audio, les observations vidéo et les cibles sont numérotées de manière bijective avec les individus. Ainsi, lorsque nous parlons de l'observation audio (ou vidéo) 1 ou de la cible 1, nous parlons des observations et de la cible associées à l'individu 1.

4.3.1 Cas du suivi d'individus proches

Le scénario est présenté à la figure 4.8(a). Deux individus se déplacent le long de deux trajectoires très proches l'une de l'autre.

La figure 4.8(b) représente les trajectoires telles qu'elles sont estimées séparément par les capteurs logiques audio et vidéo (observations bruitées). La figure 4.8(c) représente les trajectoires estimées par notre méthode. Les figures 4.8(d) et 4.8(e) représentent respectivement la mesure de performance de l'association observations audio/cibles et celle d'association observations vidéo/cibles.

Les figures 4.8(f), 4.8(g), 4.8(h) et 4.8(i) représentent respectivement les probabilités pignistiques des hypothèses d'association des observations audio et des observations vidéo.

Comme le montrent les figures 4.8(c), 4.8(d) et 4.8(e), bien que les individus se déplacent en étant très proches l'un de l'autre et malgré la présence d'un bruit important, l'association observation-cible est bien réalisée. Les mesures de performance d'association ont des valeurs importantes qui sont comprises entre 0.7 et 0.8.

Remarquons cependant dans la figure 4.8(e) qu'à l'instant $t = 22s$ la mesure de performance d'association observations vidéo-cibles est un peu faible par rapport aux autres valeurs. Ceci est dû au fait que la probabilité pignistique de l'hypothèse d'association de la première observation vidéo avec la première cible à laquelle elle est associée est faible (cf. 4.8(h)). A cet instant, la probabilité pignistique de l'hypothèse de la non association de cette observation, $(BetP(\emptyset^*))$, est non négligeable, indiquant le fait qu'il s'agit peut être d'une nouvelle observation. Ceci s'explique car la mesure de similarité visuelle entre la première observation vidéo et la première cible est faible.

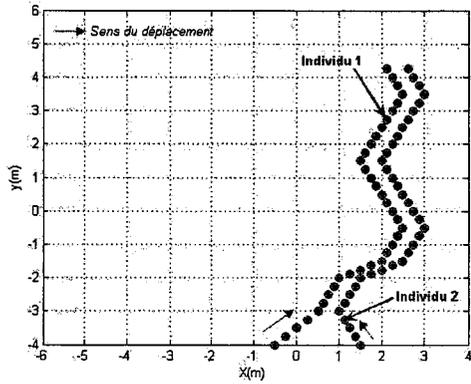
Sur les figures 4.8(g), 4.8(h) et 4.8(i), nous remarquons que les probabilités pignistiques des hypothèses d'association des observations avec les cibles auxquelles elles sont effectivement associées est importante. L'association observations/cibles est effectuée sans aucune ambiguïté bien que les deux individus soient très proches.

4.3.2 Cas du suivi d'individus perçus grâce à l'une ou l'autre des deux modalités

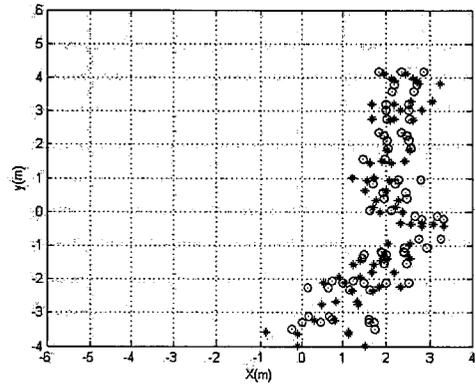
Le scénario est présenté à la figure 4.9(a). Les individus se déplacent le long de leur trajectoire mais ne sont perçus que temporairement par l'un ou l'autre des deux capteurs. Le capteur logique vidéo subit des occultations partielles : une occultation visuelle du premier individu à l'instant $t = 25s$ et un silence momentané du deuxième individu entre l'instant $t = 5s$ et $t = 13s$.

La figure 4.9(b), présente les positions estimées par les capteurs logiques audio et vidéo. La figure 4.9(c) représente les trajectoires estimées en utilisant la méthode proposée.

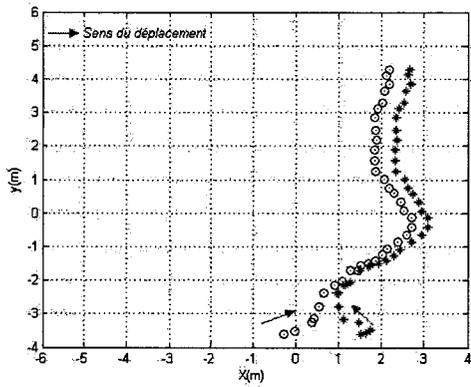
Les figures 4.9(d) et 4.9(e) représentent respectivement la mesure de performance de



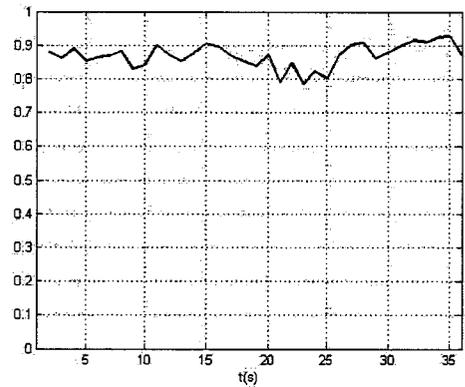
(a) Trajectoires réelles



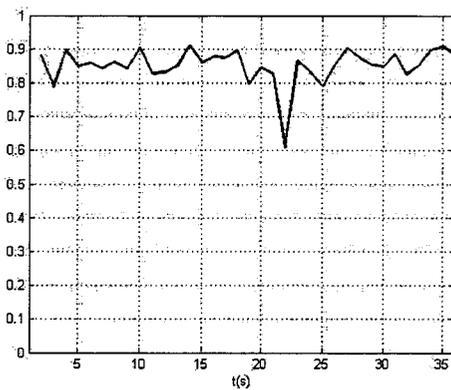
(b) Positions estimées par les deux capteurs logiques audio "*" et vidéo "o"



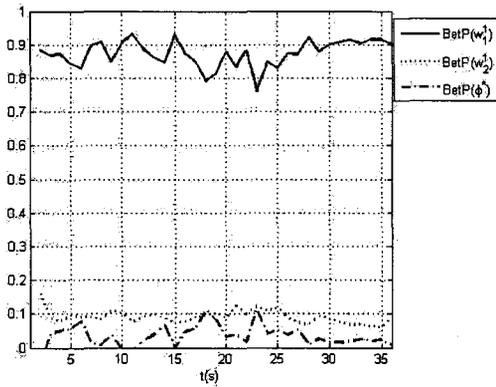
(c) Trajectoires estimées par notre méthode



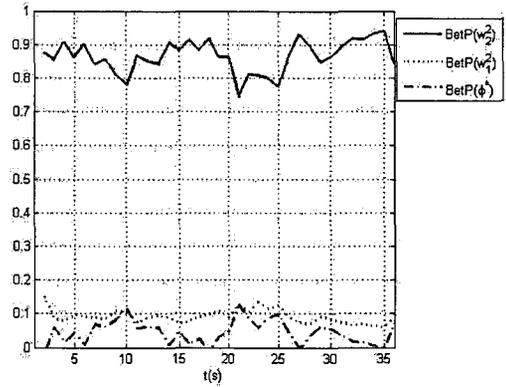
(d) Mesure de performance d'association audio-cible



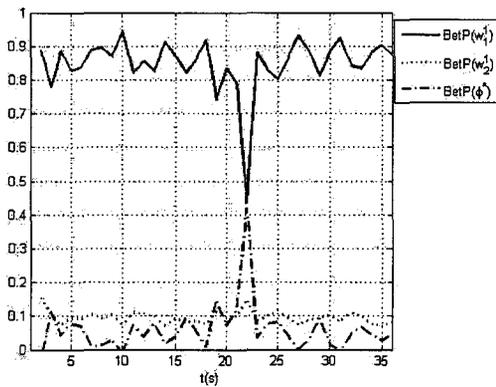
(e) Mesure de performance d'association vidéo-cible



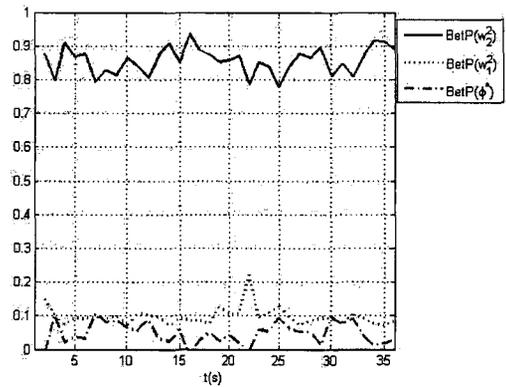
(f) Probabilités pignistiques des hypothèses d'association de la première observation audio



(g) Probabilités pignistiques des hypothèses d'association de la deuxième observation audio



(h) Probabilités pignistiques des hypothèses d'association de la première observation vidéo



(i) probabilités pignistiques des hypothèses d'association de la deuxième observation vidéo

FIG. 4.8 : Suivi de deux individus proches.

l'association observations audio/cibles et celle de l'association observations vidéo/cibles.

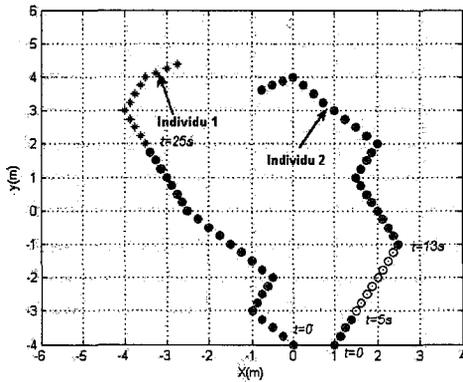
Les figures 4.9(f), 4.9(g), 4.9(h) et 4.9(i) font état respectivement les probabilités pignistiques de l'association entre les observations et les deux cibles.

D'après les figures 4.9(c), 4.9(d) et 4.9(e), l'absence d'une modalité est compensée par la présence de l'autre modalité. Ainsi à l'instant $t = 25s$ l'individu occulté par un élément de l'environnement est perçu par le capteur logique audio. Ceci permet au module d'association entre les observations audio et les cibles prédites de poursuivre le processus de suivi de l'individu.

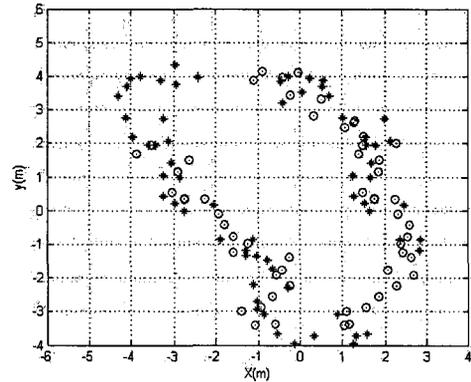
Si nous analysons le comportement du conflit relatif à l'association de la cible 2 présenté sur les figures 4.9(g), nous constatons les points suivants :

- comme le présente la figure 4.9(g), le début du processus de suivi est caractérisé par une probabilité pignistique sur l'hypothèse d'association de la deuxième cible avec la deuxième observation audio maximale.
- à l'instant $t = 5$, instant auquel le second individu cesse de parler, la probabilité pignistique, $BetP(\emptyset^*)$, de l'hypothèse de la non association de la cible 2 avec aucune des observations croit brutalement. Ce constat reste valable jusqu'à l'instant $t = 13s$.
- pendant la période $t = 5s$ à $t = 13s$, la cible 2 est associée à la deuxième observation vidéo (cf. figure 4.9(i)), ce qui permet de prolonger son suivi en l'absence de mesure. Cette cible est donc maintenue.
- à partir de l'instant $t = 13s$, c'est-à-dire dès que l'individu réémet un signal de parole, la probabilité pignistique est maximale sur l'hypothèse qui associe à la cible 2 la seconde observation audio (cf. figure 4.9(g)).

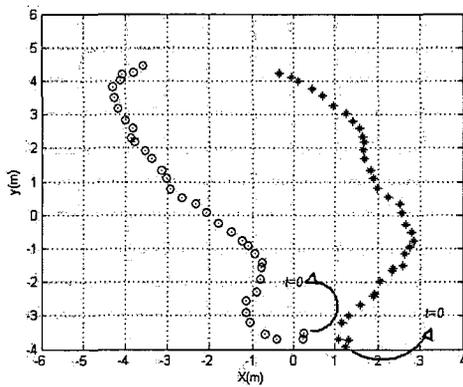
Une analyse identique est réalisée à partir de la figure 4.9(h). Les constatations faites pour la cible 2 sont semblables à celle de la cible 1. A partir de l'instant $t = 25s$, l'individu 1 est invisible pour le capteur logique vidéo, ce qui se traduit par une valeur élevée de probabilité pignistique $BetP(\emptyset^*)$ sur l'hypothèse de la non association de la cible 1. Cette valeur reste élevée puisque l'individu est occulté jusque la fin du scénario. Pendant cette même période, l'individu parle. Il est donc perçu par le capteur logique audio. Comme le montre la figure 4.9(f), la cible 2 est associée à l'observation audio 2 et son suivi peut être prolongé.



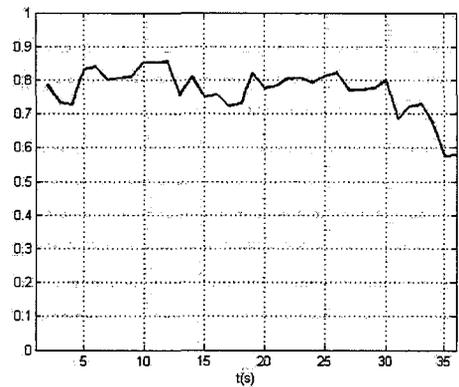
(a) Trajectoires réelles



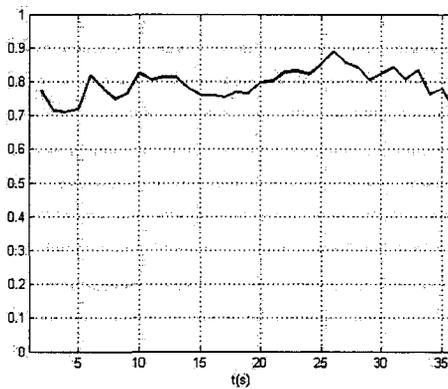
(b) Positions estimées par les deux capteurs logiques audio "*" et vidéo "o"



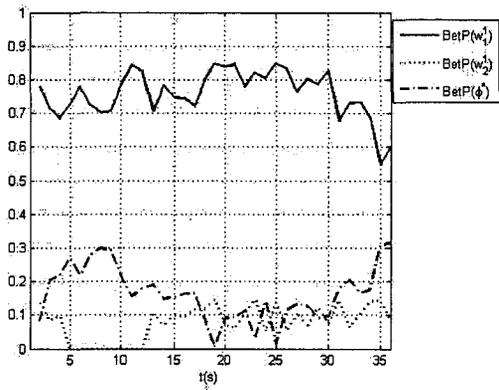
(c) Trajectoires estimées par notre méthode



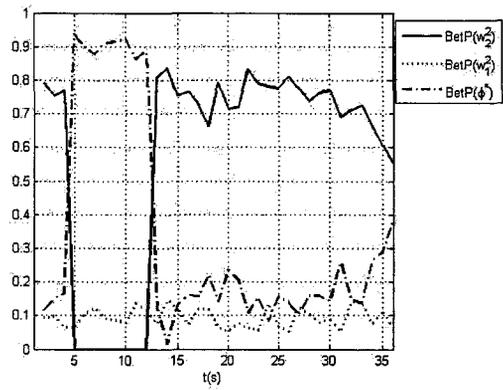
(d) Mesure de performance d'association audio-cible



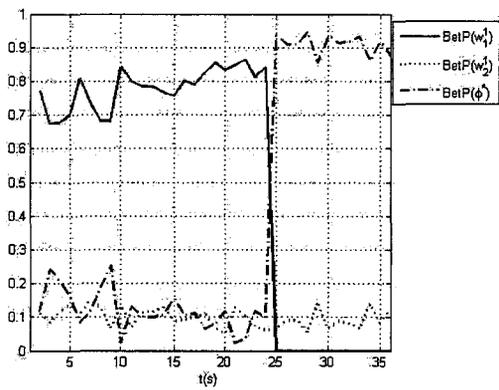
(e) Mesure de performance d'association vidéo-cible



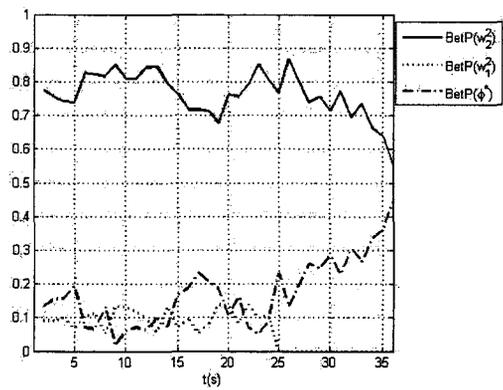
(f) Probabilités pignistiques des hypothèses d'association de la première cible avec les observations audio



(g) Probabilités pignistiques des hypothèses d'association de la deuxième cible avec les observations audio



(h) Probabilités pignistiques des hypothèses d'association de la première cible avec les observations vidéo



(i) Probabilités pignistiques des hypothèses d'association de la deuxième cible avec les observations vidéo

FIG. 4.9 : Suivi d'individus perçus grâce à l'une ou l'autre des deux modalités.

4.3.3 Cas de l'apparition des individus

Le scénario est présenté sur la figure 4.10(a). Nous étudions plus particulièrement la situation où chaque individu est détecté, dès son apparition dans la scène, par l'un des deux capteurs logiques.

Le scénario se décompose de la manière suivante : au début du scénario deux individus se déplacent dans la scène et sont susceptibles d'être perçus par les deux capteurs logiques. A l'instant $t = 11s$, un troisième individu entre dans la scène à partir de la position $(4, -3)$. A son entrée dans la scène, cet individu ne parle pas jusqu'à l'instant $t = 25s$, instant où il se met à parler. A l'instant $t = 17s$, un quatrième individu parlant entre dans la pièce à la position $(3.5, 6)$. Il est visible pour le capteur logique vidéo. Enfin, à l'instant $t = 21s$, un cinquième individu entre en parlant à la position $(-5.5, 6)$.

La figure 4.10(b), présente les positions des individus estimées par les capteurs logiques audio et vidéo. La figure 4.10(c) illustre les trajectoires estimées en utilisant la méthode proposée.

Les figures 4.9(d) et 4.9(e) représentent respectivement la mesure de performance de l'association observations audio/cibles et celle de l'association observations vidéo/cible.

Les figures 4.10(f), 4.10(g), 4.10(h), 4.10(i) et 4.10(j) présentent respectivement l'évolution des probabilités pignistiques sur les hypothèses d'association des observations audio. Les figures 4.10(k), 4.10(l), 4.10(m) et 4.10(n) représentent les probabilités pignistiques des hypothèses d'association des observations vidéo.

Apparition de l'individu 1 et 2

D'après les figures 4.10(c), 4.10(d) et 4.10(e), nous constatons que tous les nouveaux individus sont détectés, qu'ils soient perçus par un seul capteur logique ou simultanément par les deux capteurs. Les nouvelles cibles sont ainsi créées et leur suivi est efficace durant tout le temps où les individus correspondants sont perçus.

Si un nouvel objet est détecté par le système d'association, il est difficile de décider s'il s'agit d'un bonne ou une mauvaise détection. La solution adoptée afin de pallier cet inconvénient est de considérer la détection comme bonne, de créer la cible correspondante et d'augmenter la confiance associée si des observations viennent appuyer son existence pendant une période suffisamment longue (période de validation). Si ce n'est pas le cas, la détection est considérée comme fausse et la cible est détruite. Le réglage de la période de validation est délicate. Elle est directement liée au taux de mauvaise et de non détection.

Nous remarquons sur les figures 4.10(f) et 4.10(g) que les observations audio 1 et 2 ont

une probabilité pignistique importante sur leur hypothèse d'association avec les cibles 1 et 2. La vérité nous indique que ces deux cibles se réfèrent aux deux individus qui ont effectivement produit les observations audio 1 et 2. Ceci est maintenu jusqu'à la fin du processus de suivi. La figure 4.10(f) montre que la première observation audio est associée à la première cible. Tandis que la figure 4.10(g) indique que la deuxième observation est associée à la deuxième cible.

La même analyse est faite pour les deux premières observations vidéo qui correspondent aux deux premiers individus. La figure 4.10(k) indique que la première observation vidéo est associée à la première cible, elle-même associée à la première observation audio. Ceci est en accord avec la description du scénario. De même, la figure 4.10(l) montre que la deuxième cible, associée à la deuxième observation audio, est elle-même associée à la deuxième observation vidéo. Ceci se révèle être en accord avec la description du scénario.

Apparition de l'individu 3

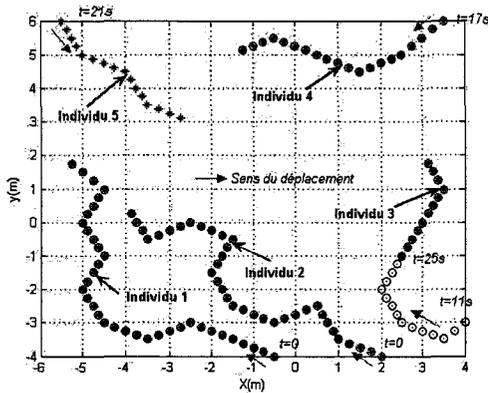
L'analyse de la figure 4.10(m) montre que l'association des observations vidéo et des cibles conduit à une valeur élevée de la probabilité pignistique sur l'hypothèse de non association. Ceci permet de conclure à l'apparition d'un troisième individu à l'instant $t = 11s$. Une nouvelle cible est initialisée à cet instant. A l'instant $t = 12$ suivant la création de la cible 3, une nouvelle observation vidéo vient appuyer son existence : la probabilité pignistique sur l'hypothèse de son association avec la cible 3 est élevée et la probabilité pignistique de l'hypothèse de sa non association est diminuée.

A l'instant $t = 25$, cet individu est détecté par le capteur logique audio : la probabilité pignistique sur l'hypothèse d'association de l'observation audio avec cette cible augmente (cf. figure 4.10(h)).

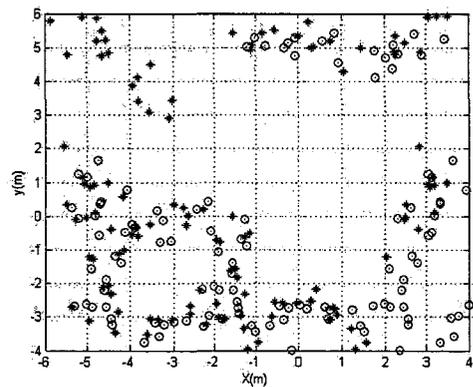
Apparition de l'individu 4

Analysons maintenant les figures 4.10(i) et 4.10(n) relatives aux graphes de probabilités pignistiques des hypothèses d'association des observations audio et vidéo qui correspondent au quatrième individu.

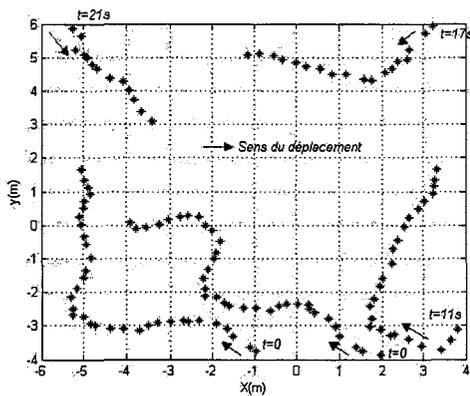
Une analyse des graphes de la figure 4.10(i) et 4.10(n) permet de conclure à l'apparition à l'instant $t = 17$ d'un quatrième individu : une grande valeur de probabilité pignistique est affectée aux hypothèses de la non association des observations audio et vidéo qui correspondent à cet individu. La cible 4 est alors initialisée. Les observations audio et vidéo qui suivent sont associées de manière correcte à la cible 4 : la probabilité d'association affectée sur l'hypothèse de leur non association est faible et la probabilité pignistique sur



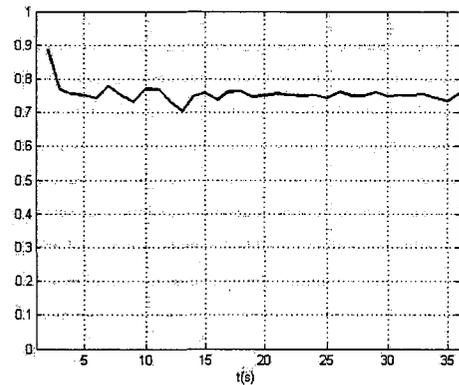
(a) Trajectoires réelles



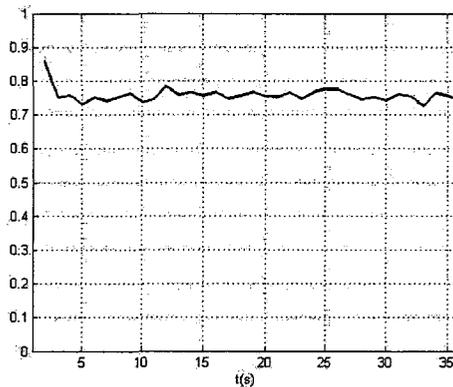
(b) Positions estimées par les deux capteurs logiques audio "*" et vidéo "o"



(c) Trajectoires estimées par notre méthode



(d) Mesure de performance d'association audio-cible

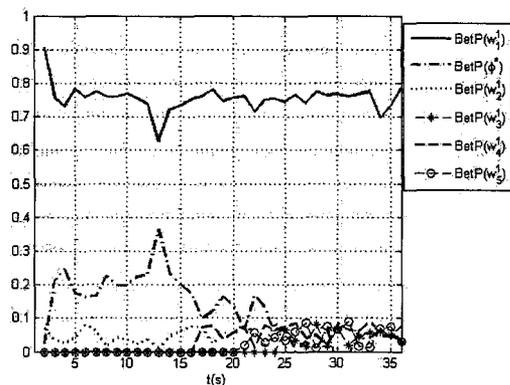


(e) Mesure de performance d'association vidéo-cible

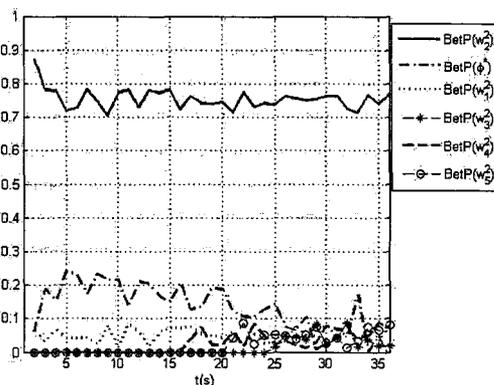
l'hypothèse de leur association avec la cible 4 est élevée.

Apparition de l'individu 5

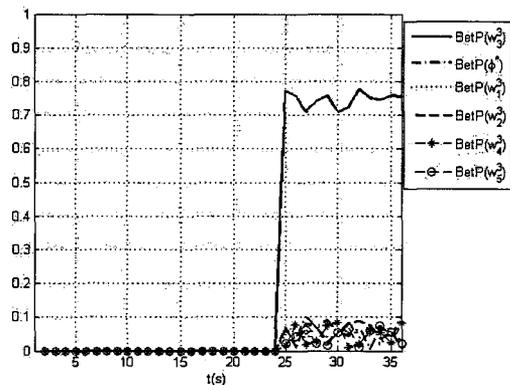
L'analyse réalisée dans le cas de l'individu 3 reste valable. Le cinquième individu est détecté à l'instant $t = 21$ grâce à la modalité audio (cf. figure 4.10(j)).



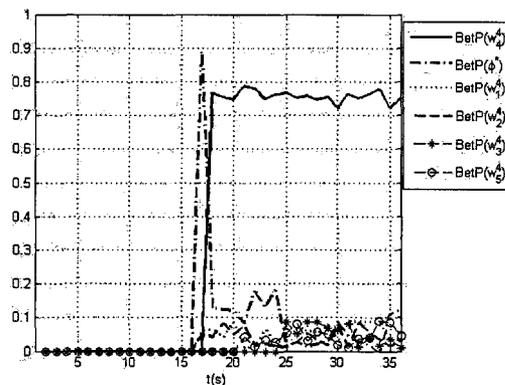
(f) probabilités pignistiques des hypothèses d'association de la première observation audio



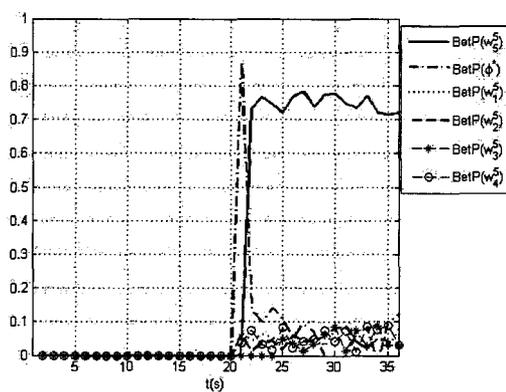
(g) probabilités pignistiques des hypothèses d'association de la deuxième observation audio



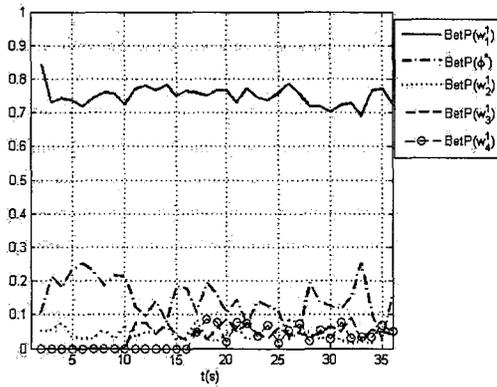
(h) probabilités pignistiques des hypothèses d'association de la troisième observation audio



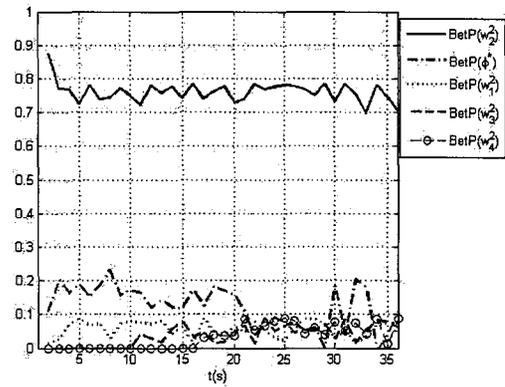
(i) probabilités pignistiques des hypothèses d'association de la quatrième observation audio



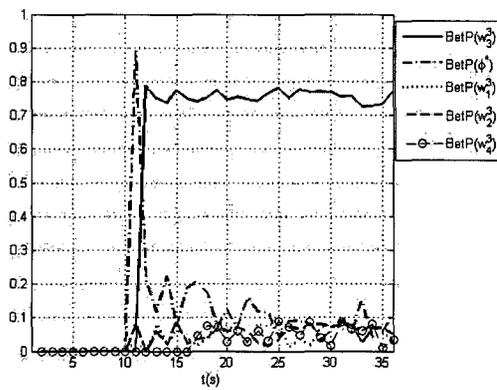
(j) probabilités pignistiques des hypothèses d'association de la cinquième observation audio



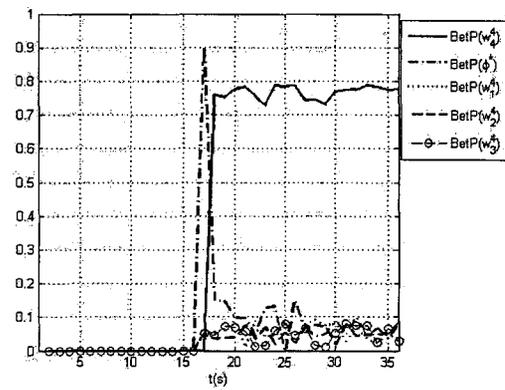
(k) probabilités pignistiques des hypothèses d'association de la première observation vidéo



(l) probabilités pignistiques des hypothèses d'association de la deuxième observation vidéo



(m) probabilités pignistiques des hypothèses d'association de la troisième observation vidéo



(n) probabilités pignistiques des hypothèses d'association de la quatrième observation vidéo

FIG. 4.10 : Apparitions des individus.

4.3.4 Cas de l'apparition/disparition, simultanée et à des positions voisines

La figure 4.11(a), illustre ce scénario. Le premier individu entre dans la scène depuis la position $(-4.5, -4)$ et la quitte à l'instant $t = 12s$ de la position $(-6, -1.75)$. A ce même instant, un deuxième individu entre dans la scène à la position $(-6, -1.5)$. A l'instant $t = 27s$, le premier individu disparaît puis réapparaît à la position où il avait quitté la scène.

Dans cette figure et pour des raisons de clarté, nous avons schématisé les trajectoires des deux individus en utilisant uniquement un cercle ou une étoile, bien que les deux individus soient détectés tout au long du scénario simultanément par les deux capteurs logiques audio et vidéo.

Les figures 4.11(b) et 4.11(c) représentent respectivement les trajectoires des individus estimées séparément par les deux capteurs logiques et leurs trajectoires estimées grâce à notre méthode.

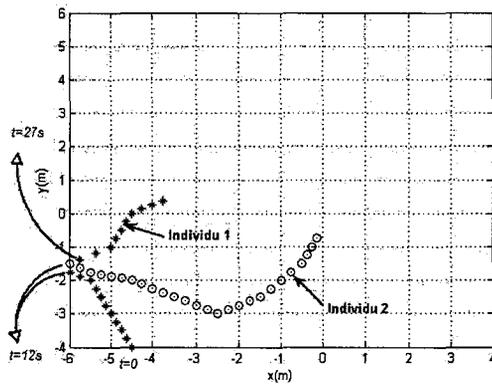
Les figures 4.11(d) et 4.11(e) donnent les mesures de performance de l'association des observations audio avec les cibles et celles de l'association des observations vidéo avec les cibles.

Les figures 4.11(f), 4.11(g), 4.11(h) et 4.11(i) nous indiquent qu'à l'instant $t = 13$, un individu entre dans la scène pendant qu'un second en sort à la même position. Une nouvelle cible est initialisée. Les deux nouvelles observations audio et vidéo n'ont pas été associées à la cible existante bien que l'information de position portée par les observations modales soient proches de l'information de position de la cible existante. Les caractéristiques modales contenues dans les vecteurs d'observations sont suffisamment discriminantes pour éviter ce problème.

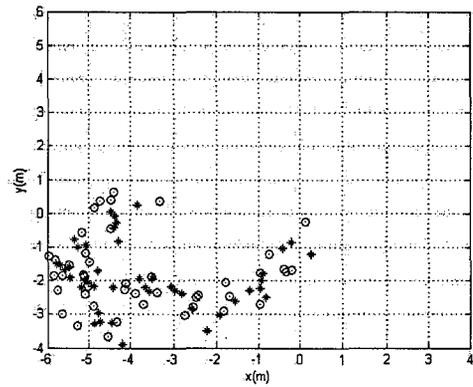
La cible existante n'est pas immédiatement éliminée. Elle est propagée par le processus d'association durant un certain temps. Toutefois, l'incertitude et l'imprécision sur sa position sont augmentées. Ainsi la probabilité pignistique sur l'hypothèse de sa non association reste importante jusqu'à l'instant $t = 27s$ où la cible est associée à de nouvelles observations audio et vidéo (cf. figure 4.11(f) et 4.11(g)).

4.4 Conclusion

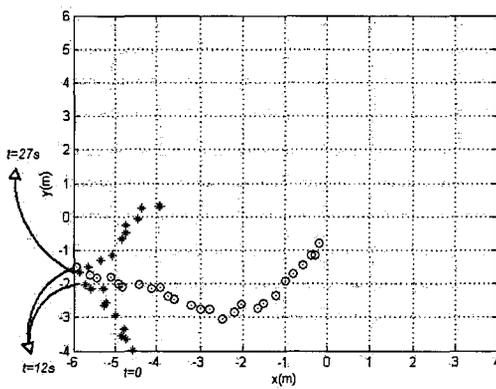
Dans ce dernier chapitre nous avons présenté les résultats obtenus avec notre méthode dans le contexte d'une application de suivi audio/vidéo de plusieurs individus. Dans un premier temps, nous avons décrit les capteurs logiques audio et vidéo composant le système de perception sur lequel nous avons envisagé le processus de suivi/association. Une brève



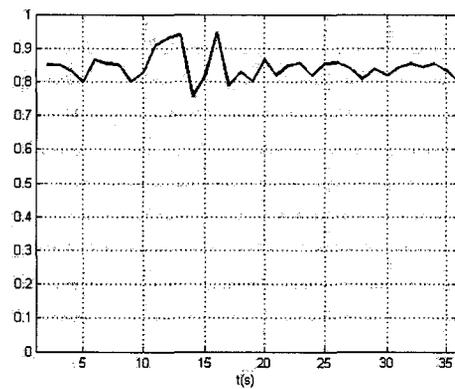
(a) Trajectoires réelles



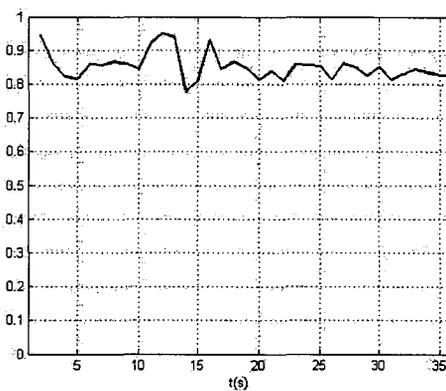
(b) Positions estimées par les deux capteurs logiques audio "*" et vidéo "o"



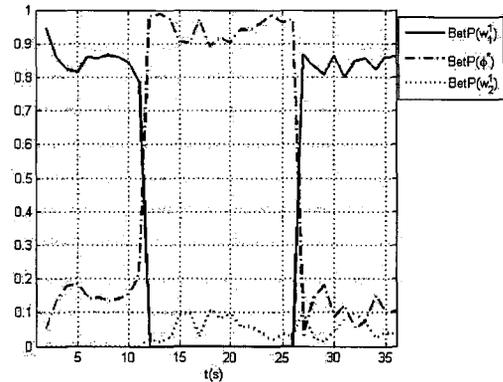
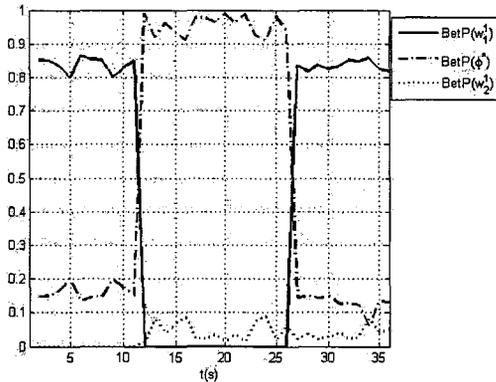
(c) Trajectoires estimées par notre méthode



(d) Mesure de performance d'association audio-cible

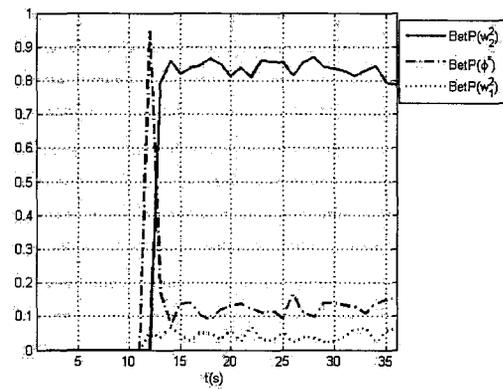
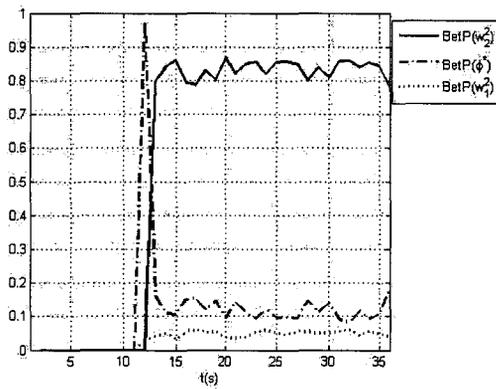


(e) Mesure de performance d'association vidéo-cible



(f) probabilités pignistiques des hypothèses d'association de la première cible avec les observations audio

(g) probabilités pignistiques des hypothèses d'association de la première cible avec les observations vidéo



(h) probabilités pignistiques des hypothèses d'association de la deuxième observation audio

(i) probabilités pignistiques des hypothèses d'association de la deuxième observation vidéo

FIG. 4.11 : Apparition et disparition simultanées de deux individus.

étude des fonctionnalités de détection, de segmentation et de caractérisation intégrées dans chaque capteur logique a montré la grande difficulté de développer ou d'automatiser certaines d'entre elles.

C'est dans ce contexte que nous avons réalisé la phase de validation de la méthode proposée sur des signaux réels (images et sons) mais en intégrant certaines limites. Ainsi, les vecteurs d'observations ont été obtenus à partir de procédures manuelles - segmentation vidéo et séparation de sources acoustiques - de procédures automatiques - obtention des histogrammes et des modèles acoustiques - et de procédures de simulation - obtention des informations de position.

Nous avons mis en évidence l'intérêt de notre méthode dans la gestion de situations complexes rencontrées lors d'une application de suivi multi-objet dans un contexte multi-capteur et plus particulièrement les situations où les objets sont perçus de manière discontinue au cours du temps et les situations où des objets disparaissent et de nouveaux objets apparaissent.

La méthode proposée permet de déceler de telles situations en analysant de manière spécifique le conflit généré lors de la phase de combinaison des fonctions de croyance élémentaires. Cette technique opère une redistribution du conflit sur un cadre de discernement plus spécifique intégrant une hypothèse supplémentaire. Finalement, la distribution des probabilités obtenue lors du passage au niveau pignistique est plus discriminante et augmente la fiabilité de la décision.

Conclusion et perspectives

4.5 Conclusion

Dans ce mémoire, nous avons présenté une méthode d'association crédibiliste de données dans le contexte du suivi multi-objet par un système de perception multi-capteur. Contrairement aux études classiques qui proposent de raisonner sous l'hypothèse du monde ouvert étendu, nous avons choisi de travailler, comme le propose Smets, sous l'hypothèse du monde ouvert et dans le contexte de modèle des Croyances Transférables.

Ce mémoire s'articule autour de quatre chapitres. Le premier chapitre est un état de l'art des méthodes de suivi/association multi-objet. Dans le cadre bayésien, la méthode traditionnelle du filtrage de Kalman convient si la loi conditionnelle commandant l'évolution de l'état d'un objet à suivre ne s'éloigne pas trop d'une loi normale. Dans un contexte multi-objet et lorsque la loi conditionnelle n'est plus gaussienne, des méthodes telles que celles fondées sur les techniques de filtrage particulière constitue une solution intéressante. Toutes ces méthodes nécessitent de résoudre le problème de l'association des mesures aux objets connus. Nous avons choisi d'étudier cette étape d'association dans le contexte de la théorie des fonctions de croyance afin de prendre en compte un contexte multi-objet et de gérer l'apparition et la disparition d'objets.

Cette théorie est capable de prendre en compte l'imperfection liées aux données (données imprécises, données incertaines et données manquantes) sans avoir une connaissance a priori précise du modèle associé à chacune d'entre elles. Dans le chapitre 2, après une brève présentation des bases nécessaires à la bonne compréhension de cette théorie, nous avons présenté la démarche la plus couramment décrite dans la littérature. Cette démarche fait l'hypothèse du monde ouvert étendu dans lequel le cadre de discernement est considéré comme fermé. L'étude de la masse de croyance affectée à l'élément de fermeture permet alors de déceler si une mesure appartient à un objet connu ou si elle correspond à un objet inconnu.

Plusieurs travaux ont montré l'efficacité de telles méthodes. Toutefois, comme le pré-

cise Smets, des précautions sont à prendre quant à la cardinalité de l'élément de fermeture ajouté au cadre de discernement. Cette cardinalité intervient directement lors de la prise de décision. Dans le chapitre 3, nous avons présenté une méthode d'association formulée dans le cadre du monde ouvert considéré comme non exhaustif. La détection des phénomènes d'apparition et de disparition des objets est réalisée par une analyse spécifique de la masse conflictuelle résultante de la phase de combinaison des fonctions de croyance. Ce conflit est décomposé en deux parties. Une partie du conflit est directement reliée aux phénomènes d'apparition et de disparition tandis que la seconde partie relève une origine indéterminée. La première partie du conflit est éliminée grâce à une étape de raffinement. Le tout étant réalisé au niveau crédal, la valeur totale de la masse conflictuelle est réduite. Ce qui permet de procéder à une étape de normalisation lors du passage au niveau pignistique.

Le quatrième chapitre fait l'état des résultats obtenus lors de la phase d'évaluation de notre méthode. Nous nous sommes placés dans le cadre d'une application de suivi d'individus par un système de perception bimodal audio et vidéo. Dans un premier temps, nous avons décrit chaque capteur logique composant le système de perception. Grâce à ce système, chaque individu est caractérisé par des attributs acoustiques et par une signature vidéo qui permet de le reconnaître et de le discriminer des autres individus tout au long de son trajet. Dans notre travail, nous avons proposé de modéliser la voix de chaque individu par un vecteur de coefficients LPC dont la distribution est supposée suivre une loi normale. Ce même individu est caractérisé par son visage que nous modélisons par un histogramme couleur. La validation de la méthode proposée est réalisée sur des données à la fois réelles et synthétisées. Les résultats de ces simulations sont très encourageants et témoignent de l'efficacité de notre nouvelle méthode d'association.

4.6 Perspectives

Comme nous l'avons précisé dans le chapitre 4, face à la difficulté de développer les deux capteurs logiques audio et vidéo, l'évaluation a été réalisée sur la base d'attributs modaux dont une partie a été simulée et une autre partie extraite de manière semi-automatique. Dans ce contexte, nous avons considéré chaque objet (et plus particulièrement le visage et la voix de chaque individu) comme étant caractérisé par un modèle invariable dans le temps : ainsi, pour chaque individu, son visage est modélisé par un histogramme couleur normalisé constant et sa voix est représentée par une loi gaussienne dans l'espace des

paramètres retenus (12LPC). Aujourd'hui, ce choix ne nous permet pas de suivre des visages qui ne sont pas perçus de face par les caméras. En effet, pour un même individu, l'histogramme normalisé couleur d'un visage tourné de quelques degrés par rapport à la caméra est sensiblement différent de celui qui est appris alors que ce même visage est perçu de face. Nous retrouvons cette propriété dynamique dans le cadre du modèle de la voix d'un individu dont les paramètres acoustiques (LPC) varient au cours de son discours. Par conséquent, outre la possibilité de choisir des attributs et des modèles plus adaptés afin de modéliser au mieux chaque individu, nous envisageons de prendre en compte leur propriété dynamique, c'est-à-dire leur faculté à varier dans le temps. Par exemple, il semblerait intéressant d'intégrer un modèle d'apparence 3D des visages. De la même manière prendre en compte les dérivés premiers et secondes des paramètres acoustiques paraît une piste intéressante. Par ailleurs, il pourrait être intéressant d'analyser de quelle manière chaque modèle peut être adapté au fur et à mesure que de nouvelles observations sont disponibles et que les associations entre les observations et les objets connus sont réalisées. Le corpus d'évidence sur lequel les fonctions de croyance élémentaires sont générées s'adapterait donc au cours du temps en fonction d'un ensemble de paramètres dynamiques (les conditions de luminosité, les caractéristiques géométriques et optiques des caméras etc.).

Récemment, Smets a proposé une formulation de son modèle de croyance transférable dans le domaine continu. Cette approche permet d'adresser un nombre plus important d'applications qui requièrent de sauvegarder la nature continue des informations modélisées. Dans le cadre du modèle des croyances transférables discret, une variable continue peut être approchée grâce à une discrétisation plus fine de son domaine de définition. Toutefois, cela reste possible au détriment de la cardinalité du cadre de discernement et du référentiel de définition. Par exemple, estimer la position angulaire d'un objet mobile entre $-\pi$ et π avec une précision de $\frac{\pi}{360}$ nécessite 360 hypothèses dans le cadre de discernement et 2^{360} propositions dans le référentiel de définition. Or, pour des raisons de puissance de calcul et de capacité mémoire des calculateurs, respecter la contrainte temps-réel demande de définir un cadre de discernement dont la cardinalité n'excède pas 20 hypothèses. Le modèle de croyance transférable continu est un outil prometteur pour prendre en compte intelligemment la nature continue des paramètres. A ce jour, cette approche a été appliquée dans le domaine du positionnement de véhicule [C15] et celui de la classification [C81]. Nous pourrions appliquer cette approche continue à la modélisation de chaque objet que nous souhaitons suivre et d'adapter notre méthode de gestion du conflit.

Comme nous l'avons spécifié très tôt dans le mémoire, la procédure d'association a lieu sous l'hypothèse que chaque objet de la scène ne donne lieu qu'à une seule observation. Cette contrainte est particulièrement forte et facilite à la fois les processus d'association que nous avons appelé APC et ACP. Or, dans le contexte du traitement des images, il est assez courant que lors d'une étape de segmentation, une région émane de deux objets différents ou encore que plusieurs régions de l'image proviennent d'un même objet. Dans ce contexte, l'étape de décision sur l'association telle que nous l'avons proposée sur la base du maximum de probabilité pignistique ne convient plus et doit être adaptée. Une solution serait de retenir plusieurs solutions possibles et d'éliminer les moins probables au fur et à mesure des observations. Une stratégie de type MHT pourrait être adoptée telle que le propose Murillion et al. dans [C55].

Il est important de noter que nous avons évalué la méthode de gestion du conflit dans le contexte d'une application de suivi de visage parlant. Bien entendu, cette méthode s'adapte parfaitement à d'autres contextes applicatifs tels que la détection et du suivi des véhicules et des piétons dans l'environnement d'un véhicule [C51, C55], l'aide à la navigation sur cartographie numérique [T6] et la localisation de véhicule par reconnaissance d'amers visuels [C40]. Dans chacune de ces applications, l'apparition d'une masse conflictuelle témoigne d'un phénomène particulier. Dans le cadre d'une application de détection d'obstacle, le conflit peut être le reflet de l'apparition ou la disparition d'un véhicule dans le champ de vision de la caméra. Il s'agit de l'exemple le plus proche de celui étudié au cours de cette thèse. Dans le contexte de la localisation par amer visuel et de la navigation sur cartographie numérique, l'apparition d'un conflit peut signifier que le véhicule modifie la trajectoire qu'on lui avait adressée. Par ailleurs, ces différents contextes applicatifs font souvent état de problèmes non linéaires et l'estimation des modèles nécessitent l'implantation de méthodes adaptées telles que celles présentées dans le chapitre 1. Depuis plusieurs années, les recherches ont montré que les méthodes de filtrage particulières sont des outils particulièrement adaptés à l'estimation des distributions de probabilités de phénomènes non linéaires. L'association de ces méthodes et de modèle des croyances transférables pourraient s'avérer une solution pertinente à la prise en compte de ces phénomènes non linéaires dont l'observation est souvent imprécise, incertaine et peut se révéler incomplète.

Annexes

Annexe A

Opérateur de gestion de conflit

A.1 Introduction

Dans le cadre de la théorie des fonctions de croyance, la fusion des informations fournies par plusieurs sources est réalisée grâce à la combinaison des fonctions de croyance relatives à chaque source. Lors de cette phase de combinaison, une masse conflictuelle peut apparaître. Ce conflit doit être éliminé afin de préserver les propriétés des jeux de masse initiaux.

Dans le cadre d'une application de fusion de données utilisant la théorie des fonctions de croyance, les origines de conflit peuvent être classées selon qu'elles proviennent d'un problème dû aux sources d'information ou d'une mauvaise modélisation de l'environnement :

- Conflit dû aux sources d'information :
 - . Une mesure aberrante issue d'un capteur qui peut être due à la défaillance du capteur lors de son acquisition ou à l'utilisation du capteur hors de son domaine de fonctionnement.
 - . Les sources ne font pas référence au même objet.
 - . La source d'information non fiable qui peut s'interpréter comme le dysfonctionnement d'un capteur ou d'un expert.
 - . Les sources ne sont pas indépendantes.
- Conflit dû à la modélisation de l'environnement :
 - . Modélisation imprécise des fonctions de croyance : un mauvais choix de la fonction de distance ou de vraisemblance.
 - . La non exhaustivité du cadre de discernement : la non prise en compte d'une ou de plusieurs hypothèses durant la construction du cadre de discernement.
 - . La contrainte d'exclusivité des hypothèses du cadre de discernement n'est plus

vérifiée.

Les causes de conflit peuvent survenir simultanément, ce qui rend l'opération de détection de son origine et son interprétation difficile. Ceci constitue un des challenges de la théorie des fonctions de croyance. Le conflit doit être traité de manière adaptée selon ses origines.

Le conflit engendré lors de la combinaison de plusieurs sources d'informations par l'opérateur de Dempster est masqué par l'opération de normalisation. La règle de combinaison de Dempster est une opération purement conjonctive. Elle renforce la croyance sur les propositions dont les sources sont en accord et annule les propositions en conflit. Plus le conflit est important, plus la décision est prise que sur une faible partie des informations disponibles. En cas de conflit important l'étape de normalisation est impossible.

Plusieurs opérateurs de combinaison ont été proposés dans la littérature afin de résoudre le problème du conflit. Ces opérateurs sont généralement classés selon l'origine du conflit en deux classes. On trouve les règles qui supposent que l'origine du conflit est due à des sources non fiables et celles qui supposent que l'origine du conflit est due à un cadre de discernement non exhaustif. Dans cette thèse nous nous attacherons à classer les règles de combinaison selon la proposition de Smets [A74]. Dans ses travaux, les règles de combinaison sont classées en fonction de l'algèbre des sous-ensembles du cadre de discernement dans laquelle les masses de croyances sont définies. Il définit donc deux grandes classes : les règles définies dans un treillis distributif (distributive lattice) et les règles définies dans l'algèbre booléenne (boolean algebras). Cette dernière classe à son tour est classée en deux sous-classes : les règles de combinaison normalisées et les règles de combinaison non normalisées. Rappelons ici que l'algèbre booléenne dans le cadre de la théorie des fonctions de croyance est l'ensemble des sous-ensembles du cadre de discernement.

Dans les sections suivantes nous adopterons les notations suivantes. Soit Ω un cadre de discernement. Soit m_i une fonction de croyance i , où i est un entier représentant son indice. La règle de combinaison conjonctive est notée $\cap : m_1 \cap m_2$ est la combinaison conjonctive des fonctions de croyance m_1 et m_2 . m_\cap représente la fonction de croyance résultante de la combinaison de toutes les fonctions de croyance m_i .

A.2 Les règles de combinaison définies dans l'algèbre booléenne

A.2.1 Les règles de combinaison normalisées

A.2.1.1 Règle de combinaison de Yager

Yager suppose que le conflit est généré par un défaut de fiabilité d'une partie des sources d'informations [A83, A84, A85]. Il est certain que la solution du problème est dans le cadre de discernement. Yager propose donc de réaffecter toute la masse conflictuelle à la proposition représentant la méconnaissance Ω . Cet opérateur transforme l'information conflictuelle en information de méconnaissance. Ainsi le résultat m_{\cap}^y de la combinaison, par la règle de Yager, des fonctions de croyance m_i est donné par :

$$\begin{aligned} m_{\cap}^y(A) &= m_{\cap}(A) \quad \forall A \subseteq \Omega \\ m_{\cap}^y(\Omega) &= m_{\cap}(\Omega) + m_{\cap}(\emptyset) \\ m_{\cap}^y(\emptyset) &= 0 \end{aligned} \tag{A.1}$$

Cette règle de combinaison est commutative mais elle n'est pas associative. Il est donc nécessaire de définir un ordre de fusion des sources.

A.2.1.2 Règle de combinaison de Shafer

La règle de combinaison proposée par Shafer [L17] permet de transférer une partie de la croyance vers l'ensemble représentant la méconnaissance totale Ω en appliquant un coefficient d'affaiblissement sur les sources d'information non totalement fiables. Si nous notons, α cet indice de fiabilité avec $\alpha \in [0, 1]$, la distribution de masse m_i affaiblie, notée m_i^A , est définie de la manière suivante :

$$\begin{aligned} m_i^A(A) &= \alpha \cdot m_i(A) \quad \forall A \subseteq \Omega \\ m_i^A(\Omega) &= \alpha \cdot m_i(\Omega) + (1 - \alpha) \end{aligned} \tag{A.2}$$

Si $\alpha = 0$, la source est totalement défailante. Si $\alpha = 1$, la source est totalement fiable. Comme la règle précédente, cette règle de combinaison est commutative mais elle n'est pas associative. Un ordre de fusion des sources est donc nécessaire.

A.2.1.3 Règle de combinaison disjonctif

La règle de combinaison disjonctive est inspirée des outils développés par Dubois et Prade [A23] dans la théorie des possibilités. Elle affecte le résultat de la combinaison des sources sur des unions d'hypothèses de la manière suivante :

$$m_{1 \cup 2}(A) = \sum_{B \cap C = \emptyset, B \cup C = A} m_1(B).m_2(C) \quad \forall A \subseteq \Omega \quad (\text{A.3})$$

Cet opérateur est utilisé lorsqu'on sait que l'une des sources n'est pas fiable mais on ne sait pas laquelle. Avec cet opérateur, la masse sur l'ensemble vide est toujours nulle, par contre, la qualité de la distribution de masse obtenue est moins précise. L'opérateur est commutatif mais il n'est pas associatif.

A.2.1.4 Règle de combinaison de Dubois et Prade

Dubois et Prade [A24] proposent de redistribuer la masse conflictuelle sur l'union des propositions à l'origine du conflit. En fait, l'information de conflit est transformée en une information de doute entre les deux propositions sources de ce conflit. L'opérateur est donc défini par la relation suivante :

$$m_{1 \cap 2}^D(A) = m_{1 \cap 2}(A) + \sum_{B \cap C = \emptyset, B \cup C = A} m_1(B).m_2(C) \quad (\text{A.4})$$

Cette règle de combinaison utilise une approche conjonctive quand les sources sont en accord et une approche disjonctive en cas de conflit. Mais comme le précédent opérateur, cet opérateur est commutatif mais pas associatif.

A.2.1.5 Règle de combinaison de Lefevre

Dans [A45, A44], Lefevre propose un cadre générique pour unifier les opérateurs décrits précédemment. Par ailleurs, ce cadre permet de définir d'autres règles de redistribution de la masse conflictuelle en fonction de l'objectif précis.

L'objectif de cet opérateur générique est de distribuer la masse conflictuelle sur un ensemble de proposition proportionnellement à une valeur de pondération. Ce poids varie selon la proposition considérée et des masses ayant engendré le conflit. Ainsi, après cette

transformation, la fonction de croyance résultante m_{\cap}^I est définie par :

$$\begin{cases} m_{\cap}^I(A) = m_{\cup}(A) + m_{\cap}^c(A) & \forall A \subseteq \Omega, A \neq \emptyset \\ m_{\cap}^I(\emptyset) = 0 \end{cases} \quad (\text{A.5})$$

où $m_{\cap}^c(A)$ s'écrit :

$$m_{\cap}^c(A) = w(A, m_1, \dots, m_i, \dots) \cdot m_{\cap}(\emptyset) \quad \forall A \subseteq \Omega \quad (\text{A.6})$$

Les coefficients de pondération vérifient la relation :

$$\sum_{A \subseteq \Omega} w(A, m_1, \dots, m_i, \dots) = 1 \quad (\text{A.7})$$

Les opérateurs précédents peuvent être réécrits en définissant de manière spécifique les valeurs des poids de pondération.

Dans [C41], les auteurs proposent une règle similaire dans laquelle la masse placée sur l'ensemble vide peut ne pas être nulle.

A.2.1.6 La règle de combinaison unifiée d'Inagaki

Inagaki [A34] a défini une classe paramétrique de règles de combinaison qui inclue les opérateurs de Dempster et de Yager. Soit m_{\cup}^I le résultat de cette règle appliquée à la fonction de croyance m_{\cap} . m_{\cup}^I est définie de la manière suivante :

$$m_{\cup}^I(A) = (1 + K \cdot K^*) m_{\cap}(A) m_{\cup}^I(\Omega) = (1 + K^* \cdot K) m_{\cap}(\Omega) + (1 + K^* \cdot K - K^*) \cdot m_{\cap}(\Omega) \quad (\text{A.8})$$

avec $K = m_{\cap}(\emptyset)$ et $0 \leq k^* \leq \frac{1}{1-K-m_{\cap}(\Omega)}$. Pour $k^* = \frac{1}{1-K}$, la règle de combinaison d'Inagaki est équivalente à celle de Dempster. Pour $k^* = 0$, la règle coïncide avec celle de Yager. La détermination de la valeur de k^* est une étape importante pour l'implémentation de cet opérateur. Son choix reste difficile à justifier. Par ailleurs, la seule règle de combinaison associative de cette classe paramétrique d'opérateurs est la règle de Dempster.

A.2.1.7 Les règles de combinaison utilisant une mesure d'intersection

Cette règle est proposée pour la première fois par Zhang dans [L22]. Il voulait fournir une alternative à la règle de Dempster qui prend en compte la mesure d'intersection des éléments focaux du cadre de discernement.

Dans ces règles de combinaison, les masses à combiner sont donc pondérées par une

mesure d'intersection entre les sous-ensembles du cadre de discernement. Pour deux fonctions de masses de croyance m_1 et m_2 , ces règles peuvent s'écrire sous la forme suivante :

$$m(X) = k^* \sum_{A \cap B = X} r(A, B) m_1(A) m_2(B) \quad (\text{A.9})$$

Dans cette équation $r(A, B)$ est la mesure d'intersection entre deux ensemble finis A et B et k^* est un facteur de normalisation qui assure que la somme de toutes les masses de croyance est égale à 1. Avec ces règles, la masse sur l'ensemble vide est ignorée : $m(\emptyset) = 0$.

D'après Smets [A74], ces règles sont associatives si la mesure d'intersection utilisée vérifie l'équation suivante : $r(A, B \cap C) = r(A \cap B, C)$.

Les règles de combinaison de Zhang et de Mahler entrent dans cette catégorie :

Règle de combinaison du centre du Zhang (Center combination rule)

Zhang [L22] définit la mesure d'intersection $r(A, B)$ entre deux ensemble finis A et B comme étant le rapport entre la cardinalité de leur intersection et le produit de leur cardinalité :

$$r(A, B) = \frac{|A \cap B|}{|A| \cdot |B|} \quad (\text{A.10})$$

En utilisant une autre mesure d'intersection entre les éléments focaux du cadre de discernement, nous pouvons générer d'autres règles de combinaison. Par exemple $r(A, B)$ peut être calculée en prenant le rapport entre la cardinalité de leur intersection et la cardinalité de leur union.

Règle de combinaison de Mahler

Dans [A52] Mahler fournit une autre version de la règle de Zhang en utilisant une mesure d'intersection calculée sur la base des fonctions de crédibilité des masses à combiner. La fonction de croyance résultante est définie par :

$$m^M(X) = m_{1 \cap 2}(X) + k^* \sum_{A \cap B = X} \left[\frac{bel(X)}{bel(A) \cdot bel(B)} m_1(A) m_2(B) \right] \quad (\text{A.11})$$

A.2.2 Les règles de combinaison non normalisées

A.2.2.1 Règle de combinaison conjonctif et modèle des croyances transférables

Dans la littérature cet opérateur est connu sous le non règle de combinaison conjonctive. Récemment dans [A74], l'auteur a ajouté à cette nomenclature le terme TBM pour éviter toute confusion avec les autres règles. Ici, TBM est l'abréviation du terme anglophone de modèle des croyances transférables (Transferable Belief model).

Smets propose de ne pas redistribuer le conflit tel que le font les autres opérateurs. Cette masse conflictuelle est propagée et maintenue d'une combinaison à l'autre. Si la masse conflictuelle dépasse une valeur seuil, un système expert est sollicité afin de trouver son ou ses origines et pour réduire sa valeur. Cet opérateur est défini de la manière suivante :

$$\begin{cases} m_{1\cap 2}(A) &= \sum_{B\cap C=A} m_1(B).m_2(C) \quad \forall A \subseteq \Omega \\ m_{1\cap}(\emptyset) &= \sum_{B\cap C=\emptyset} m_1(B).m_2(C) \end{cases} \quad (\text{A.12})$$

La règle de combinaison conjonctive TBM est commutative et associative. Le système expert vérifie toutes les hypothèses sous lesquelles la règle de combinaison conjonctive TMB est appliquée, notamment la fiabilité des sources, leur indépendance, l'égalité des objets auxquels les sources font référence et la non exhaustivité du cadre de discernement. Cependant, le choix du seuil à partir duquel le système expert est déclenché constitue le problème majeur pour cette règle de combinaison.

A.2.3 Les règles de combinaison définies dans un treillis distributif

La théorie de Dezert-SMarandache (DSmT) décrite dans [L16] a été proposée en 2001. C'est une alternative à la théorie des croyances qui permet de manipuler des données incertaines et paradoxales (conflictuelles). Tout comme la théorie des fonctions de croyance, elle a la particularité de définir un cadre de discernement qui doit inclure les solutions à un problème posé. Toutefois, les hypothèses qui le composent, ne sont pas nécessairement exhaustives et exclusives.

Selon ses fondateurs, la DSmT a été développée pour résoudre les trois limitations suivantes de la théorie des fonctions de croyance :

- Une masse de croyance conflictuelle trop importante modifie la structure de l'information portée par toutes les distributions de croyance. L'étape de décision devient difficile et une étape de gestion du conflit doit être mise en place.
- Chaque hypothèse du cadre de discernement doit être définie sans ambiguïté. Les hypothèses modélisés sous la forme de notions vagues et floues ne sont pas manipulées par cette théorie.
- La théorie suppose une nature universelle au cadre de discernement : les différentes sources d'informations doivent en avoir la même interprétation.

Dans la DSmT, le cadre de discernement Ω est appelé *cadre de discernement généralisé* puisque les hypothèses élémentaires ne sont pas nécessairement disjointes et exhaustives.

Le référentiel de la définition 2^Ω de la théorie des fonctions de croyance est remplacé par le *hyper-powerset* noté D^Ω obtenu en prenant non seulement toutes les unions possibles des hypothèses du cadre de discernement mais aussi toutes les intersections possibles entre elles (non nulles car les hypothèses sont non exclusives). Par exemple, si on considère le cadre de discernement Ω constitué seulement de deux hypothèses : $\Omega = \{\omega_1, \omega_2\}$ nous avons : $2^\Omega = \{\emptyset, \omega_1, \omega_2, \omega_1 \cup \omega_2\}$ et $D^\Omega = \{\emptyset, \omega_1, \omega_2, \omega_1 \cup \omega_2, \omega_1 \cap \omega_2\}$. Soit m_1 , une fonction de croyance. Selon qu'elle est définie dans le cadre d'une ou l'autre des théories, m_1 vérifie :

- Pour la théorie des fonctions de croyance : $m(\omega_1) + m(\omega_2) + m(\omega_1 \cup \omega_2) = 1$.
- Pour la DSMT : $m(\omega_1) + m(\omega_2) + m(\omega_1 \cup \omega_2) + m(\omega_1 \cap \omega_2) = 1$.

La DSMT peut être interprétée comme une généralisation de la théorie des fonctions de croyance. En effet, dans la DSMT, l'exclusivité des hypothèses du cadre de discernement apparaît comme un cas particulier.

La génération de l'hyper-powerset D^Ω correspond au problème de Dedekind d'énumération des fonctions monotones booléennes croissantes (fonctions logiques exprimables avec les opérateurs logiques *ET* et *OU*). C'est pour cette raison que l'hyper-powerset est désigné par la lettre *D* pour faire référence au mathématicien R. Dedekind.

Il faut noter que la cardinalité de D^Ω est toujours majorée par 2^{2^n} où n est la cardinalité du cadre de discernement Ω . Dans le cas où $n \geq 10$, la solution générale du problème de Dedekind n'est pas encore définie. Ce qui constitue aujourd'hui une problématique à résoudre en mathématique.

La masse de croyance, la crédibilité, la plausibilité ainsi que la probabilité pignistique sont définies de la même manière que dans la théorie des croyances à l'exception que les propositions sont maintenant définies dans D^Ω .

Dans la suite de cette section, nous décrivons les deux opérateurs de combinaison définis dans cette théorie. Ces opérateurs sont définis dans le treillis de Dedekind :

- La règle de combinaison DSMT libre (DSMT rule of combination for free-DSMT models)
- La règle de combinaison DSMT hybride (hybrid DSMT rule).

A.2.3.1 Le modèle libre DSMT

Le modèle libre DSMT noté $M^f(\Omega)$, correspond au cas où toutes les hypothèses du cadre de discernement Ω sont vagues et floues, c'est-à-dire non exclusives. Dans ce modèle l'intersection entre les hypothèses est différente de l'ensemble vide. La règle de combinaison est définie par la relation suivante :

$$m^{free}(A) = \sum_{B \cap C = A} m_1(B) \cdot m_2(C) \quad \forall A \in D^\Omega \quad (\text{A.13})$$

Par définition du modèle, la combinaison n'engendre aucun conflit : $m^{free}(\emptyset) = 0$. Cette règle est équivalente à la règle conjonctive. Elle est donc commutative et associative.

A.2.3.2 Le modèle hybride DSm (Hybrid DSm model)

Des contraintes peuvent être introduites sur quelques sous-ensembles du D^Ω . Si nous connaissons par exemple l'impossibilité de quelques propositions $A \in D^\Omega$ nous pouvons les représenter par l'ensemble vide. Le modèle ainsi obtenu est appelé le modèle hybride DSm et il est noté $M^h(\Omega)$. Si ce modèle est totalement contraint (toutes les intersections entre les hypothèses sont vides), il coïncide avec le modèle de Shafer dans lequel toutes les hypothèses du cadre de discernement sont exclusives. Ce dernier modèle est noté $M^0(\Omega)$. L'ensemble des éléments du D^Ω qui ont été forcés à l'ensemble vide sont notés par \emptyset_M .

Soit $\Phi = \{\emptyset, \emptyset_M\}$ l'ensemble de tous les éléments de D^Ω forcés à l'ensemble vide par les contraintes du modèle hybride $M^h(\Omega)$ y compris l'ensemble vide classique \emptyset .

Dans ce modèle, les auteurs définissent la règle de combinaison hybride qui peut être appliquée dans le cadre d'un modèle partiellement contraint à un modèle totalement contraint (le modèle de Shafer). Dans le cas de combinaison de deux évidences m_1 et m_2 , cette règle est donnée par la formule suivante :

$$m(X) = \begin{cases} S_1(X) + S_2(X) + S_3(X) & \forall X \subseteq \Omega, X \notin \Phi \\ 0 & X \in \Phi \end{cases} \quad (\text{A.14})$$

Le premier terme S_1 correspond à la règle conjonctive du modèle libre défini par :

$$S_1(X) = \sum_{A, B \in D^\Omega, A \cap B = X} m_1(A) m_2(B) \quad (\text{A.15})$$

Le deuxième terme S_2 représente un transfert de la masse portée par tous les ensembles vides (les ensembles vides forcés et l'ensemble vide classique) l'ignorance partielle ou totale. Le lecteur pourra se référer à [L16] pour plus de détails. Le dernier terme S_3 a pour rôle de transférer les masses affectées aux ensembles vides forcés à des ensembles non

vides. Ce terme est défini par la relation suivante :

$$S_3(X) = \sum_{A, B \in D^\Omega, A \cap B \in \Phi, A \cup B = X} m_1(A)m_2(B) \quad (\text{A.16})$$

Il faut noter que, lorsqu'elle est appliquée à un modèle hybride totalement contraint, la règle de combinaison hybride DSm n'est pas équivalente à la règle de Dempster bien que la cadre de discernement D^Ω soit égal à 2^Ω .

A.3 Quelques remarques

Dans [A30] et [C33], Haenni critique toutes les règles de combinaison apparues dans la littérature. Selon lui, le conflit généré lors d'une combinaison des sources d'informations par la règle de combinaison de Dempster est dû à une mauvaise modélisation du problème à résoudre et non à la règle elle-même. Donc l'introduction et l'utilisation d'un autre opérateur restent injustifiées puisque le problème de conflit peut être résolu en mettant en place un modèle plus représentatif du phénomène observé.

Smets dans [A74] critique les opérateurs de gestion de conflit qui tentent de résoudre le cas d'une masse conflictuelle totale. Selon lui le vrai défi de la combinaison des évidences est lorsque le conflit apparaît au fur et à mesure des combinaisons. Lorsque les sources sont entièrement conflictuelles, aucun opérateur ou système expert ne peut résoudre le conflit. Tout ce que nous pouvons conclure c'est qu'il y a un conflit et nous ne connaissons pas son origine. Dans le cadre probabiliste, ce phénomène apparaît lorsque la probabilité sur les hypothèses ω et sur $\bar{\omega}$ sont égales à 1 en même temps.

A.4 Conclusion

Cette annexe a été consacrée à une présentation des règles de combinaison proposées afin de résoudre le problème du conflit. Pour les lecteurs intéressés, ils peuvent trouver deux excellentes études comparatives dans [C73, A74]. Les autres règles qui ne sont pas décrites dans cette annexe peuvent être trouvées dans [A61, A91, A87, C27, C72, A20, C17, A21, A35, C9, C44, L12, A47, C76, C82, A79, A95, C71].

En règle générale, les opérateurs de combinaison sont tous fondés sur la combinaison conjonctive. La différence essentielle entre ces règles réside sur la manière de redistribuer la masse conflictuelle sur les autres sous-ensembles du cadre de discernement. Toutefois ces règles de combinaison présentent quelques problèmes :

- La masse conflictuelle est réaffectée sur des propositions moins spécifiques que celles d'origine. De ce fait, la distribution de masse résultante est de qualité dégradée.
- Tous ces opérateurs présentés sont commutatifs, mais ils ne sont pas tous associatifs. Seules quelques règles vérifient cette propriété. Il est donc nécessaire pour la majorité de ces règles de définir une stratégie de fusion dans laquelle on établit l'ordre de fusion des différentes sources. La définition de cette stratégie dépendra fortement de l'application envisagée.
- Toutes ces règles ne sont pas idempotentes. La seule règle idempotente est la règle de combinaison prudente (cautious combination rule) [L20] qui est aussi commutative et associative.
- L'interprétation du conflit selon qu'il provient de la fiabilité défaillante de sources, d'un cadre de discernement non exhaustif ou d'une autre origine n'est pas justifiée avec ces règles de combinaison.



Bibliographie

Livres et chapitres de livres

- [L1] R. Boite, H. Bourlard, T. Dutoit, J. Hancq et H. Leich. *Traitement de la parole*. Presses Polytechniques Universitaires Romandes, Lausanne, 2000.
- [L2] E. Blasch. *Derivation of a belief filter for high range resolution radar, simultaneous target tracking and identification*. A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of philosophy, 1999.
- [L3] S. Blackman et R. Popoli. *Design and analysis of modern tracking systems*. Artech House, 1999.
- [L4] M. Brandstein. *A framework for speech source localization using sensor arrays*. Ph.D. thesis, Brown Univ., Providence, RI, 1995.
- [L5] S. Ben-Yacoub, Y. Abdeljaoued et E. Mayoraz. *Fusion of Face and Speech Data for Person Identity Verification*. IDIAP Research Report, Mars 1999.
- [L6] B. Cobb et P. Shenoy. Dans *A comparison of methods for transforming belief function models to probability models*, pages 255–266. Springer-Verlag, Berlin, 2003.
- [L7] K. Dahia. *Nouvelles méthodes en filtrage particulière - Application au recalage de navigation inertielle par mesures altimétriques*. Thèse de doctorat, 2005.
- [L8] R. Duda, P. Hart et D. Stork. *Pattern Classification*. Wiley, New York, 2000.
- [L9] C. Hue. *Méthode séquentielles de Monte-Carlo pour le filtrage non linéaire multi-objets dans un environnement bruité - Application au pistage multicibles et à la trajectographie d'entités dans les séquences d'images 2D*. Thèse de doctorat, 2003.
- [L10] J. F. Fisher III, T. Darrell, W. Freeman et P. Viola. *Learning joint statistical models for audio-visual fusion and segregation*. Advances in Neural Information Processing Systems 13, 2000.
- [L11] A. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, New York, 1970.
- [L12] W. Liu et A. Bundy. *The combination of different pieces of evidence using incidence calculus*. Technical report, Department of Artificial Intelligence, University of Edinburg, 1992.
- [L13] J. Mariani. *Analyse, Synthèse et Codage de la parole, Traitement du langage Parlé*. Hermes, 2002.
- [L14] H. Martin et H. Hunke. *Locating and tracking of human faces with neural networks*. Technical Report of CMU, CMU-CS-94-155, 1994.

- [L15] H. Rowley et S. Buluja. *Human face detection in visual scenes*. Technical Report of CMU, CMU-CS-95-158R., 1995.
- [L16] F. Smarandache et J. Dezert. *Advances and Applications of DS_mT for Information Fusion*. Rehoboth, American Research Press, 2004.
- [L17] G. Shafer. *Mathematical theory of evidence*. Princeton University Press, Princeton, New Jersey, 1976.
- [L18] G. Shafer. *Calliope, La parole et son traitement automatique*. Masson, 1989.
- [L19] L. Weiru. *Audio-Visual Speaker Tracking with Importance Particle Filters*. DIAP Research Report, 2002.
- [L20] L. Weiru. *Analysing the degree of conflict among belief functions*. Tech Rep, 2004.
- [L21] K. Yow et R. Cipolla. *Feature-based human face detection*. Technical Report No. 249, 1996.
- [L22] L. Zhang. Dans *Advances in the Dempster-Shafer Theory of Evidence.*, pages 51–69. John Wiley & Sons, Inc., 1994.

Articles dans des revues

- [A1] A. Appriou. Probabilités et incertitude en fusion de données multi-senseurs. *Revue scientifique et technique de la defense*, 11 :27–40, 1991.
- [A2] A. Ayoun et P. Smets. Data association in multi-target detection using the transferable belief model. *International Journal of Intelligent Systems*, 16(1) :1167–1182, Septembre 2001.
- [A3] B. Atal. Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification. *Journal of the Acoustical Society of America*, 55(6) :1304–1312, 1974.
- [A4] P. Aarabi et S. Zaky. Robust sound localization using multi-source audio-visual information fusion. *Information Fusion*, 3(2) :209–223, Septembre 2001.
- [A5] M. Brandstein, J. Adcock et H. Silverman. A practical time-delay estimator for localizing speech sources with a microphone array. *Computer Speech and Language*, 9 :153–169, 1995.
- [A6] Y. Bennani et P. Gallinari. Neural networks for discrimination and modeling of speakers. *Speech Communication*, 17(1-2), 1995.
- [A7] M. Beal, N. Jovic et H. Attias. Graphical model for audio-visual object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(7) :828–836, Juillet 2003.
- [A8] F. Bourgeois et J. Lassalle. An extension of the munkres algorithm for the assignment problem to rectangular matrices. *Communications of the Association for Computing Machinery*, 14(12) :802–806, dec 1971.
- [A9] S. Blackman. Multiple hypothesis tracking for multiple target tracking. *IEEE aerospace and Electronic*, 90(1) :5–18, Janvier 2004.
- [A10] F. Bimbot et L. Mathan. Text-free speaker recognition using an arithmetic harmonic sphericity measure. *Eurospeech conference*, 1 :169–172, 1993.

- [A11] F. Bimbot et I. Magrin-Chagnolleau. Second-order statistical measures for text-independent speaker identification. *Speech Communication*, 17(1-2) :177–192, 1995.
- [A12] B. Berdugo, J. Rosenhouse et H. Azhari. Speakers direction finding using estimated time delays in the frequency domain. *Signal Processing*, 82 :19–30, 2002.
- [A13] Y. Bar-Shalom et E. Tse. Tracking in a cluttered environment with probabilistic data association. *Automatica*, 11 :451–460, 1975.
- [A14] J. Campbell. Speaker recognition : A tutorial. *Proceedings of the IEEE*, 85(9) :1437–1462, 1997.
- [A15] A. Cichocki, R. Bogner, L. Moszczynski et K. Pope. Modified hérault-jutten algorithms for blind separation of sources. *Digital Signal Processing*, 7(2) :80 – 93, Avril 1997.
- [A16] S. Choi, A. Cichocki, H.-M. Park et S.-Y. Lee. Blind source separation and independent component analysis : A review. *Neural Information Processing - Letters and Reviews*, 16(1), Janvier 2005.
- [A17] A. Cichocki, J. Karhunen, W. Kasprzak et R. Vigario. Neural networks for blind separation with unknown number of sources. *Neurocomputing*, 24 :55–93, 1999.
- [A18] G. Chow et X. Li. Towards a system for automatic facial features detection. *Pattern Recognition*, 26(12) :1735–1799, 1993.
- [A19] S.-H. Cha et S. N. Srihari. On measuring the distance between histograms. *Pattern Recognition*, 35(6) :1355–1370, Juin 2002.
- [A20] M. Daniel. Distribution of contradictive belief masses in combination of belief functions. *Information, Uncertainty and Fusion*, pages 431–446, jul 2000.
- [A21] M. Daniel. Associativity in combination of belief functions ; a derivation of minc combination. *Soft Computing*, 7(5) :288–296, 2003.
- [A22] T. Denoeux. A k-nearest neighbor classification rule based on dempster shaffer theory. *IEEE Transactions on Systems, Man and Cybernetics*, 25(5) :804–813, Janvier 1995.
- [A23] D. Dubois et H. Prade. A set-theoretic view of belief functions : logical operations and approximations by fuzzy sets. *International Journal of General Systems*, 12 :193–226, 1986.
- [A24] D. Dubois et H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4 :244–264, 1988.
- [A25] D.Zotkin, R.Duraiswami et L.Davis. Joint audio-visual tracking using particle filters. *EURASIP journal on Applied Signal Processing*, 2002(11) :1154–1164, Janvier 2002.
- [A26] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele et P. Yanker. Query by image and video content : the qbic. *IEEE Computer magazine*, 28(9) :23–32, 1995.
- [A27] D. Gavrilu. The visual analysis of human movement : A survey. *Computer Vision and Image Understanding*, 73(1) :82–98, 1999.

- [A28] H. Gish et M. Schmidt. Second-order statistical methods for text-independent speaker identification. *Speech Communication*, 17(1-2), 1995.
- [A29] J. Herault et B. Ans. Réseaux de neurones a synapses modifiables : Décodage de messages sensoriels composites par un apprentissage non supervise et permanent. *CRAS Paris*, 299(13) :525–528, 1984.
- [A30] R. Haenni. Are alternatives to dempster's rule of combination real alternatives ? : Comments on :about the belief function combination and the conflict management problem - lefevre et al. *Information Fusion*, 3(3) :237–239, 2002.
- [A31] E. Hjelm et B. Low. Face detection : A survey. *Computer Vision and Image Understanding*, 83 :236–274, 2001.
- [A32] T. Hofmann, J. Puzicha et J. Buhmann. Textured image segmentation in a deterministic annealing framework. *IEEE Transaction on Pattern Analysis and Machine Intelligent*, 20(8), 1998.
- [A33] B. Heisele, A. Verri et T. Poggio. Learning and vision machine. *Proceedings of the IEEE*, 90(7), 2002.
- [A34] T. Inagaki. Interdependence between safety-control policy and multiple-sensor schemes via dempster- shafer theory. *IEEE Transaction on Reliability*, 40(2) :182–188, 1991.
- [A35] A. Josang. The consensus operator for combining beliefs. *Artificial Intelligence Journal*, 142 :157–170, 2002.
- [A36] S. Julier et J. Uhlmann. A new extension of the kalman filter to nonlinear systems. *Proceeding of the SPIE*, pages 182–193, 1997.
- [A37] T. Kailath. The divergence and bhattacharyya distance measures in signal selection. *IEEE Trans*, 15(1) :52–60, 1967.
- [A38] R. Kalman. A new approach to linear filtering and prediction problems. *ournal of Basic Engineering, Transactions of the ASME*, 82 :35–45, 1960.
- [A39] R. Kennes. Computational aspects of the mÃbius transformation of graphs. *IEEE Transactions on Systems, Man, and Cybernetics*, 22 :201–223, 1992.
- [A40] S. Kullback et R. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22 :79–86, 1951.
- [A41] M. Kirby et L. Sirovich. Application of the karhunen loeve procedure for the characterization of human faces. *Pattern Analysis and Machine Intelligence*, 12(1) :12–16, 1990.
- [A42] H. Kuhn. The hungarien method for assigment problem. *Naval Research Logistics Quaterly*, 2 :83–97, 1955.
- [A43] M. Kass, A. Witkin et D. Terzopoulos. Snakes : Active contour models. *International Journal of Computer Vision*, 4 :321–331, 1998.
- [A44] E. Lefevre, O. Colot et P. Vannoorenberghe. Belief functions combination and conflict management. *Information Fusion Journal*, 3(2) :149–162, 2002.
- [A45] E. Lefevre, O. Colot, P. Vannoorenberghe et D. D. Bruccq. Information et combinaison : les liaisons conflictuelles. *Revue Traitement du Signal*, 18(3) :161–177, 2001.

- [A46] A. Laine et J. Fan. Texture classification by wavelet packet signatures. *IEEE Transaction on Pattern Analysis and Machine Intelligent*, 15 :1186–1191, 1993.
- [A47] W. Liu et J. Hong. Re-investigating Dempster’s idea on evidence combination. *Knowledge and Information Systems*, 2 :223–241, 2000.
- [A48] T.-W. Lee, M. Lewicki, M. Girolami et T. Sejnowski. Blind source separation of more sources than mixtures using overcomplete representations. *IEEE Signal Processing Letters*, 6(4), Avril 1999.
- [A49] D. Li, I.-K. Sethi, N. Dimitrova et T. McGee. Classification of general audio data for content-based retrieval. *Pattern Recognition Letters*, Décembre 2000.
- [A50] Z. Liu, Y. Wang et T. Chen. Audio feature extraction and analysis for scene segmentation and classification. *VLSI Signal Processing Syst. Signal, Image, Video Technol*, 20 :61–79, Octobre 1998.
- [A51] K. Lam et H. Yan. Locating and extracting the eye in human face images. *Pattern Recognition*, 29(5) :771–779, 1996.
- [A52] R. Mahler. Combining ambiguous evidence with respect to abiguous a priori knowledge, i : Boolean logic. *IEEE Transaction on Systems, Man and Cybernetics, Part A : Systems and Humans*, 26(1) :27–41, 1996.
- [A53] J. Makhoul. Linear prediction : A tutorial review. *Proceedings of the IEEE*, 63(4) :561–580, Avril 1975.
- [A54] J. Makhoul. Stable and efficient lattice method for linear prediction. *EEE Transaction on Acoustics, Speech and Signal processing*, 25(5) :423–426, Octobre 1997.
- [A55] K. Matusita. Decision rules based on the distance for problem of fit, two samples and estimation. *The Annals of Mathematical Statistics*, 26 :631–640, 1955.
- [A56] J. Mao et A. Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition*, 25 :173–188, 1992.
- [A57] N. Matsuo, H. Kitagawa et S. Nagata. Speaker position detection system using audio-visual information. *Fujitsu Sci. Tech. J.*, 35(2) :212–220, Décembre 1999.
- [A58] D. Maio et D. Maltoni. Real-time face location on gray-scale static images. *Pattern Recognition*, 33 :1525–1539, 2000.
- [A59] R. Mukai, H. Sawada, S. Araki et S. Makino. Blind source separation for moving speech signals using blockwise ica and residual crosstalk subtraction. *IEICE Transaction on Fundamentals of Electronics, Communications and Computer Sciences*, E87-A(8) :1941–1948, 2004.
- [A60] S. Makino, H. Sawada, R. Mukai et S. Araki. Blind source separation of convolutive mixtures of speech in frequency domain. *IEICE Transaction on Fundamentals of Electronics, Communications and Computer Sciences*, E88-A(7) :1640–1655, 2005.
- [A61] K. Murphy. Combining belief functions when evidence conflicts. *Decision Support Systems*, 29 :1–9, 2000.

- [A62] J. Miao, B. Yin, K. Wang, L. Shen et X. Chen. A hierarchical multiscale and multiangle system for human face detection in a complex background using gravity center template. *Pattern Recognition*, 32(7) :1237–1248, 1999.
- [A63] J. Puzicha, T. Hofmann et J. Buhmann. Histogram clustering for unsupervised segmentation and image retrieval. *Pattern Recognition*, 20(9) :899–909, 1999.
- [A64] A. Pentland, R. W. Picard et S. Sclaroff. Photobook : content-based manipulation of image databases. *Internat. J. Comput. Vision*, 18(3) :233–254, 1996.
- [A65] D. Reid. An algorithm for tracking multiple targets. *IEEE transaction on automatic control*, 24(6) :843–854, Décembre 1979.
- [A66] Y. Rubner, J. Puzicha, C. Tomasi et J. M. Buhmann. Empirical evaluation of dissimilarity measures for color and texture. *Computer Vision and Image Understanding Journal*, 84(1) :25–43, Octobre 2001.
- [A67] D. Rabinkin, R. Renomeron, A. Dahl, J. French, J. Flanagan et M. Bianchi. A dsp implementation of source location using microphone arrays. *Journal of the Acoustical Society of America*, 99(4) :2503, Avril 1996.
- [A68] M. Swain et D. Ballard. Color indexing. *International Journal on Computer Vision*, 7(1) :11–32, 1991.
- [A69] J. Smith et S. Chang. Tools and techniques for color image retrieval in storage and retrieval for image and video databases. *Conference SPIE*, pages 426–437, 1996.
- [A70] R. Sinkhorn. Diagonal equivalence to matrices with prescribed row and column sums. *The American Mathematical Monthly*, 74(4) :402–405, 1967.
- [A71] H. Silverman et S. Kirtman. A two-stage algorithm for determining talker location from linear microphone array data. *Computer, Speech and Language*, 6(2) :129–152, 1992.
- [A72] P. Smets et R. Kennes. The transferable belief model. *International Journal of Intelligent Systems*, 66(2) :191–234, 1994.
- [A73] P. Smets et R. Kruse. The transferable belief model for belief representation. *Uncertainty Management in Information Systems*, pages 343–368, 1996.
- [A74] P. Smets. Analyzing the combination of conflicting belief functions. *submitted*, 2005.
- [A75] S. Satoh, Y. Nakamura, et T. Kanade. Name it : Naming and detecting faces in news videos. *IEEE Multimedia*, 6(1) :22–35, 1999.
- [A76] K. Sobottka et I. Pitas. A novel method for automatic face segmentation, facial feature extraction and tracking. *Signal Processing : Image Communication*, 12(3) :263–281, 1998.
- [A77] N. Strobel, S. Spors et R. Rabenstein. Joint audio-video object localization and tracking. *IEEE Signal Processing Magazine*, 18(1) :22–31, Janvier 2001.
- [A78] E. Saber et A. M. Tekalp. Frontal view face detection and facial feature extraction using color, shape, and symmetry based cost functions. *Pattern Recognition Letters*, 17(8) :669–680, 1998.

- [A79] B. Tonn. An algorithmic approach to combining belief functions. *International Journal of Intelligent Systems*, 11 :463–476, 1996.
- [A80] H. Wang et S. Chang. A highly efficient system for automatic face region detection in mpeg video. *IEEE Transaction on Circuits and systems for Video Technology*, 7(4) :615–628, 1997.
- [A81] Y. Wang, Z. Liu et J.-C. Huang. Multimedia content analysis using both audio and visual clues. *IEEE Signal Processing Magazine*, Novembre 2000.
- [A82] Y. Wang et B. Yuan. A novel approach for human face detection from color images under complex background. *Pattern Recognition*, 34 :1983–1992, 2001.
- [A83] R. Yager. Hedging in the combination of evidence. *Journal of Information and Optimization Science*, 4(1) :73–81, 1983.
- [A84] R. Yager. On the relationships of methods of aggregation of evidence in expert systems. *Cybernetics and Systems*, 16 :1–21, 1985.
- [A85] R. Yager. On the dempster-shafer framework and new combination rules. *Source Information Sciences : An International Journal*, 41(2) :93–137, Mars 1987.
- [A86] H. yan. Human face image processing techniques. *International Journal of Images and Graphics*, 1(2) :197–215, 2001.
- [A87] J. Yen. Can evidence be combined in the dempster-shafer theory. *International Journal of Approximate Reasoning*, 2(3) :346–347, 1988.
- [A88] H. Yao et W. Gao. Face detection and location based on skin chrominance and lip chrominance transformation from color images. *Pattern Recognition*, 34 :555–564, 2001.
- [A89] M. Yang, D. Kriegman et N. Ahuja. Detecting faces in images : A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1) :34–58, 2002.
- [A90] O. Yylmaz et S. Rickard. Blind separation of speech mixtures via time-frequency masking. *IEEE Transactions on Signal Processing*, 52(7) :1830–1847, Juillet 2004.
- [A91] Y. Yao et S. Wong. Representation, propagation and combination of uncertain information. *International Journal of General Systems*, 23 :59–83, 1994.
- [A92] D. Ying et N. Yasuaki. Face-texture model based on sgld and its application in face detection in a color scene. *Pattern Recognition*, 29(6) :1007–1017, 1996.
- [A93] J.-M. Ye, X.-L. Zhu et X.-D. Zhang. Adaptive blind separation with an unknown number of sources. *Neural Computation*, 16(8) :1641 – 1660, Août 2004.
- [A94] M. Zhang et M. Er. Simple unified algorithm for estimating and tracking location of desired talker by microphone arrays. *Circuits Systems Signal Processing*, 15(6) :749–762, 1996.
- [A95] B. Zeigler. Some properties of modified dempster-shafer operators in rule based inference systems. *International Journal of General Systems*, 14 :345–356, 1988.

Communications

- [C1] H. Asoh, F. Asano, K. Yamamoto, T. Yoshimura, Y. Motomura, N. Ichimura, I. Hara et J. Ogata. An application of a particle filter to bayesian multiple sound source tracking with audio and video information fusion. Dans *Proc. Fusion*, pages 805–812, 2004.
- [C2] F. Asano, Y. Motomura, H. Asoh, T. Yoshimura, N. Ichimura, K. Yamamoto, N. Kitawaki et S. Nakamura. Detection and separation of speech segment using audio and video information fusion. Dans *8th European Conference on Speech Communication and Technology : Eurospeech*, 2003.
- [C3] M. Augustejin et T. Skujca. Identification of human faces through texture-based feature recognition and neural network technology. Dans *IEEE Conference on Neural Networks*, pages 392–398, Rhodes (Greece), 1993.
- [C4] M. Bett, R. Gross, H. Yu, X. Zhu, Y. Pan, J. Yang et A. Weibel. Multimodal meeting tracker. Dans *Recherche d'Informations assistée par Ordinateur*, Paris, 2000.
- [C5] S. Blackman. Multiple target tracking with radar application. Artech House, 1986.
- [C6] Y. Bar-Shalom et T. Fortmann. Tracking and data association. Academic Press, 1989.
- [C7] Y. Bar-Shalom et X. Li. Multi target- multisensor tracking principles and techniques. YBS publishing, Storrs, CT, 1995.
- [C8] P. Bofill et M. Zibulevsky. Blind separation of more sources than mixtures using sparsity of their short-time fourier transform. Dans *International Workshop Independent Component Anal. Blind Signal Separation*, pages 87–92, Helsinki, Finland, Juin 2000.
- [C9] M. Cattaneo. Combining belief functions issued from dependent sources. Dans *Symposium on Imprecise Probabilities and Their Applications*, 2003.
- [C10] A. Clerentin, L. Delahoche, E. Brassart et C. Cauchois. An uncertainty propagation architecture for the localization problem. Dans *Workshop on Performance Metrics for Intelligent Systems PerMIS2002, National Institut of Standards and Technology*, pages 876–880, Washington, Août 2002.
- [C11] A. Clerentin, L. Delahoche et C. Pagar. Omnidirectional sensors cooperation for multi-target tracking. Dans *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, Baden-Baden, Août 2001.
- [C12] R. Cappelli, D. Maio et D. Maltoni. Combining fingerprint classifiers. Dans *International Workshop on Multiple Classifier Systems*, pages 351–361, Juillet 2000.
- [C13] D. Chai et K. Ngan. Locating facial region of a head and shoulders color images. Dans *International Workshop on Automatic Face and Gesture Recognition*, pages 124–129, 1998.
- [C14] B. Cobb et P. Shenoy. On transforming belief function models into probability models. Dans *Technical Report 293*, University of Kansas School of Business, 2004.

- [C15] F. Caron, P. Smets, E. Duflos et P. Vanheeghe. Multisensor data fusion in the frame of the tbm on reals application to land vehicle positioning. Dans *International Conference on Information Fusion*, Philadelphie, Juillet 2005.
- [C16] N. Checka, K. Wilson, M. Siracusa et T. Darrell. Multiple person and speaker activity tracking with a particle filter. Dans *International Conference on Acoustics, Speech, and Signal Processing*, 2004.
- [C17] M. Daniel. Associativity and contradiction in combination of belief functions. Dans *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, pages 133–140, Universidad Politecnica de Madrid, jul 2000.
- [C18] M. Daniel. Transformations of belief functions to probabilities. Dans *Workshop on Uncertainty Processing in Expert Systems*, pages 77–90, 2003.
- [C19] A. Dempster. A generalization of bayesian inference. Dans *Journal of Royal Statistical Society, series B*, 30, pages 205–247, 1968.
- [C20] T. Denoeux. Construction de fonctions de croyance à partir de similarités. Dans *Réunion Française sur les Fonctions de Croyance*, Université de technologie de Compiègne, Mars 2005.
- [C21] T. Darrell, J. Fisher, P. Viola et B. Freeman. Audio-visual segmentation and the cocktail party effect. Dans *International Conference on Multimodal Interfaces*, Beijing, Octobre 2000.
- [C22] H. Dang, C. Han et D. Gruyer. Combining imm filtering and ds data association for multitarget tracking. Dans *IEEE International Conference in Information Fusion*, pages 876–880, Stockholm, Juillet 2004.
- [C23] Y. Dai et Y. Nakano. Extraction for facial images from complex background using color information and sgld matrices. Dans *International workshop on automatic Face and Gesture Recognition*, pages 238–242, 1995.
- [C24] L. Davis, V. Philomin et R. Duraiswami. Tracking humans from a moving platform. Dans *International Conference on Pattern Recognition*, pages 4171–4178, Spain, 2000.
- [C25] S. Fine, J. atil et R. Gopinath. A hybrid gmm/svm approach to speaker identification. Dans *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2001.
- [C26] J. Fisher et T. Darrell. Signal level fusion for multimodal perceptual user interface. Dans *Workshop on Perceptive User Interfaces*, 2001.
- [C27] S. Ferson et V. Kreinovich. Representation, propagation, and aggregation of uncertainty. Dans *SAND Report [in progress]*.
- [C28] D. Gruyer et V. Berge-Cherfaoui. Multi-objects association in perception of dynamical situation. Dans *Fifteenth Conference on Uncertainty in Artificial Intelligence, UAI'99*, Stockholm, Juillet 1999.
- [C29] H. Graf, E. Cosatto, D. Gibbon, M. Kocheisen et E. Petajan. Multimodal system for locating heads and faces. Dans *International Conference on Automatic Face and Gesture Recognition*, pages 88–93, 1996.
- [C30] H. Graf, T. Chen, E. Petajan et E. Cosatto. Locating faces and facial parts. Dans *International Workshop on Automatic Face and Gesture Recognition*, pages 41–46, 1995.

- [C31] D. Gruyer, C. Royère et V. Berge-Cherfaoui. Credibilist multi-sensors fusion for the mapping of dynamic environment. Dans *IEEE International Conference in Information Fusion*, paris, Juillet 2000.
- [C32] D. Gruyer, C. Royère et V. Berge-Cherfaoui. Heterogeneous multi-criteria combination with partial or full information. Dans *6th International Conference on Information Fusion, FUSION'03*, Australia, Juillet 2003.
- [C33] R. Haenni. Shedding new light on zadeh's criticism of dempster's rule of combination. Dans *International Conference on International Fusion*, Philadelphie, Juillet 2005.
- [C34] P. Heckbert. color image quantization for frame buffer display. Dans *SIGGRAPH*, pages 297–307, 1982.
- [C35] J. H'erault, C. Jutten et B. Ans. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. Dans *Actes du X^{eme} colloque GRETSI*, pages 1017–1022, Nice, France, Mai 1985.
- [C36] J. Hershey et J. Movella. Using audio-visual synchrony to locate sounds. Dans *Neural Information Processing Systems*, 1999.
- [C37] W. Huadong, M. Siegel et S. Ablay. Sensor fusion using dempster-shafer theory ii : Static weighting and kalman filter-like dynamic weighting. Dans *IEEE Instrumentation and Measurement Technology conference*, Anchorage, AK, USA, 2003.
- [C38] W. Huandong, M. Siegel, R. Stiefelhagen et J. Yang. Sensor fusion using dempster-shafer theory. Dans *IEEE Instrumentation and Measurement Technology conference*, Anchorage, AK, USA, 2002.
- [C39] G. Iyengar, H. Nock et C. Neti. Audio-visual synchrony for detection of monologues in video archives. Dans *International Conference on Acoustics, Speech, and Signal Processing*, 2003.
- [C40] M. Jabbour, V. Cherfaoui et P. Bonnifait. Management of landmarks in a gis for an enhanced localisation in urban areas. Dans *IEEE Intelligent Vehicle Symposium*, 2006.
- [C41] A. Josang, M. Daniel et P. Vannoorenberghe. Strategies for combining conflicting dogmatic beliefs. Dans *International Conference on International Fusion*, pages 1133–1140, Cairns, Queensland, Australia, 2003.
- [C42] A. Jacquin et A. Eleftheriadis. Automatic location and tracking of faces and facial features in video sequence. Dans *International Workshop on Automatic Face and Gesture Recognition*, pages 237–242, Zurich, Switzerland, 1995.
- [C43] T. Jebara, K. Russell et A. Pentland. Mixtures for eigenfeatures for real-time structure from texture. Dans *IEEE International Conference on Computer Vision*, pages 128–135, 1998.
- [C44] A. Joshi, S. Sahasrabudhe et K. Shankar. Sensitivity of combination schemes under conflicting conditions and a new method. Dans *Advances in Artificial Intelligence*, pages 39–48, 1995.
- [C45] W. Kellermann. A self-steering digital microphone array. Dans *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 3581–3584, Toronto, Canada, 1991.

- [C46] S.-H. Kim, N. k. Kim, S. Ahn et H. g. Kim. Object oriented face detection using range and color information. Dans *International Workshop on Automatic Face and Gesture Recognition*, pages 76–81, 1998.
- [C47] L. Khoudour, T. Leclercq, J.-L. Bruyelle et A. Flancquart. Local camera network for surveillance in newcastle airport. Dans *The IEE IDSS*, London, 2003.
- [C48] T. Kurien. Issues in the design of practical multitarget tracking algorithms. in y. bar-shalom (ed.) multitarget-multisensor tracking : Advanced applications. MA : Artech House, 1990, ch. 3.
- [C49] J. LeBlanc et P. Leon. Source separation of speech signals using kurtosis maximization. Dans *Allerton Conference on Communications, Control, and Computing*, Moticello, Septembre 1997.
- [C50] Y. Lemeret, E. Lefevre et D. Jolly. Tracking cars and prediction of their trajectories. Dans *IEEE International Conference on Machine Intelligence*, 2005.
- [C51] Y. Lemeret, E. Lefevre et D. Jolly. An association algorithm for tracking multiple moving objects. Dans *IEEE Intelligent Transport System Conference*, 2006.
- [C52] N. Megherbi et S. Ambellouis. Extraction de visage par contour actif et algorithme génétique. Dans *Conference Traitement et Analyse de l'Information Méthodes et Applications*, pages 345–351, Hammamet, Tunisia, Octobre 2005.
- [C53] N. Megherbi et S. Ambellouis. A genetic snake model to automatic human face and head boundary extraction. Dans *International Symposium On Programming and Systems*, pages 289–296, Alger, Algeria, Mai 2005.
- [C54] J. Malik, S. Belongie, J. Shi et T. Leung. Textons, contours and regions : Cue integration in image segmentation. Dans *IEEE International Conference on Computer Vision (ICCV'99)*, 1999.
- [C55] B. Mourllion, D. Gruyer, C. Royere et S. Theroude. Multi-hypotheses tracking algorithm based on the belief theory. Dans *IEEE International Conference in Information Fusion*, Philadelphia, PA, USA, Juillet 2005.
- [C56] E. Menegatti, E. Mumolo, M. Nolich et E. Pagello. A surveillance system based on audio and video sensory agents cooperating with a mobile robot. Dans *International Conference on Intelligent Autonomous Systems*, pages 335–343, The Netherlands, 2004.
- [C57] H. Nock, G. Iyengar et C. Neti. Speaker localisation using audio-visual synchrony : An empirical study. Dans *International Conference on Image and Video Retrieval*, 2003.
- [C58] J. Oglesby et J. Mason. Radial basis function networks for speaker recognition. Dans *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1991.
- [C59] S. Park et J. Aggarwal. Segmentation and tracking of interacting human body parts under occlusion and shadowing. Dans *IEEE Workshop on Motion and Video Computing*, pages 105–111, Orlando, 2002.
- [C60] J. Puzicha, T. Hofmann et J. Buhmann. Non-parametric similarity measures for unsupervised texture segmentation and image retrieval. Dans *computer Vision and Pattern Recognition Conference*, pages 267–272, 1997.

- [C61] D. Pham, C. Servière et H. Boumaraf. Blind separation of speech mixtures based on non stationarity. Dans *ISSPA*, Paris, Juillet 2003.
- [C62] G. Pass, R. Zabih et J. Miller. Comparing images using color coherence. Dans *ACM International Multimedia Conference*, pages 65–73, New York, 1996.
- [C63] R. Qian et T. Huang. Object detection using hierarchical mrf and map estimation. Dans *IEEE International Conference on Image Processing*, pages 131–135, 1998.
- [C64] W. H. Q.B. Sun et J. Wu. face detection based on color and local symmetry. Dans *In Proceedings of the Third International Workshop on Automatic Face and Gesture Recognition*, pages 130–135, 1998.
- [C65] M. Rombaut et V. B. Cherfaoui. Decision making in data fusion using dempster schaffer's theory. Dans *IFAC Symposium on Intelligent Components and Instrumentation for Control Applications*, Annecy, France, Juin 1997.
- [C66] J.-L. Rouas, J. Louradour et S. Ambellouis. Audio events detection in public transport vehicle. Dans *IEEE Intelligent Transport System Conference*, 2006.
- [C67] M. Rahurkar, A. Sethi et T. S. Huang. Robust speaker tracking by fusion of complementary features from audio and video modalities. Dans *Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, 2005.
- [C68] Y. Rubner, C. Tomasi et L. J. Guibas. A metric for distributions with applications to image databases. Dans *International Conference in Computer Vision*, pages 59–66, 1998.
- [C69] D. Reisfeld et Y. Yeshurum. Robust detection of facial features by generalized symmetry. Dans *Proceedings of the 11th International Conference on Pattern Recognition*, pages 117–120, Hague, Netherland, 1992.
- [C70] D. Sturim, M. Brandstein et H. Silverman. Tracking multiple talkers using microphone-array measurements. Dans *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume I, pages 371–374, Munich, Germany, 1997.
- [C71] F. Smarandache et J. Dezert. A simple proportional conflict redistribution rule. Dans *rxiv Archives, Los Alamos National Laboratory*, jul 2004.
- [C72] F. Smarandache et J. Dezert. Information fusion based on new proportional conflict redistribution rules. Dans *International Conference on International Fusion*, Philadelphie, Juillet 2005.
- [C73] K. Sentz et S. Ferson. Combination of evidence in dempster-shafer theory. SANDIA Tech. Report, SAND2002-0835, Avril 2002.
- [C74] M. Schmidt et H. Gish. Radiaspeaker identification via support vector machines. Dans *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1996.
- [C75] H. Sawhney et J. Hafner. Efficient color histogram indexing. Dans *International Conference on Image Processing*, pages 66–70, New York, 1994.
- [C76] S. Shi, M. Hull et D. Bell. A new rule for updating evidence. Dans *In Z. W. Ras and M. Zemankova, editors, Methodologies for Intelligent Systems, 8th International Symposium, ISMIS '94*, pages 95–104, 1994.

- [C77] P. Smets. The nature of the unnormalized beliefs encountered in the transferable belief model. Dans *Conference on Uncertainty in Artificial Intelligence*, pages 292–297, San Mateo, California, 1992.
- [C78] N. Strobel, T. Meier et R. Rabenstein. Speaker localization using steered filtered-and-sum beamformers. Dans *Vision, Modeling, and Visualization' 99*, pages 195–202, Erlangen, 1999.
- [C79] K. Sobottka et I. Pitas. A fully automatic approach to facial feature, detection and tracking. Dans *International Conference on Audio- and Video-based Biometric Person Authentication*, Switzerland, Mars 1997.
- [C80] N. Strobel et R. Rabenstein. Robust speaker localization using a microphone array. Dans *European Signal Processing Conference*, volume III, pages 1409–1412, 2000.
- [C81] P. Smets et B. Ristic. Belief function theory on the continuous space with an application to model based classification. Dans *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, Italy, Juillet 2004.
- [C82] J. Sudano. A generalized belief fusion algorithm. Dans *International Conference on International Fusion*, pages 1126–1132, 2003.
- [C83] A. Tsukamoto, C. Lee et S. Tsuji. Detection and pose estimation of human face with synthesized templates. Dans *Asian Conference on Computer Vision*, pages 138–186, 1993.
- [C84] M. Turk et A. Pentland. Face recognition using eigenfaces. Dans *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 586–591, Juin 1991.
- [C85] V.-T. Vu, F. Bremond, G. Davini, M. Thonnat, Q.-C. Pham, N. Allezard, P. Sayd, J.-L. Rouas, S. Ambellouis et A. Flancquart. Audio-video event recognition system for public transport security. Dans *IEE Imaging for Crime Detection and Prevention conference*, 2006.
- [C86] J. Vermaak, A. Blake, M. Gangnet et P. Pérez. Sequential monte carlo fusion of sound and vision for speaker tracking. Dans *IEEE International Conference on Computer Vision*, pages 741–746, Canada, 2001.
- [C87] H. Wang et P. Chu. Voice source localization for automatic camera pointing system in videoconferencing. Dans *IEEE International Conference Acoustics, Speech, and Signal Processing*, volume I, pages 187–90, Munich, 1997.
- [C88] K. Wilson, N. Checka, D. Demirdjian et T. Darrell. Audio-video array source separation for perceptual user interfaces. Dans *Workshop on Perceptive User Interfaces*, 2001.
- [C89] H.-C. Wu, J. Principe et D. Xu. Exploring the time-frequency microstructure of speech for blind source separation. Dans *International Conference on Acoustics, Speech, and Signal Processing*, pages 1145–1148, Mai 1998.
- [C90] E. Wong et S. Sridharan. comparison of linear prediction cepstrum coefficients and mel frequency cepstrum coefficients for language identification. Dans *International Symposium on Intelligent Multimedia, Video and Speech Processing*, Hong Kong, Mai 2001.
- [C91] M. H. Yang et N. Ahuja. Detecting human faces in color images. Dans *IEEE International Conference on Image Processing*, pages 127–139, Chicago, Octobre 1998.

- [C92] K. Yow et R. Cipolla. Enhancing human face detection using motion and active contours. Dans *Asian Conference on Computer Vision*, pages 515–522, 1998.
- [C93] J. Yang, W. Lu et A. Waibel. Skin-color modeling and adaption. Dans *Proceedings of the AVSP'97 workshop*, pages 45–54, Rhodes (Greece), Septembre 1997.
- [C94] Y.Bar-Shalom et T. Fortmann. tracking and data association. Academic Press, 1988.
- [C95] J. Yang et A. Waibel. A real-time face tracker. Dans *Third Workshop Applications of Computer Vision*, pages 142–147, 1996.
- [C96] T. Yokoyama, Y. Yagi et M. Yachida. Active contour model for extracting human faces. Dans *International Conference on Pattern Recognition*, pages 673–676, Brisbane, Australaria, Septembre 1998.
- [C97] T. Yokoyama, Y. Yagi et M. Yachida. Facial contour extraction model. Dans *IEEE International Conference on Face and Gesture Recognition*, pages 254–259, Nara, Japan,, Avril 1998.
- [C98] J. Yang, X. Zhu, R. Gross, J. Kominek, Y. Pan et A. Weibel. Multimodal people id for a multimedia meeting browser. Dans *Proceedings of ACM Multimedia*, 1999.

Thèses de Doctorat

- [T1] G. Cook. *Data Selection and Model Combination in Connectionist Speech Recognition*. PhD thesis, Department of Engineering, University of Cambridge, UK, 1997.
- [T2] P. Fabiani. *Représentation dynamique de l'incertain et stratégie de prise d'informations pour un système autonome en environnement évolutif*. PhD thesis, ONERA-Toulouse, 1996.
- [T3] D. Gruyer. *Etude du traitement de données imparfaites pour le suivi multi-objets : Application aux situations routières*. PhD thesis, Université de Technologie de Compiègne, 1999.
- [T4] S. Johnson. *Speaker Tracking*. PhD thesis, Department of Engineering, Cambridge University, UK, 1997.
- [T5] E. Lefevre. *Fusion adaptée d'informations conflictuelles dans le cadre de la théorie de l'évidence*. PhD thesis, Institut National des Sciences Appliquées de Rouen, 2001.
- [T6] C. Royère. *Contribution à la résolution du conflit dans le cadre de la théorie de l'évidence. Application à la perception et à la localisation des véhicules intelligents*. PhD thesis, Université de Technologie de Compiègne, 2002.

Articles de l'auteur en rapport avec le sujet de thèse

- [1] N. Megherbi et S. Ambellouis. A genetic snake model to automatic human face and head boundary extraction. In *Proceeding of The Seventh International Symposium On Programming and Systems*, pages 289–296, Alger, Algeria, May 2005.
- [2] N. Megherbi et S. Ambellouis. Extraction de visage par contour actif et algorithme génétique. In *Proceeding of Traitement et Analyse de l'Information Méthodes et Applications TAIMA*, pages 345–351, Hammamet, Tunisia, October 2005.
- [3] N. Megherbi, S. Ambellouis, O. Colot et F. Cabestaing. Multimodal data association based on the use of belief functions for multiple target tracking. In *Proceeding of the 2005 IEEE Eighth International conference on information fusion*, pages 345–351, Philadelphia, July 2005.
- [4] N. Megherbi, S. Ambellouis, O. Colot et F. Cabestaing. Joint audio-video people tracking using belief theory. In *Proceeding of the 2005 IEEE International conference on Advanced Video and Signal-based Surveillance*, pages 135–140, Como, Italy, September 2005.
- [5] N. Megherbi, S. Ambellouis, O. Colot et F. Cabestaing. Data association in multi-target tracking using belief theory : Handling target emergence and disappearance issue. In *Proceeding of the 2005 IEEE International conference on Advanced Video and Signal-based Surveillance*, pages 517–521, Como, Italy, September 2005.

Articles de l'auteur hors du cadre de la thèse

- [6] N. Megherbi et K. Benmahammed. Orientation field estimation of fingerprint images. In *Proceeding of the 2th National Seminar in electric*, pages 122–125, Biskra, Algeria, September 1999.
- [7] H. Megherbi, N. Megherbi et K. Benmahammed Un algorithme génétique pour la binarisation des images des empreintes digitales. In *Proceeding of the 1th National Seminar in control and signal processing*, pages 284–289, Djelfa, Algeria, 2001.
- [8] N. Megherbi et H. Megherbi. An asexual genetic algorithm to fingerprint image binarization. In *Proceeding of the 4th National Seminar in data processing*, pages 43–50, Biskra, Algeria, 2004.

