



Vers le suivi d'objets dans un cadre évidentiel : représentation, filtrage dynamique et association

Wafa Rekik

► To cite this version:

Wafa Rekik. Vers le suivi d'objets dans un cadre évidentiel : représentation, filtrage dynamique et association. Intelligence artificielle [cs.AI]. Université Paris Sud - Paris XI; Université de Sfax (Tunisie), 2015. Français. NNT : 2015PA112040 . tel-01166041

HAL Id: tel-01166041

<https://theses.hal.science/tel-01166041>

Submitted on 22 Jun 2015

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

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

République Tunisienne
Ministère de l'Enseignement Supérieur,
de la Recherche Scientifique



Université de Sfax
École Nationale d'Ingénieurs de Sfax

Université PARIS-SUD



*Ecole Doctorale SFAX
Sciences et Technologies
Thèse de DOCTORAT
En Génie Electrique*

*Ecole Doctorale STITS
Institut d'Electronique
Fondamentale
Thèse de DOCTORAT
En Physique*

N° d'ordre: 2015–609

THESE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE DE SFAX

*Discipline : Génie Electrique
et*

DOCTEUR DE L'UNIVERSITE PARIS SUD

Discipline : Physique

Présentée et soutenue publiquement à

L'Ecole Nationale d'Ingénieurs de Sfax

Par

Wafa REKIK
(Mastère ENT)

Représentation, Filtrage Dynamique et Association dans un Cadre Evidentiel : Vers le Suivi d'Objets

Soutenu le 23 Mars 2015, devant le jury composé de :

- | | | |
|------|---|-----------------------|
| M. | Ezzedine BEN BRAIK (Pr. ENSIT, Université de Tunis, Tunisie) | , Président |
| M. | Arnaud MARTIN (Pr, Université de Rennes 1, France) | , Rapporteur |
| M. | Imed Riadh FARAH (MC. ISAMM, Université de Manouba, Tunisie) | , Rapporteur |
| M. | Jean DEZERT (Dr, Chercheur Senior, ONERA, France) | , Examineur |
| Mme. | Sylvie LE HEGARAT-MASCLE (Pr, Université Paris Sud, France) | , Directrice de Thèse |
| M. | Ahmed BEN HAMIDA (Pr. ENIS, Université de Sfax, Tunisie) | , Directeur de Thèse |
| M. | Abdelaziz KALLEL (MC. ENET' COM, Université de Sfax, Tunisie) | , Invité |

A mes chers grands parents,

A mes chers parents ‘Fatma’ et ‘Abdelmonem’ pour les efforts et les sacrifices qu’ils n’ont jamais cessé de consentir pour mon instruction et mon bien-être,

A mon cher mari ‘Aymen’ pour son amour, sa patience, sa confiance en moi, ses encouragements et ses vraies qualités humaines,

A mon futur bébé ‘Manar’ pour la joie et le bonheur qu’elle nous apportera,

A mes chères soeurs ‘Wiem’, ‘Wissal’, ‘Wisseem’ pour leur soutien, amour et compréhension,

A ma belle famille,

A tous ceux que j’aime et à tous ceux qui me sont chers,

Je dédie cette thèse...

Wafa

Remerciement

Ce travail faisant l'objet de la présente **thèse de doctorat** s'inscrit dans le cadre d'une thèse en **co-tutelle** entre l'Unité de recherche **ATMS** (*Advanced Technologies for Medicine & Signals*) de l'**ENIS** (Ecole Nationale des Ingénieurs de Sfax) et le Laboratoire **IEF** (Institut d'Electronique Fondamentale) notamment le département de recherche **ACCIS** (Architectures, Contrôle, Communication, Images, Systèmes) de l'**Université Paris Sud** et du **CNRS**.

Je profite de l'occasion qui m'est offerte par le biais de la présentation de ce travail pour exprimer mes gratitudes et mes reconnaissances envers tous ceux qui ont permis à cette thèse de voir le jour.

D'abord, je tiens à exprimer mes remerciements les plus sincères à Madame **Sylvie Le Hégarat-Mascle** ma Directrice de thèse, Professeur à l'Université Paris Sud et Responsable de l'équipe Vision & Image de ACCIS, et à mon Directeur de thèse Monsieur **Ahmed Ben Hamida**, Professeur à l'ENIS et Directeur de recherche de l'Unité ATMS, pour la confiance qu'ils m'ont accordée en acceptant d'encadrer et diriger mes travaux, leurs conseils judicieux, leurs aides et surtout pour le travail effectué durant mes années de thèse. Qu'ils trouvent ici un témoignage de ma profonde reconnaissance.

Je tiens à remercier vivement mon superviseur Monsieur **Abdelaziz Kallel** HDR, Maître des Conférences à l'ENET'COM Sfax (Tunisie) et membre de l'Unité ATMS qui a tout essayé pour faire avancer et réussir ce travail.

J'adresse également mes remerciements à Monsieur **Roger Reynaud**, Professeur à l'Université Paris Sud et Directeur du département de recherche ACCIS pour m'avoir accepté à ACCIS durant ma période en France.

Je tiens à exprimer mon profond respect à Monsieur **Ezzedine Ben Braik**, Professeur à l'Ecole Nationale Supérieure d'Ingénieurs de Tunis (Tunisie) qui m'a fait l'honneur d'accepter la présidence de mon jury.

J'exprime ma gratitude à Monsieur **Imed Riadh Farah**, Maître des Conférences à l'Institut Supérieur des Arts Multimédias de la Mannouba (Tunisie), et Monsieur **Arnaud Martin**, Professeur à l'Université de Rennes I (France), d'avoir accepté la lourde tâche de rapporteur et d'avoir consacré un temps précieux à l'examen de ce rapport. Pour la même occasion, j'adresse mes remerciements à Monsieur **Jean Dézert**, Chercheur Senior à l'ONERA (France) d'avoir accepté d'examiner et juger cette thèse.

J'adresse également mes remerciements aux différents membres de notre Unité de recherche **ATMS** ainsi qu'à tout le personnel du département **Génie Electrique de**

l'ENIS, pour leurs encouragements, leurs aides, leurs soutiens ainsi que leurs disponibilités.

Je m'adresse également des expressions de reconnaissance aux membres de l'équipe de recherche **ACCIS** ainsi que tout le personnel de l'**Université Paris Sud** pour leurs aides ainsi que leurs disponibilités.

Finalement, je tiens à exprimer mes reconnaissances à tous mes professeurs de l'**ENIS** pour la formation qu'ils ont eu le soin de m'apporter tout au long de mes études universitaires.

Table des matières

Remerciement	ii
Table des matières	iv
Listes des Figures	viii
Liste des Tableaux	x
Abréviations	xi
1 Introduction générale	1
1.1 Contexte applicatif	1
1.2 Fusion de données par fonctions de croyance	2
1.3 Plan de la thèse	3
2 Généralités sur le <i>tracking</i>	5
2.1 Introduction	5
2.2 Détection d'objets	6
2.3 Représentation d'objets	7
2.4 Techniques de suivi d'objet(s)	8
2.4.1 Filtre de Kalman	9
2.4.2 Filtre particulière SIR	11
2.4.3 Fenêtrage	13
2.4.4 Association	13
2.4.4.1 NN et GNN	14
2.4.4.2 PDAF et JPDAF	15
2.4.4.3 MHT	16
2.5 Présentation des données	17
2.5.1 Séquence <i>Digitéo ComFuCet</i>	17
2.5.2 Séquence <i>SafeAround</i>	18
2.6 Métriques d'évaluation	19
2.6.1 Définitions des VP, FP et FN pour le sous-problème d'association	19
2.6.2 Définition des VP, FP et FN pour le sous-problème de filtrage	20
2.7 Conclusion	20
3 Généralités sur les Fonctions de croyance	23

3.1	Introduction	23
3.2	Choix du TBM	24
3.3	Aperçu sur le TBM	25
3.4	Niveau Crédal : partie statique	26
3.4.1	Cadre de discernement	26
3.4.2	Fonction de masse (bba)	27
3.4.3	Transformations de la fonction de masse	28
3.5	Niveau Crédal : partie dynamique	30
3.5.1	Notions de redondance et de complémentarité des sources	30
3.5.2	Notion d'indépendance des sources	31
3.5.3	Affaiblissement des croyances	31
3.5.4	Principe du moindre engagement et ordonnancement	32
3.5.5	Conditionnement	34
3.5.6	Déconditionnement	34
3.5.7	Combinaison des croyances	35
3.5.8	Décomposition canonique	37
3.5.9	Opérations sur les cadres de discernement	39
3.5.9.1	Première application pratique : combinaison de bbas définies dans des cadres de discernement indépendants	41
3.5.9.2	Deuxième application pratique : conditionnement	42
3.6	Niveau Décisionnel	42
3.6.1	Maximum de probabilité pignistique	43
3.6.2	Maximum de plausibilité / Maximum de crédibilité	44
3.7	Conflit	44
3.8	Conclusion	45
4	Approche crédibiliste pour la construction dynamique d'objets	46
4.1	Introduction	46
4.2	Définition du problème	47
4.3	Représentation crédibiliste des objets [1, 2]	49
4.3.1	Combinaison disjonctive	50
4.3.2	Conditionnement temporel	51
4.3.3	Conditionnement spatial	52
4.4	Généralités sur l'association de données	53
4.4.1	Fonction coût	53
4.4.2	Types d'association	54
4.4.2.1	Type 1-1	54
4.4.2.2	Type 1-N	55
4.5	Etat de l'art des méthodes d'association dans le cadre du TBM	56
4.5.1	Approches fondées sur la définition de bbas d'associations	56
4.5.1.1	Approche de Gruyer et al : bbas d'association spécialisées [3–5]	59
4.5.1.2	Approche de Mercier et al [6]	61
4.5.1.3	Approche de Denœux et al. : association optimale [7, 8]	62
4.5.2	Bbas de description des objets	64
4.5.2.1	Coûts crédibilistes	64
4.5.2.2	Plausibilité globale	65

4.6	Contribution sur l'association de données	68
4.6.1	Fonctions coût	68
4.7	Exemple jouet	71
4.8	Résultats	73
4.8.1	Validation qualitative	73
4.8.2	Validation quantitative	74
4.9	Conclusion	78
5	Estimation dynamique du cadre de discernement	80
5.1	Introduction	80
5.2	Etat de l'art des méthodes d'estimation du cadre de discernement	81
5.2.1	Estimation du cadre de discernement à partir de sources homogènes	82
5.2.2	Estimation automatique du cadre de discernement dans le cas d'une classification non supervisée	84
5.2.3	Estimation automatique du cadre de discernement à partir de sources hétérogènes	85
5.3	Définition du problème (cas général)	87
5.3.1	Produit cartésien	87
5.3.2	Hypothèse d'indépendance	88
5.3.3	Cadre de discernement commun	89
5.3.4	Application 1 : classification au niveau pixel	90
5.3.5	Application 2 : détection d'objets d'intérêt	90
5.4	Méthode proposée [9–11]	90
5.4.1	Actualisation du produit cartésien Ω	91
5.4.2	Actualisation des bbas définies sur Ω	92
5.4.3	Actualisation du cadre de discernement Θ	94
5.4.3.1	Contrainte sur le cardinal de Θ	95
5.4.4	Exemple jouet	95
5.5	Application	99
5.5.1	Données en entrée et relation entre objets $\equiv$	100
5.5.2	Définition de la bba élémentaire	100
5.5.3	Aspects temporels	102
5.5.4	Aspects numériques	104
5.5.5	Récapitulatif des paramètres de la méthode	106
5.6	Résultats	108
5.6.1	Résultats sur des données simulées	108
5.6.2	Résultats sur des données réelles	109
5.7	Conclusion	113
6	Vers le suivi d'objets	115
6.1	Introduction	115
6.2	Approche de suivi	116
6.2.1	Détections fragmentaires	117
6.3	Exemples de résultats	119
6.3.1	Résultats sur des données simulées	120
6.3.2	Résultats sur des données réelles	120
6.4	Conclusion	122

7 Conclusion générale	124
7.1 Bilan des travaux	124
7.2 Perspectives	126
A Annexe au chapitre 3	128
A.1 Algorithme Hongrois	128
A.2 Modification de l'algorithme Hongrois	132
A.3 Détection de changement	132
A.3.1 Estimation de l'image de fond	134
A.3.1.1 Filtre sigma-delta	134
A.3.1.2 Modèle codebook	135
A.3.2 Détection des différences par rapport à l'image de fond	136
B Annexe au chapitre 4	137
B.1 Démonstration de l'équation 5.7	137
Bibliographie	139
Résumé	154
Abstract	155
Liste des publications	156

Table des figures

2.1	Représentations d'un objet : (a) centroïde , (b) points multiples, (c) rectangle, (d) ellipse, (e) forme articulée, (f) squelette, (g) contour sous forme de points de contrôle, (h) contour complet, (i) silhouette [12].	8
2.2	Description schématique du Filtre de Kalman [13].	10
2.3	Propagation des particules par l'algorithme SIR-figure tirée de [14].	12
2.4	Le module de suivi [14].	13
2.5	Exemple d'une association par NN-figure tirée de [14]	15
2.6	Exemples d'images de la séquence <i>Digiteto ComFuCet</i>	18
2.7	Exemples d'images de la séquence <i>SafeAround</i>	19
2.8	Dans l'exemple cité, $n_{tp}=4$, $n_{fp}=2$ et $n_{fn}=1$	20
3.1	Exemple d'un grossissement Θ de Ω associé à un raffinement ρ de Θ	40
3.2	Extension vide et marginalisation	42
4.1	Illustration des verrous scientifiques au niveau de la construction d'objets : (i) déplacement d'objets; (ii) fragmentation d'objets; (iii) détections non labellisées.	47
4.2	Associations 1-1, 1-N, M-N.	55
4.3	Algorithme de la construction dynamique d'objets.	69
4.4	Cas d'école (les objets sont représentés sans traits; les éléments focaux des bbas sont dessinés en traits gras) : (A,D,H) localisation spatiale des objets et des détections à (A) $t = 0$, (D) $t = 1$, (H) $t = 2$; (B-C,E-G) éléments focaux des bbas d'objets issues des différentes étapes de la méthode proposée.	72
4.5	Exemples de construction d'objets à 2 instants. A un instant donné t , les détections de même couleur sont associées à un même objet. En traits gras, sont représentées les détections à t et, en traits fins, les détections précédentes.	73
4.6	Exemples de construction d'objets à 2 instants. A un instant donné t , les détections de même couleur sont associées à un même objet. En traits gras, sont représentées les détections à t et, en traits fins, les détections précédentes.	74
4.7	Evaluation quantitative de la méthode proposée en fonction du retard à la décision : taux de rappel et précision calculés en termes de label (A) et fragments (B).	77
5.1	Exemples de résultats de détection à deux instants successifs t et $t + 1$ qui illustrent l'aspect fragmentaire des détections, les fausses détections ainsi que les non détections. Chaque objet physique est englobé par une boîte de couleur différente. Les fausses détections sont entourées en rouge. . . .	81

5.2	Comparaison des deux critères de décision.	98
5.3	Définition de la <i>bba</i> élémentaire : (a) évolution de la masse élémentaire β (notée x dans la figure) par rapport à la taille τ_a d'un objet certain ; (b) croissance monotone de la masse $m_{ij}(o_j)$ sur un objet o_j en fonction de sa taille $ A(o_j) $ (notée a_j dans la figure) et du paramètre τ_a	103
5.4	Exemples de l'évolution de la masse sur un objet par rapport au temps. Les paramètres utilisés sont : $\tau_a = 2500$, $\eta = 0.98$, $\tau_{it} = 25$, $\epsilon = 0.1$	103
5.5	Indicateurs de performance (S_e et PPV) estimées sur des données simulées en utilisant le <i>BetP</i> comme critère de décision ; lignes : (p_{nd}, p_{fa}) variable, $ \mathcal{T} $ fixe ; colonnes : (p_{nd}, p_{fa}) fixe, $ \mathcal{T} $ variable.	110
5.6	Indicateurs de performance (S_e et PPV) estimés sur des données simulées en utilisant le <i>BetP</i> comme critère de décision ; première lignes : η variable, $\epsilon = 0.1$; deuxième ligne : ϵ variable, $\eta = 0.98$, $(p_{nd}, p_{fa}) = (0.5, 0.25)$, $ \mathcal{T} = 6$	111
5.7	Données de <i>SafeAround</i> : évolution des masses sur les différents objets en fonction du temps illustrant des exemples de renforcement des croyances sur les objets persistants (cas des hypothèses A , B et C), affaiblissement des croyances sur les objets disparus (cas de l'hypothèse D et des <i>fausses alarmes</i>) et fusion des fragments des objets (cas des hypothèses B et D).	113
5.8	Données de <i>SafeAround</i> : visualisation des hypothèses de Θ projetées sur l'image à l'instant donné ; 1 ^{ère} colonne : composition en <i>RVB</i> des détections des 3 derniers instants ; 2 ^{ème} colonne : illustration des hypothèses de Θ (à t) en couleurs et des fausses alarmes en blanc.	114
6.1	Schéma global du suivi d'objets.	117
6.2	Résultat de détection issu de l'algorithme [15] et en actualisant l'image de fond, appliqué sur la séquence <i>Digiteo ComFuCet</i> (cf. Section 2.5.1) : le résultat de la détection au niveau pixel issu du premier critère NFA est représenté en bleu. Les détections fragmentaires (taille maximale égale à 10×20) issues du deuxième critère NFA sont en rouge.	118
6.3	Exemple de résultat de suivi d'objets sur des données simulées. Les détections sont celles issues de l'algorithme [15] et en actualisant l'image de fond. Les différents objets d'intérêt prennent différents labels.	120
6.4	Exemple de résultat de suivi d'objets sur la séquence <i>Digiteo ComFuCet</i> (cf. Section 2.5.1). Les détections sont issues de l'algorithme [15] et en actualisant l'image de fond. Les fausses alarmes sont présentées en noir. Les différents objets d'intérêt prennent différents labels.	121
A.1	Premier exemple illustrant l'algorithme Hongrois.	130
A.2	Deuxième exemple illustrant l'algorithme Hongrois.	131
A.3	Exemple illustrant l'algorithme Hongrois modifié.	133

Liste des tableaux

4.1	Probabilités pignistiques calculées pour chaque objet de $\mathcal{E}_1$	61
4.2	Probabilités pignistiques calculées pour chaque détection de $\mathcal{E}_2$	61
4.3	Matrice pignistique globale	61
5.1	bbas ‘partielles’ correspondant aux objets détectés par S_1 et S_2	96
5.2	Actualisation des bbas dans $\Omega_{1,2}$ et estimation de $\Theta_{1,2}$. Pour plus de lisibilité, les hypothèses disjonctives $\{o'_1, \bar{o}'_1\}$ et $\{o''_2, \bar{o}''_2\}$ ont été notées respectivement, $\mathcal{O}'_1$ et $\mathcal{O}''_2$, et $(\emptyset, \emptyset), (o'_1, \emptyset), (\bar{o}'_1, \emptyset)$ et $(\mathcal{O}'_1, \emptyset)$ n’étant pas des éléments focaux ni pour $m_1^{\Omega_{1,2}}$, ni pour $m_2^{\Omega_{1,2}}$ ni pour $m_{1,2}^{\Omega_{1,2}}$, ils ne figurent pas dans la table.	96
5.3	Actualisation de $m_{1,2}$ suite à la fusion de o'_1 et o''_2 et illustration des bbas ‘partielles’ correspondant aux objets détectés par S_3 . Les hypothèses disjonctives $\{o'_1, \bar{o}'_1\}$ et $\{o'''_2, \bar{o}'''_2\}$ ont été notées $\mathcal{O}'_1$ et $\mathcal{O}'''_2$, respectivement.	99
5.4	Actualisation des bbas dans $\Omega_{1,2,3}$ et estimation de $\Theta_{1,2,3}$. Les hypothèses disjonctives $\{o'_1, \bar{o}'_1\}$ et $\{o'''_2, \bar{o}'''_2\}$ ont été notées $\mathcal{O}'_1$ et $\mathcal{O}'''_2$, respectivement. $(\emptyset, \emptyset), (o'_1, \emptyset), (\bar{o}'_1, \emptyset)$ et $(\mathcal{O}'_1, \emptyset)$ n’étant pas des éléments focaux ni pour $m_{1,2}^{\Omega_{1,2,3}}$ ni pour $m_3^{\Omega_{1,2,3}}$ ni pour $m_{1,2,3}^{\Omega_{1,2,3}}$, ils ne figurent pas dans la table.	99

Abréviations

bba	b asic b elief a ssignment
ddp	d ensité d e p robabilité
EKF	E xtended K alman F ilter
FN	F aux N égatifs
FP	F aux P ositifs
GNN	G lobal N earest N eighbor
GSSF	G eneralized S imple S upport F unction
HOG	H istogram of O riented G radients
ISSF	I nverse S imple S upport F unction
JPDAF	J oint P robabilistic D ata A ssociation F ilter
KI	K nown as I mpossible
KP	K nown as P ossible
NFA	N umber of F alse A larms
MHT	M ultiple H ypothesis T racking
NN	N earest N eighbor
PDAF	P robabilistic D ata A ssociation F ilter
PMI	P rincape du M inimum d' I nformation
PPV	P ositive P redictive V alue
RSSF	a square R oot S imple S upport F unction
SIFT	S cale I nvariant F eature T ransform
SIR	S equential I mportance R esampling
SSF	S imple S upport F unction
SVM	S upport V ector M achines
TBM	T ransferable B elief M odel
TI	T raitement d' I mages

UH	U nknown H ypotheses
UKF	U nscented K alman F ilter
VP	V rais P ositifs
VT	V érité T errain

Chapitre 1

Introduction générale

1.1 Contexte applicatif

Dans notre société, les systèmes intelligents sont de plus en plus présents pour aider les hommes dans leurs travaux ou leurs loisirs. La surveillance de sites et la vidéosurveillance sont des exemples de systèmes d'assistance à un téléopérateur. Grâce à ces systèmes de surveillance, d'une part le nombre d'opérateurs effectuant cette tâche relativement ingrate peut diminuer et, d'autre part, cette tâche est effectuée dans des conditions plus confortables (à distance et avec l'aide d'un générateur d'alarmes). Typiquement, le but du système est de détecter les intrus (ou encore les événements) et remonter une alarme ou une menace. Seules les menaces pertinentes sont présentées à l'opérateur avec en outre un indicateur de 'sévérité' de ces menaces. Ceci permet de diminuer le nombre des fausses alarmes et donc d'éviter de submerger l'opérateur. Ce dernier peut alors se concentrer sur l'identification et la localisation des intrus.

Plus spécifiquement les systèmes de vidéosurveillance exploitent des données images afin d'en extraire des informations (l'image brute en est un exemple mais également les intrus détectés par un algorithme adéquat) qui servent de support à une décision. Afin à la fois d'être plus robustes et d'extraire des caractéristiques plus haut niveau comme la trajectoire, les algorithmes de traitement du flux vidéo se sont très tôt intéressés au suivi (*tracking* en anglais) des objets d'intérêt ('intrus' dans l'application citée). Maintenant, pour l'application qui nous intéresse, le nombre d'objets d'intérêt peut être supérieur à un, et les approches de suivi sont dites multi-objets. Les principaux défis sont alors liés

soit au nombre de ces objets et les ambiguïtés qui en découlent (croisement des objets, occultations etc.), soit même à la discrimination de ces objets et leur énumération à partir des détections en relation non univoque (bijective) avec les objets.

La capacité d'un algorithme de suivi multi-objets dépend alors de : (i) la représentation des objets à partir de laquelle des caractéristiques discriminatives (telles que la couleur, texture, flot optique, bord) sont extraites, qui elles-mêmes dépendent de la détection d'objets influant sur leur précision et fiabilité ; (ii) l'association des détections avec les différents objets (*data association* en anglais). La représentation mentionnée n'est pas restreinte à la description en termes de caractéristiques mais peut inclure l'imprécision et l'incertitude des connaissances. L'association est un problème bien connu dans des applications radar et/ou militaire. Elle a pour but de mettre en correspondance deux ensembles de détections, l'un correspondant à l'instant courant et l'autre issu des détections précédentes qui, associées entre elles, construisent les objets au cours du temps.

Les détections étant naturellement imparfaites (imprécises et incertaines), la robustesse de la surveillance provient d'une hypothèse implicite que les objets d'intérêt sont ceux persistants dans le temps et qu'un filtrage temporel permet de robustifier les alarmes remontées par le système de surveillance. Ainsi, l'information temporelle est utilisée pour distinguer les objets d'intérêt des fausses alarmes et déterminer de façon fiable l'ensemble des objets d'intérêt ou des alarmes. Notons cependant que la notion de fausse alarme elle-même est ambiguë et doit être définie par rapport à une application donnée. Par exemple, un chat entrant dans un site surveillé (e.g. chantier) est-il une alarme ? Du point de vue de la détection de changement, la réponse est probablement oui, mais du point de vue de la dangerosité la réponse est généralement non.

1.2 Fusion de données par fonctions de croyance

La fusion de données présente l'ensemble des méthodes et techniques permettant la combinaison de données issues de différentes sources afin d'avoir une information plus fiable améliorant la prise de décision. Les sources peuvent être des capteurs physiques ou logiques délivrant des mesures ou des observations imprécises et/ou incertaines. Elles sont souvent partiellement complémentaires et partiellement redondantes de sorte que leur fusion fournit une description plus précise, plus complète et de certitude suffisante pour

déclencher une action pertinente. Schématiquement, la redondance permet de confirmer une décision incertaine et la complémentarité permet de raffiner le champ des décisions. La fusion a été motivée notamment par l'accroissement des capteurs/sources et le besoin de synthétiser leurs informations. Son essor est apparu avec l'informatique et l'intelligence artificielle. La fusion est premièrement apparue dans le domaine militaire, plus particulièrement sonar/radar pour la détection et le suivi d'intrus, la surveillance de champs de bataille et la détection de mines. Depuis plusieurs années, des méthodes de fusion ont été adaptées et développées pour des applications civiles telles que la télédétection, l'imagerie médicale et la robotique.

Les cadres théoriques pour la fusion de données sont nombreux à savoir la théorie des probabilités, la théorie des possibilités, la théorie des fonctions de croyance, etc. Le choix d'une théorie parmi d'autres est lié à la nature des données exploitées. En vidéosurveillance en particulier, les données sont des flux vidéo et si l'on veut les combiner il s'agira de sorties d'algorithmes de traitement d'images. Ces dernières (comme les premières) sont naturellement imparfaites. D'une part, les détections sont incertaines dans le sens où il existe des non-détections (objets partiellement détectés, souvent fragmentés, ou totalement non détectés) et des fausses alarmes (sur-détections). D'autre part, la localisation d'un objet est d'autant plus imprécise qu'il est perçu par ses détections fragmentaires, c'est-à-dire 'n'importe où' dans l'objet. Parmi les théories permettant de gérer à la fois l'incertitude (à l'instar de la théorie des probabilités par exemple) et l'imprécision (à l'instar de la théorie des possibilités par exemple), la théorie des fonctions de croyance offre un formalisme solide incluant de nombreux opérateurs permettant de manipuler des informations homogènes et hétérogènes.

1.3 Plan de la thèse

Les deux chapitres suivants sont bibliographiques. Le premier (Chapitre 2) présente des généralités sur le *tracking*. Le deuxième (Chapitre 3) expose les notions de base de la théorie des fonctions de croyance en particulier le modèle des croyances transférables, cadre théorique de ce travail. Le chapitre 4 présente l'approche développée pour la construction dynamique d'objets d'intérêt qui inclut deux contributions à savoir la représentation et l'association des détections aux objets. Par la suite vient le chapitre 5 décrivant notre approche pour l'estimation de l'ensemble d'objets d'intérêt qui est notre troisième

contribution. Dans le chapitre 6, nous illustrons l'intérêt des deux chapitres 4 et 5 dans la résolution du problème de suivi d'objets en présentant quelques résultats. Finalement, nous concluons et nous proposons des perspectives pour ce travail.

Chapitre 2

Généralités sur le *tracking*

2.1 Introduction

Le suivi (*tracking*) est une application de la vision par ordinateur ayant connu un essor remarquable ces deux dernières décennies. Par *tracking*, nous entendons l'ensemble des algorithmes permettant la détection et le suivi d'objets dans une séquence vidéo. Il touche alors plusieurs domaines, tels que la vidéosurveillance, la biométrie ou encore l'imagerie biomédicale, l'interaction homme-machine, le contrôle de trafic et les véhicules intelligents, etc.

La multiplicité des problèmes à résoudre fait du *tracking* un défi pour les chercheurs qui proposent des systèmes permettant d'accroître soit la qualité du suivi soit la rapidité du calcul nécessaire. Soulignons en effet qu'un système de suivi doit répondre à des critères de robustesse aux variations de poses des objets (par conséquent aux changements d'apparence et aux déformations), robustesse aux occultations, aux variations d'illumination de la scène et au bruit présent dans les images, etc. En outre, il doit pouvoir gérer les éventuelles apparitions et disparitions des objets dans la scène.

Plus spécifiquement sur les aspects logiciels, un algorithme de suivi inclut généralement les étapes suivantes : (i) détection d'objets à partir des images ; (ii) représentation des objets qui soit adaptée à l'extraction de caractéristiques discriminantes ; (iii) maintien d'un lien temporel entre les instances des objets (c'est-à-dire la mise en correspondances des objets à chaque instant). Notons également que les processus (i) et (iii) s'effectuent séquentiellement ou simultanément. Dans le premier cas, les objets sont d'abord détectés (cf. Section 2.2) et ensuite suivis. Dans le deuxième, les objets sont déjà présents

dans l'image précédente et toute information sur leurs positions précédentes ou d'autres caractéristiques telles que la forme peut servir à les détecter dans l'image courante.

2.2 Détection d'objets

L'étape de détection d'objets est primordiale pour initier tout processus de suivi. Plusieurs approches ont été proposées. A titre d'exemple, [12] énumère quatre type d'approches : modélisation du fond, segmentation, classifieurs supervisés et détecteurs de points-clés.

Selon le premier type d'approche, l'image observée peut être interprétée comme la somme de deux images appelées respectivement *foreground* et *background*. L'image de *background* est généralement une image de la séquence ne contenant pas d'objets d'intérêt (scène vide) et de luminosité représentative de la scène ultérieurement. Cependant, en pratique, ces deux contraintes sont difficiles à atteindre. Les scènes réelles sont également sujettes à des changements au niveau des objets statiques (agitation des branches des arbres, vagues de la mer, fermeture et ouverture d'une porte, etc.) ou encore des variations des conditions d'acquisition (luminosité, nuages, etc.). Ainsi, pour ne pas percevoir ces modifications en tant que des changements, des approches telles que le filtre sigma-delta (cf. Section A.3.1.1) et le modèle codebook (cf. Section A.3.1.2) assimilent progressivement ces modifications dans l'image de *background* et génèrent un modèle de *background* à chaque instant.

Le deuxième type d'approche permettant la détection des objets est la segmentation de l'image. Les pixels sont alors regroupés en régions, constituant une partition de l'image, l'hypothèse étant alors que les objets à détecter correspondent à certaines de ces régions. Parmi les techniques de segmentation proposées récemment, nous citons le Mean-Shift [16], les Graph-Cuts [17] ainsi que les contours actifs [18].

Dans le cadre de la classification supervisée, les objets (classes) que nous souhaitons détecter sont représentés par des caractéristiques définies de façon supervisée, soit à partir de connaissances et modèles a priori sur les objets, soit par apprentissage. Ces caractéristiques (par exemple, texture, histogramme de couleur etc.) doivent être suffisamment discriminantes pour discerner un objet (une classe) d'un autre. A partir de ces caractéristiques, des méthodes de discrimination telles que les réseaux de neurones [19], les arbres de décision, les SVM (*Support Vector Machines*) [20] permettent d'identifier les

classes (les objets).

La dernière approche se distingue des précédentes au sens où elle ne considère pas l'image comme un champ dense, mais se focalise sur les pixels correspondant aux points-clés, aussi appelés points d'intérêt. Ces points de l'image présentent des propriétés remarquables pour la reconnaissance et le suivi notamment grâce à des propriétés telles que l'invariance au changement d'illumination et la prise de vue. Le calcul de ces points implique généralement deux étapes dont une détection et une caractérisation. Par exemples, les descripteurs SIFT (*Scale Invariant Feature Transform*) [21] sont parmi les plus utilisés.

2.3 Représentation d'objets

La représentation (de ou) des objets à suivre affecte l'efficacité d'un algorithme de *tracking*. Celle-ci définit en effet les caractéristiques discriminantes extraites pour discerner un objet du fond ou des autres objets. Le choix d'une représentation est fortement lié à l'application traitée : si les objets à suivre occupent des petites régions dans l'image, le plus courant est de les assimiler à un ensemble de points ou simplement par leur centroïde. Des objets simples et rigides sont couramment représentés par des primitives de formes géométriques telles qu'un rectangle ou une ellipse [22]. Le contour (bord de l'objet) [23] ou la silhouette (région à l'intérieur du contour) sont généralement utilisés pour représenter des objets de forme non rigide et complexe. Finalement, il est possible de modéliser un objet articulé (ensemble de sous parties rassemblées par des jonctions) par un ensemble de formes géométriques (cylindre ou ellipse pour chacune des sous parties) ou un squelette [24].

Hormis les représentations de forme (cf. Figure 2.1), des représentations d'apparence ont été proposées pour décrire les objets. Nous nous limitons à citer les plus utilisées. L'histogramme [22] représente la distribution de probabilités de caractéristiques (couleurs, textures, gradient, forme, etc.) calculées sur une ou plusieurs régions de l'objet. Si les caractéristiques sont des valeurs de couleurs brutes ou des valeurs d'intensité, des traitements tels que l'ajustement de la moyenne sont souvent utilisés pour pallier les problèmes de changement de l'illumination et d'ombre [25]. Notons que ces histogrammes peuvent concerner le gradient (comme dans le descripteur SIFT cité précédemment où le HOG (*Histogram of Oriented Gradients*)). Citons également la représentation par matrice de cooccurrence [26] (associée à une région de l'objet) qui a l'avantage de pouvoir fusionner

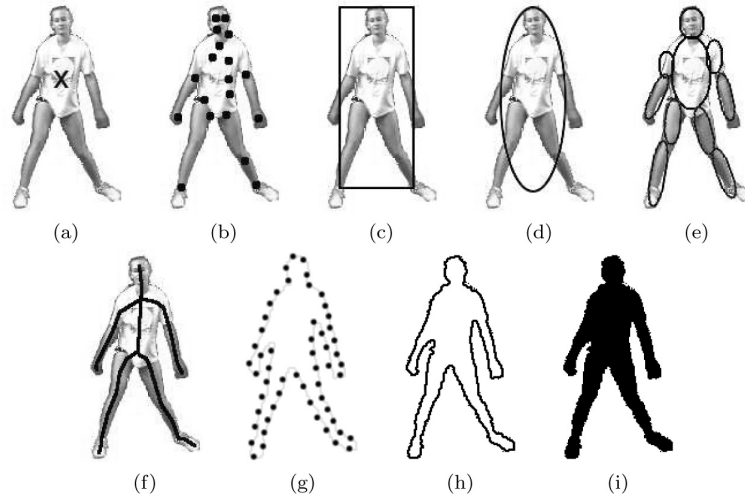


FIGURE 2.1: Représentations d'un objet : (a) centroïde , (b) points multiples, (c) rectangle, (d) ellipse, (e) forme articulée, (f) squelette, (g) contour sous forme de points de contrôle, (h) contour complet, (i) silhouette [12].

des caractéristiques multiples de la région ce qui permet de la mettre en correspondance à différentes vues et poses. Enfin, les modèles d'apparence actifs [27] modélisent simultanément la forme de l'objet et son apparence mais nécessitent une étape d'apprentissage, et les modèles d'apparence multi-vue [28, 29] permettent le codage de différentes vues d'un objet en se basant sur des approches telles que l'analyse en composantes principales et l'analyse en composantes indépendantes. Leur inconvénient est la nécessité d'avoir simultanément plusieurs vues d'un même objet.

2.4 Techniques de suivi d'objet(s)

Le *tracking* peut (selon l'application) se faire pour un seul objet (suivi mono-objet) ou le plus souvent pour plusieurs objets simultanément présents dans la scène (suivi multi-objets). Pour un algorithme donné de suivi mono-objet, la question de son extension au suivi multi-objets n'est pas toujours triviale. Cette extension est généralement possible si les objets ne se croisent pas. En effet, il est possible d'appliquer l'algorithme de suivi mono-objet pour suivre chacun des objets. Cependant, si les objets interagissent, ce qui est souvent le cas, les techniques de suivi mono-objet échouent généralement à distinguer les différents objets. Une autre complexité par rapport au suivi mono-objet est due à la variabilité du nombre d'objets dans la scène (apparitions et disparitions des objets dans la scène).

Comme dit à la section 2.1, l'objectif du suivi est d'estimer la trajectoire de chaque objet présent dans la scène (durée de la séquence vidéo) en le localisant dans chaque image de la séquence. Dans cette section, nous exposons certaines techniques classiques de suivi à savoir le filtre de Kalman, le filtre particulaire SIR (*Sequential Importance Resampling*), le filtre JPDAF (*Joint Probability Data Association Filter*) et le MHT (*Multiple Hypothesis Tracking*).

2.4.1 Filtre de Kalman

Le filtre de Kalman [30] est un algorithme de filtrage linéaire qui estime de manière optimale l'état d'un système linéaire dynamique étant donné un ensemble de mesures (ou observations), lorsque le bruit d'observation est additif gaussien. Il a été intensivement utilisé pour le suivi mono-objet où il permet d'estimer le vecteur d'état x_t décrivant généralement, l'état cinématique de l'objet à suivre tel que sa position et sa vitesse ou d'autres attributs comme ses dimensions, son type ou sa classe, connaissant le vecteur de mesures bruitées (ou encore vecteur d'observation) d_t . Il est alors possible d'estimer la ddp (densité de probabilité) $p(x_t/d_t)$ supposée gaussienne d'espérance $\hat{x}_t = E[x_t/d_t]$ et de variance $B_t = E[(x_t - \hat{x}_t)(x_t - \hat{x}_t)^T]$. Ces deux premiers moments sont calculés récursivement en passant par les deux étapes suivantes :

- une étape de prédiction qui prédit l'état actuel du système à partir d'un état initial, mais avec une grande incertitude.
- une étape de correction qui fournit un état plus précis à partir du résidu (écart entre l'état prédit et la mesure actuelle).

Soient les modèles donnés par les équations (2.1). La première équation est le modèle d'état. La deuxième présente le modèle de mesure décrivant la relation entre la mesure et l'état.

$$\begin{cases} x_t = M_t x_{t-1} + w_t, \\ d_t = H_t x_t + \epsilon_t, \end{cases} \quad (2.1)$$

où :

- M_t et H_t sont deux matrices déterministes, appelées respectivement matrice d'état et matrice de mesure.
- w_t est le bruit du modèle supposé blanc et gaussien et de variance Q_t .
- ϵ_t est le bruit des mesures supposé aussi blanc et gaussien et de variance R_t . Les deux bruits sont mutuellement indépendants.

L'étape de prédiction utilise le modèle d'état ainsi que la ddp de l'état à l'instant $t - 1$ pour déduire la ddp a priori de l'état à l'instant actuel t .

$$\begin{cases} \hat{x}_{t/t-1} = M_t \hat{x}_{t-1}, \\ B_{t/t-1} = Q_t + M_t B_t M_t^T. \end{cases} \quad (2.2)$$

L'étape de correction fait intervenir la fonction de *vraisemblance* de la mesure actuelle pour déduire la ddp *a posteriori*. Les mesures à t sont utilisées pour corriger l'état prédit afin d'obtenir une estimation plus précise.

$$\begin{cases} \hat{x}_t = \hat{x}_{t/t-1} + K_t(d_t - H_t \hat{x}_{t/t-1}), \\ B_t = B_{t/t-1} - K_t H_t B_{t/t-1}, \end{cases} \quad (2.3)$$

où $K_t = B_{t/t-1} H_t^T (R_t + H_t B_{t/t-1} H_t^T)^{-1}$ est le gain du filtre de Kalman. En le multipliant par le résidu $d_t - H_t \hat{x}_{t/t-1}$, il permet de corriger l'estimée prédite.

Schématiquement, le comportement du filtre de Kalman est illustré dans la figure 2.2. Dans cette dernière, le point bleu désigne l'état initial entouré par un cercle d'incertitude (cercle bleu). A chaque instant, nous représentons les états prédits par des points rouges entourés par des cercles d'incertitude (cercles rouges), tandis que les points verts dénotent les états mis à jour qui ont des précisions plus fines (cercles verts). Dans le cas où le

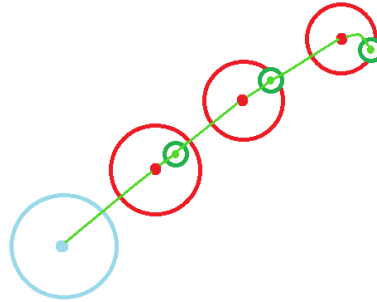


FIGURE 2.2: Description schématique du Filtre de Kalman [13].

si le système est non linéaire mais avec des bruits toujours gaussiens, le filtre de Kalman ne fournit pas la solution optimale, d'où l'apparition de ses extensions telles que l'EKF (*Extended Kalman Filter*) [31] et l'UKF (*Unscented Kalman Filter*) [32].

Le filtre de Kalman a été intensivement utilisé en vision notamment en suivi tel que le suivi de points dans des images bruitées [33] et pour l'estimation des trajectoires 3D à partir du mouvement 2D [34].

2.4.2 Filtre particulaire SIR

En vision, la plupart du temps, le système dynamique est décrit par un modèle fortement non linéaire et/ou non gaussien. Dans ce cas, le filtre de Kalman et ses extensions ne parviennent pas à approximer précisément l'état courant de l'objet suivi. Des approximations numériques, telles que celles données par les méthodes séquentielles Monte-Carlo connues aussi sous le nom de filtres particulaires ont été étudiées comme des alternatives. Le filtre particulaire SIR [35–38] cherche à approximer la densité a posteriori $p(x_t/d_{1..t})$ par un échantillon pondéré de taille k , $\{x_t^{(i)}, \pi_t^{(i)}\}_{i=1..k}$ (cf. Equation (2.4)). Chaque $x_t^{(i)}$ est une solution possible de l'état à estimer, appelée particule, à laquelle est associé un poids $\pi_t^{(i)} \in [0, 1]$ traduisant sa qualité par rapport aux mesures.

$$p(x_t/d_{1..t}) \approx \sum_{i=1}^k \pi_t^{(i)} \delta_{x_t^{(i)}}(x_t). \quad (2.4)$$

Dans l'équation (2.4), $d_{1..t}$ est l'ensemble des mesures jusqu'à l'instant t , δ est la fonction de Kronecker. L'approximation récursive de la ddp a posteriori passe par trois étapes essentielles : (i) exploration de l'espace d'état permettant de propager les particules. C'est une étape de prédiction de l'échantillon $\{x_t^{(i)}, \pi_t^{(i)}\}_{i=1..k}$ à l'instant t à partir de l'échantillon précédent $\{x_{t-1}^{(i)}, \pi_{t-1}^{(i)}\}_{i=1..k}$ via un échantillonnage utilisant une fonction d'importance (ou proposition), $x_t^{(i)} = f(x_t/x_{0..t-1}^{(i)}, d_{1..t})$; (ii) correction (ou évaluation) de la qualité des particules, à l'aide des observations, en calculant leurs nouveaux poids selon l'équation (2.5) avec $\sum_{i=1}^k \pi_t^{(i)} = 1$ (une étape de normalisation des poids est nécessaire) ;

$$\pi_t^{(i)} = \pi_{t-1}^{(i)} \frac{p(d_t/x_t^{(i)})p(x_t^{(i)}/x_{t-1}^{(i)})}{f(x_t/x_{0..t-1}^{(i)}, d_{1..t})}. \quad (2.5)$$

(iii) sélection facultative des particules pour éliminer les particules de faibles poids et accentuer les particules ayant des poids importants. Ce rééchantillonnage est généralement utilisé pour éviter la dégénérescence de l'échantillon. Finalement, le nuage de particules obtenu à l'instant t approxime la distribution a posteriori, $p(x_t/d_{1..t}) \approx \left\{x_t^{(i)}, \pi_t^{(i)}\right\}_{i=1}^k$. Une illustration de la propagation des particules par l'algorithme SIR est présentée dans la figure 2.3. Le filtre particulaire SIR a été intensivement utilisé dans plusieurs applications telles que le traitement de signal [39], le traitement de la parole [40], la robotique mobile [41] ainsi que le suivi visuel [42]. En suivi, il permet de représenter des densités arbitraires, de se focaliser sur des régions spécifiques de l'espace d'état et de gérer les

modèles multiples. Il est robuste aux problèmes de fouillis et d'occultations.

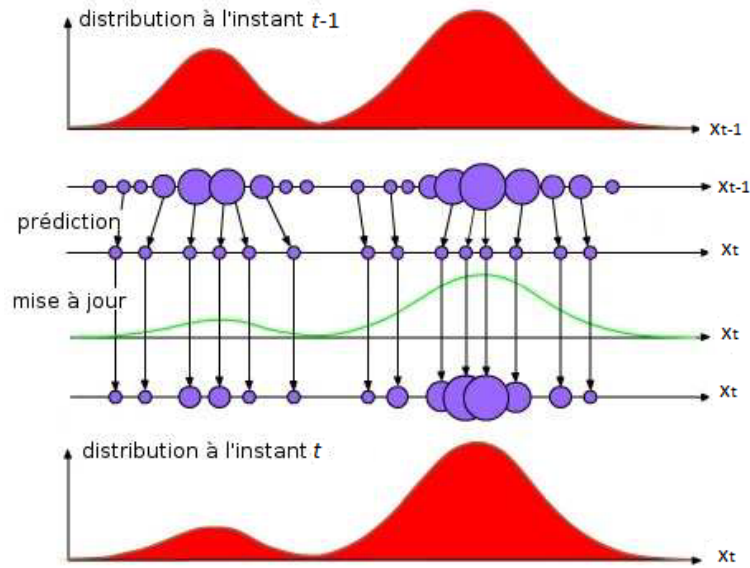


FIGURE 2.3: Propagation des particules par l'algorithme SIR-figure tirée de [14].

Le filtre de Kalman et le filtre particulaire SIR sont conçus pour suivre un seul objet en estimant son état caché à partir des mesures acquises au cours du temps. Et dans ce cas, plusieurs mesures sont souvent disponibles à chaque itération. Une mesure provient soit du seul objet suivi soit d'un dysfonctionnement des capteurs (une fausse alarme). Cependant, la prise en compte des fausses alarmes dans la mise à jour de la loi de filtrage peut entraîner la perte de l'objet suivi. En suivi multi-objets, le problème est beaucoup plus critique. En effet, outre le filtrage des fausses alarmes, il faut savoir à quel objet appartient une mesure ce qui demande une étape de mise en correspondance entre les mesures et les objets (cf. Figure 2.4). Dans ce cas, le plus simple est d'utiliser la méthode du plus proche voisin NN (*Nearest Neighbor*). Cependant, si les objets sont très proches les uns des autres, il est généralement possible qu'au moins l'une des correspondances soit erronée ce qui entraîne la divergence des filtres. Comme alternatives, nous pouvons utiliser des techniques statistiques qui traitent l'association de données (cf. Section 2.4.4) conjointement aux problèmes d'estimation d'état. Parmi ces dernières, les plus connues sont le JPDAF (cf. Section 2.4.4.2) et le MHT (cf. Section 2.4.4.3).

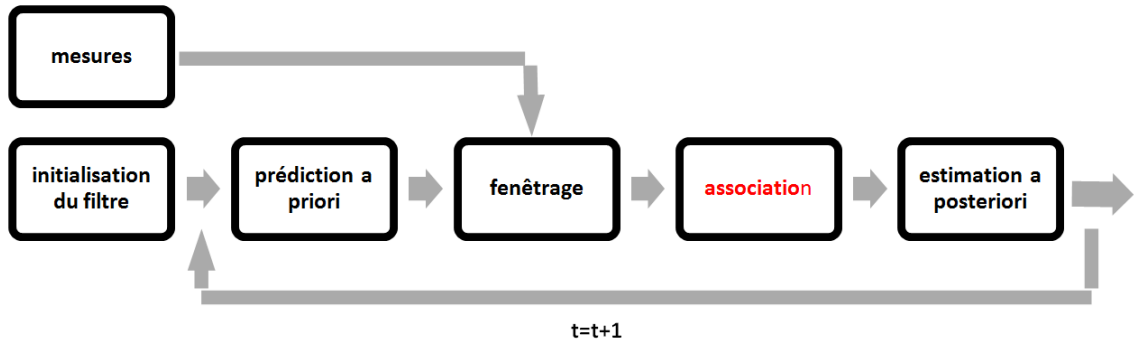


FIGURE 2.4: Le module de suivi [14].

2.4.3 Fenêtrage

Dans la figure 2.4, le module fenêtrage [14, 43] sert pour valider les mesures acquises à chaque itération diminuant ainsi la complexité des calculs et évitant des associations non probables et aberrantes, entre les mesures et les objets. Il se base sur un test de proximité et la définition d'un seuil délimitant un espace d'observation à l'extérieur duquel un objet suivi a une très faible probabilité de se trouver. Ainsi, seules les mesures à l'intérieur du volume sont gardées (validées). Pour chaque objet, étant donné le vecteur d'état prédit (calculé par filtre de Kalman ou filtre particulaire), une distance statistique comme la distance de Mahalanobis est calculée entre la mesure prédite et la mesure courante disponible de sorte que les mesures statistiquement éloignées de la mesure prédite sont éliminées (ne sont pas prises en compte dans l'association).

2.4.4 Association

Dans le cas du suivi, nous nous intéressons à un objet au cours du temps. Nous parlons alors de piste pour faire référence à la notion d'indexage temporel. Lorsque plusieurs objets (pistes) sont simultanément présents dans la scène, l'association entre les détections (observations ou mesures) et les objets (pistes) permet de déterminer à quel objet (piste) se rattache une nouvelle détection (observation ou mesure) avant de l'actualiser :

– association type 1-1 :

s'il s'agit d'un suivi mono-capteur, une détection validée est soit une fausse alarme, soit elle provient d'au plus un objet. De même, un objet génère au plus une détection. Ainsi, le problème d'association est tel qu'au maximum une détection validée est associée au plus à un objet..

- association type 1-N :

ce type d'association apparaît lorsqu'à un objet peut correspondre plusieurs détections, le cas par exemple du suivi multi-capteurs.

Par ailleurs, une association est soit locale soit globale.

- association locale :

elle consiste à définir une relation entre un objet o_i et une détection d_j . Soient n objets o_i et l détections d_j . L'hypothèse d'association locale s'exprime comme 'la détection d_j est associée avec l'objet o_i , $\forall i \in \{1...n\}$, $\forall j \in \{1...l\}$ ' ou encore 'la détection d_j n'est pas associée'. Notons que le NN [44] et le PDAF (*Probabilistic Data Association Filter*) [45] sont deux méthodes d'association locale.

- association globale :

c'est un ensemble d'associations locales compatibles entre elles et représentatives de la situation actuelle. Si nous disposons par exemple de 2 objets $\{o_1, o_2\}$ et 3 détections $\{d_1, d_2, d_3\}$, une hypothèse d'association globale peut être ' d_1 est associée avec o_2 , d_2 est associée avec o_1 , d_3 est non associée'.

Les approches d'association globale les plus utilisées sont GNN (*Global Nearest Neighbor*) [46], JPDAF [47] et MHT [48].

2.4.4.1 NN et GNN

La méthode NN (cf. Figure 2.5) est probablement celle qui est la plus simple pour l'association. Pour chaque objet, les distances séparant son état prédit des mesures validées c'est-à-dire appartenant à sa fenêtre de validation, sont calculées. Seule la mesure minimisant la distance est associée à l'objet. Cependant, la nécessité d'un traitement séquentiel rend le résultat d'association dépendant de l'ordre dans lequel les objets sont traités. Cette méthode est en plus, sensible au seuil à partir duquel nous décidons d'accepter ou rejeter une association. Par ailleurs, la NN demande que le nombre d'objets soit connu ainsi qu'un faible taux de fausses alarmes. Son avantage majeur est son faible coût calculatoire. Comme alternative, la méthode GNN, plus utilisée que la NN, a l'avantage de résoudre le problème d'association de manière globale. Ainsi, une matrice de coût est établie dont les valeurs sont les coûts locaux d'association pouvant être une distance séparant la mesure prédite d'un objet et une mesure validée. Le problème d'association est résolu en minimisant le coût global (somme des coûts locaux) faisant recours à des algorithmes d'affectation tels que l'algorithme Hongrois [49].

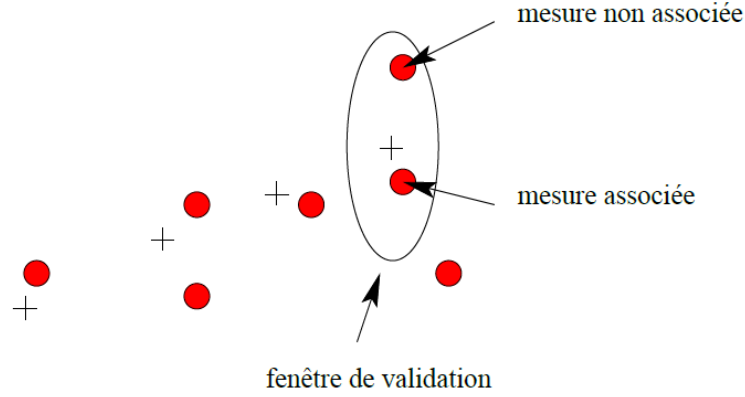


FIGURE 2.5: Exemple d'une association par NN-figure tirée de [14]

2.4.4.2 PDAF et JPDAF

Dans des cas pratiques, plusieurs observations (ou mesures) peuvent provenir d'un même objet comme par exemple la détection d'un piéton par un télémètre laser induisant la détection disjointe de ses jambes. Ainsi, les jambes sont deux observations provenant du même objet qui est le piéton. Sachant qu'une piste est une séquence d'observations (mesures) issues d'un même objet, le problème traite de l'association multiple de type piste/multi-observations. En effet, plusieurs observations (ou mesures) peuvent s'associer à la même piste et contribuer à son actualisation. Ce type d'association non résolu par des méthodes d'association bijectives telles que le NN et le GNN, peut être traité par des approches probabilistes telles que le PDAF et le JPDAF [50] qui sont des filtres de mise à jour des pistes considérant plusieurs observations.

Soient à un instant donné t , n pistes $\{o_1, \dots, o_n\}$ et l mesures $\{d_1, \dots, d_l\}$. Le PDAF est une méthode bayésienne sous-optimale. Elle filtre un seul objet évoluant dans un environnement bruité à partir de ses observations voisines en construisant un estimateur qui tient compte des mesures dans la fenêtre de validation associée à la piste, c'est-à-dire toute l'information disponible dans celle-ci.

Soit $o_{i,i \leq n}$ une piste. Pour chaque mesure $d_{j,j \leq l}$, nous calculons la probabilité p_{ij} de l'hypothèse qui associe la mesure d_j à la piste o_i . Soit $v_{ij}(t)$ le résidu (cf. Section 2.4.1) lié à la piste o_i à l'instant t suite à la mesure d_j . L'actualisation d'une piste selon les associations possibles apparaît dans l'étape d'estimation du filtre de Kalman (cf. Section 2.4.1). En effet, le résidu $v_i(t)$ de la piste o_i est égal à la somme des résidus pondérés $v_{ij}(t)$ de

toutes les l associations possibles (cf. Equation (2.6)).

$$v_i(t) = \sum_{j=1}^l p_{ij} \cdot v_{ij}(t). \quad (2.6)$$

Notons finalement que le PDAF est assez performant lorsque le nombre de pistes est connu. Cependant, ne pas considérer les autres pistes peut causer des erreurs dans l'association d'où la perte de pistes. Une alternative est d'utiliser le JPDAF comme extension du PDAF. Le JPDAF permet de résoudre les conflits issus du chevauchement de deux fenêtres de validation (une fenêtre par piste). En effet, il diffère du PDAF dans le calcul des probabilités p_{ij} qui considère simultanément toutes les pistes. S'il n'y a pas de zones critiques de chevauchement entre les fenêtres de validation des pistes, le JPDAF revient à appliquer indépendamment le PDAF par piste.

Le JPDAF a été utilisé pour la reconstruction 3D [51] ainsi que le suivi de régions [52]. Cependant, sa limitation majeure est son incapacité de gérer la variabilité du nombre de pistes. Dans ce qui suit, nous présentons brièvement l'algorithme MHT qui veut surmonter cette limite.

2.4.4.3 MHT

Le MHT [53] est un algorithme de suivi multi-objets de type 1-1 qui fait face aux problèmes d'initialisation et d'ambiguïté dans les associations entre les observations (mesures) et les pistes. A chaque instant, étant donné un nouvel ensemble d'observations, soit : (i) l'observation (mesure) est issue d'une piste déjà existante ; (ii) l'observation (mesure) est issue d'une nouvelle piste ; (iii) l'observation (mesure) est une fausse alarme. Le MHT suppose que le nombre de pistes est inconnu et peut varier au cours du temps. Etant données toutes les hypothèses d'association possibles, le MHT crée une nouvelle piste pour chacune de ces hypothèses. L'idée de base du MHT est de retarder la prise de décision concernant les hypothèses à garder et celles à éliminer et d'attendre que de prochaines observations lèvent l'ambiguïté. Bien que cet algorithme soit coûteux en termes de temps de calcul et mémoire, il est significativement performant car il est capable de gérer la création, la destruction et la fusion de pistes. Cependant, le nombre d'hypothèses suivies croît exponentiellement avec les nouvelles détections. D'où la nécessité d'éliminer les hypothèses de trop faible probabilité afin de diminuer la complexité de l'algorithme. Cet algorithme a été utilisé pour le suivi d'objets de type humain ainsi que le suivi de

points d'intérêt. Dans ce dernier, l'algorithme Murty [54] a permis d'estimer les meilleures hypothèses en temps polynomial.

En conclusion, pour une association type 1-N (piste - observations), les approches PDAF, JPDAF sont les plus adaptées. S'il s'agit d'une association type 1-1 (piste-observation), il faut opter vers les approches GNN et MHT. Par ailleurs, les approches PDAF, JPDAF, GNN retiennent une seule et meilleure hypothèse à chaque pas de temps alors que le MHT étudie et propage plusieurs hypothèses d'association au cours du temps.

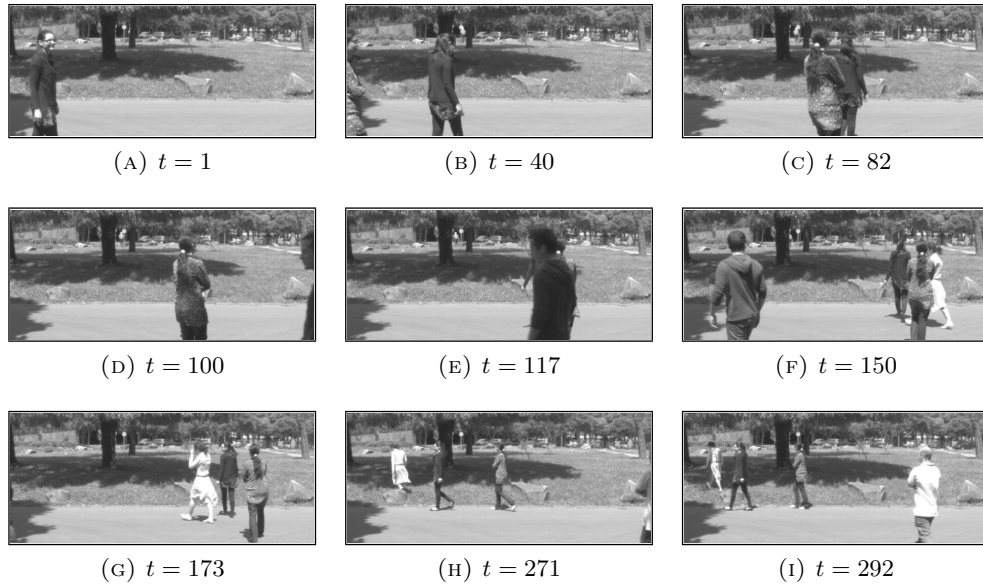
2.5 Présentation des données

Nous exposons à présent les séquences vidéo utilisées dans le reste du manuscrit, pour illustrer nos trois contributions majeures à savoir la représentation, l'association (cf. Chapitre 4) et le filtrage d'objets (cf. Chapitre 5).

2.5.1 Séquence *Digiteo ComFuCet*

La première séquence de données réelles a été acquise dans le cadre du projet *Digiteo ComFuCet* à la fréquence de 25 images/seconde. Ce projet avait comme application l'assistance à la conduite automobile dans le cas de véhicules dotés d'un système de vision par ordinateur. Nous nous sommes intéressés à un extrait d'une quinzaine de secondes sur lequel la vérité terrain a été définie manuellement. Cet extrait a été choisi car significatif par rapport à l'ensemble de la séquence. Dans l'extrait, jusqu'à six 'objets' sont simultanément présents. Ils correspondent à six personnes et une voiture qui traversent ou évoluent dans la scène en se croisant.

Certaines images de la séquence sont présentées dans la figure 2.6. Elles illustrent plusieurs difficultés au niveau de la séquence : (i) celle-ci est acquise dans un environnement extérieur sujet à des variations rapides de luminosité. Ces problèmes de changement de l'illumination et d'ombre influent considérablement sur les caractéristiques des objets telles que l'histogramme de couleur (ou de texture). Nous prenons comme exemple la personne qui passe de l'ombre (cf. Figure 2.6e) au soleil (cf. Figure 2.6f) ; (ii) les objets sont sujets au cours du temps, à des variations en termes de couleur et de texture notamment en raison de modifications de leur orientation (personne de profil puis de dos

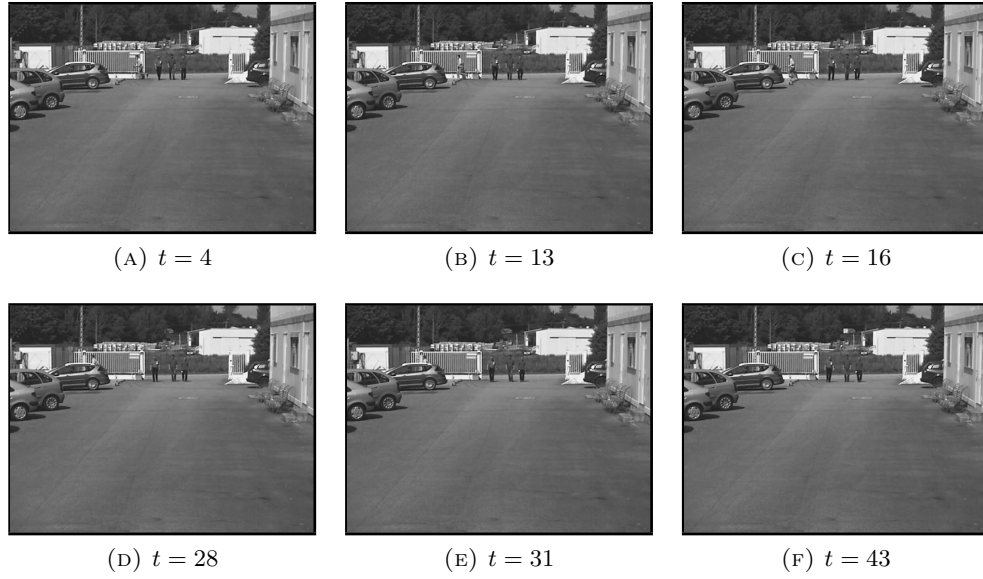
FIGURE 2.6: Exemples d'images de la séquence *Digiteo ComFuCet*.

par exemple); (iii) les objets se croisent à plusieurs reprises impliquant des occultations d'un objet par un autre.

2.5.2 Séquence *SafeAround*

La deuxième séquence de données réelles a été acquise à la fréquence de 1 image/seconde dans le cadre du projet *SafeAround*. Ce projet avait pour but la définition d'un système de protection de périmètre [55] qui doit détecter les intrus (typiquement les véhicules, les individus et les groupes). La figure 2.7 présente quelques images de la séquence où : (A) un groupe de quatre personnes entrent dans la scène; (B) un parmi le groupe se déplace derrière une voiture stationnée; (C) des voitures circulent au fond de la scène; (D) trois objets persistent dans la scène à savoir les trois personnes (appelées par la suite *A*, *B*, *C*) passant en avant; (E) un objet présent au début de la scène et disparaît après, à savoir la personne (appelée *D*) se déplaçant derrière la voiture stationnée; (F) deux objets absents au début de la séquence, apparaissent et disparaissent successivement à savoir le camion sur la route et une voiture (appelés respectivement *E* et *F*).

Comparée à la séquence *Digiteo ComFuCet*, cette séquence est moins complexe en termes de déplacements et croisements d'objets ainsi que de variation de l'illumination (bien qu'également acquise en extérieur).

FIGURE 2.7: Exemples d'images de la séquence *SafeAround*.

2.6 Métriques d'évaluation

Nos contributions sont évaluées quantitativement en termes de sensibilité S_e et PPV (*Positive Predictive Value*). Ces métriques de performance sont définies à partir du nombre de VP (vrais positifs), nombre de FP (faux positifs) et nombre de FN (faux négatifs) notés respectivement n_{tp} , n_{fp} et n_{fn} (cf. Equation (2.7)). Ces derniers dont la définition dépend du problème traité (e.g. association, filtrage), sont calculés à partir d'une VT (vérité terrain) et un résultat obtenu.

$$S_e = \frac{n_{tp}}{n_{tp} + n_{fn}},$$

$$PPV = \frac{n_{tp}}{n_{tp} + n_{fp}}. \quad (2.7)$$

2.6.1 Définitions des VP, FP et FN pour le sous-problème d'association

Nous illustrons à travers la figure 2.8 qui montre la VT (cf. Figure 2.8a) correspondant à un résultat d'association un peu complexe (cf. Figure 2.8c), les définitions des VP, FP et FN. La VT et le résultat sont tels que les détections appartenant à un même objet se voient attribuer un même label (couleur). En utilisant une représentation sous forme d'un graphe biparti (cf. Figure 2.8b), nous définissons les termes VP, FP et FN

à partir des relations (illustrées sous formes d'arêtes du graphe) entre les labels (un sommet du graphe représente un label attribué à un objet) figurant dans la VT et ceux dans le résultat d'association. Lorsque deux sommets sont liés par une seule arête, nous distinguons un VP tandis que lorsque plusieurs arêtes sont présentes, une est comptée un VP, les autres sont soit des FP (lorsque deux labels ou plus dans le résultat d'association correspondent à un seul label dans la VT) soit des FN (lorsqu'un label dans le résultat d'association correspond à deux labels ou plus dans la VT).

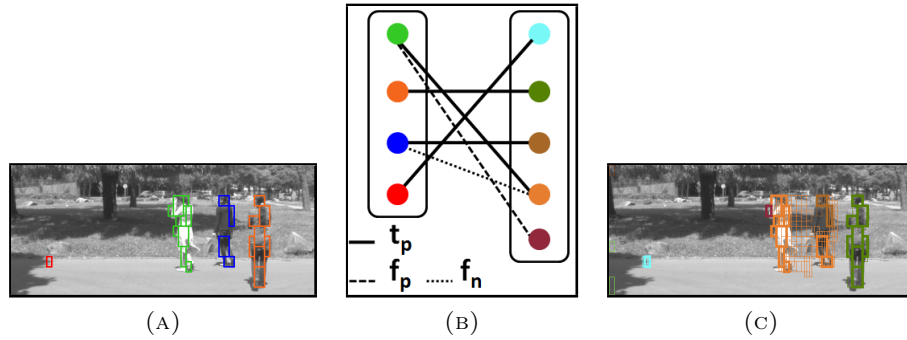


FIGURE 2.8: Dans l'exemple cité, $n_{tp}=4$, $n_{fp}=2$ et $n_{fn}=1$.

2.6.2 Définition des VP, FP et FN pour le sous-problème de filtrage

Le filtrage nous fournit à chaque instant t un ensemble d'objets d'intérêt noté Θ . Etant donnée une VT qui représente l'ensemble d'objets réellement présents dans la scène imagée au même instant t , les VP, FP et FN sont définis à partir de VT et Θ comme suit : (i) un VP est un élément de Θ qui correspond à un objet de VT ; (ii) un FP est un élément de Θ qui ne correspond pas à un objet de VT mais plutôt à une fausse alarme ; (iii) un FN est un objet de VT qui ne correspond à aucun élément de Θ .

Finalement, notons que l'évaluation de la représentation d'objets n'est pas faite directement mais plutôt au travers des deux autres contributions 'association et filtrage'.

2.7 Conclusion

Ce chapitre avait pour objet d'exposer des généralités sur le *tracking*, contexte général de ce travail, allant des méthodes de détection, de représentation jusqu'aux techniques classiques de suivi d'objets telles que le filtre de Kalman et le filtre particulaire, qui

elles-mêmes nécessitent une étape d'association entre les objets (pistes) et les détections (mesures), pour laquelle divers techniques existent en littérature telles que le GNN, le JPDAF et le MHT.

Pour conclure sur ce chapitre bibliographique, nous distinguons deux catégories de *tracking* : (i) le *tracking* radar ; (ii) le *tracking* image ou encore vidéo. Plusieurs techniques de suivi (e.g. filtre de Kalman, MHT) ont été appliquées d'abord en radar puis adaptées pour le *tracking* image [56]. En *tracking* image, la taille des objets dans la vidéo varie considérablement en fonction du temps et avec le déplacement des objets. Par ailleurs, la détection des objets est généralement imparfaite induisant leur fragmentation, un problème qui ne se pose pas en *tracking* radar. En radar, les observations sont le plus souvent assimilées à des points qui représentent généralement les centres des objets. De nos jours et avec le progrès scientifique dans le domaine radar, une résolution fine du capteur devient possible et permet de caractériser un objet (e.g. voiture) non seulement par le centre mais aussi par la forme. Ainsi, il est possible de représenter un objet (cible) par une forme géométrique (e.g. ellipse). Les objets (e.g. avion, bateau) sont supposés aussi rigides ou semi-rigides. En effet, en radar, la taille des objets ne varie pas entre deux images successives sauf si l'objet disparaît brusquement (e.g. sortir du champ de vision). Dans [57], l'EKF et l'UKF ont été utilisés pour le *tracking* radar et les matrices d'observations concernent la position, la vitesse et la longueur des axes majeurs de l'objet suivi. Dans [58], un objet est modélisé par une distribution de probabilité spatiale (loi de poisson) et le filtre de Kalman permet de prédire la position. En *tracking* image, l'hypothèse de la rigidité d'objets est très faible car l'apparence d'un objet varie généralement d'une image à une autre. Et contrairement au *tracking* radar où la description d'un objet se limite à l'information géométrique et la longueur du signal, en *tracking* image, les informations couleur, texture, etc. sont accessibles.

Finalement, notre problème renferme les verrous scientifiques du *tracking* radar ainsi que celui de l'image. En effet, d'une part, nous ne disposons pas d'information autre que la localisation spatiale des détections (forme géométrique) pour construire les objets ce qui rejoint le *tracking* radar qui se limite à la forme géométrique et la vitesse des objets pour les suivre. D'autre part, nous sommes face aux différents problèmes du *tracking* image notamment la fragmentation, l'occultation, la fusion et la séparation d'objets.

A la fin de ce chapitre, nous avons présenté les données (séquences vidéo) ainsi que les métriques de performance qui nous ont servi pour tester nos différentes contributions et

les évaluer quantitativement. Le chapitre suivant est aussi bibliographique et concerne la théorie des fonctions de croyance, cadre théorique de ce travail.

Chapitre 3

Généralités sur les Fonctions de croyance

3.1 Introduction

La théorie des fonctions de croyance connue aussi sous le nom de la théorie de Dempster-Shafer, a été introduite par Dempster en 1967 [59] puis développée par Shafer en 1976 [60]. En se basant sur la théorie des probabilités, l'une des premières théories de l'incertain, Dempster a introduit la notion de croyance dans la définition même d'une probabilité [61]. Selon [59], cette mesure représente un degré de croyance quantifiant un état de connaissance pouvant être a priori ou faisant référence à des expériences préalables. La connaissance étant souvent partielle et insuffisante pour estimer la probabilité P d'un sous-ensemble A de l'ensemble des hypothèses, Dempster s'est inspiré de la théorie des probabilités généralisées [62] pour mettre en place les bases de la théorie des fonctions de croyance.

La théorie des probabilités généralisées qui remplit déjà les propriétés de la théorie des probabilités [61], associe à chaque sous-ensemble A un intervalle $[P_*, P^*]$ incluant sa probabilité P . Les bornes P_* et P^* sont appelées respectivement, probabilité inférieure et probabilité supérieure. Leur différence représente l'ampleur de l'imprécision sur l'incertitude telle que l'imprécision est nulle pour une probabilité P mesurable ($P_* = P^* = P$). Dempster a introduit les deux fonctions de croyance, crédibilité Bel et plausibilité Pl

comme étant P_* et P^* et ayant des propriétés analogues. Ainsi, $Bel(A)$ et $Pl(A)$ désignent respectivement la limite inférieure et la limite supérieure sur les degrés de croyance d'une source sur la réalisation de A .

En reprenant le modèle de Dempster fondé sur la théorie des probabilités généralisées, Shafer l'a présenté en tant que théorie des croyances (*evidence theory*) [60]. Ainsi, une croyance est une représentation d'un ensemble de preuves (*body of evidence*) telle que $Bel(A) = p$ signifie que la source dispose d'un degré de croyance p en A [63]. Du point de vue de Shafer, la règle de combinaison orthogonale (cf. Section 3.5.7) proposée par Dempster [59] pour fusionner des sources d'information * indépendantes (cf. Section 3.5.2) consiste à créer une nouvelle croyance (*evidence*) à partir des croyances associées aux sources combinées. En particulier, le conditionnement [59] (cf. Section 3.5.5) revient à combiner orthogonalement une croyance actuelle avec une croyance certaine en un sous-ensemble d'hypothèses.

Cependant, ces deux modèles ont fait l'objet de plusieurs critiques [63–65]. Dans [64], Pearl mentionne que si toute fonction de croyance représente une famille de distributions de probabilité, l'inverse n'est pas vrai. Il a aussi montré la non adéquation du conditionnement avec l'interprétation des probabilités généralisées. Dans [65], Zadeh critique le modèle évidentiel en soulignant, via un contre exemple, la limite de la règle orthogonale dans la combinaison de sources fortement conflictuelles.

A la suite de ces critiques, le modèle des croyances transférables TBM (*Transferable Belief Model*) est apparu en 1990 par Smets [66], permettant de s'affranchir des contraintes probabilistes. Nous avons choisi ce modèle pour développer notre travail. Par ailleurs, nous exposerons ses notions fondamentales dans ce chapitre bibliographique en mettant l'accent sur les outils de base qui nous ont servi.

3.2 Choix du TBM

Le TBM est un modèle très riche en opérateurs qui le rendent approprié à la modélisation et la combinaison d'informations incertaines, imprécises et/ou de nature différente (observations, connaissance a priori, fiabilité des sources, etc.). Il est alors possible d'intégrer des informations hétérogènes dans le même formalisme, d'utiliser de l'information contextuelle, d'exploiter la dépendance des sources ainsi que leur complémentarité et

*. Une source d'information peut représenter un capteur, un expert ou une base de données.

redondance potentielles, de combiner des sources partiellement conflictuelles, exploiter le conflit (cf. Section 3.7) entre sources pour des applications telles que la "fusion" de données ambiguës dans l'espace ou dans le temps.

Grâce à sa flexibilité, le TBM a connu un certain succès à travers plusieurs applications allant du TI (image médicale [67], télédétection [68], vidéosurveillance [69], association de données [7]) jusqu'aux applications financières.

3.3 Aperçu sur le TBM

L'ignorance telle que définie par Smets [66] prend deux formes : l'incertitude et l'imprécision. Une déclaration est imprécise si la valeur d'une variable d'intérêt n'est pas spécifiée dans son référentiel[†] (e.g. John a moins de 30 ans, John est jeune). Une déclaration est incertaine s'il est difficile d'évaluer sa vérité ou sa fausseté [66, 71]. Le modèle le plus approprié au traitement de l'incertain est la théorie des probabilités, en particulier l'approche bayésienne. Cependant, ce modèle n'est pas adapté au traitement de l'information imprécise. En cas d'ignorance totale, la théorie des probabilités suppose que toutes les hypothèses sont équiprobables même si elles sont incompatibles. Ainsi, la probabilité affectée à l'union des hypothèses est redistribuée de manière égale sur chacune des hypothèses simples. En passant par le modèle TBM, une confiance est attribuée à l'union sans privilégier aucune de ses hypothèses singletons. En fait, si une source d'information considère indifféremment un groupe d'hypothèses, il n'y a aucune raison de diviser arbitrairement la confiance qui lui a été attribuée.

Dans ce qui suit, le modèle TBM est présenté en tant que formalisme mathématique indépendant de la théorie des probabilités et distinguant deux niveaux d'abstraction : (i) un niveau crédal (du latin «credo» signifiant «je crois») siège de la représentation d'information (partie statique) et des combinaisons et des raisonnements sur ces connaissances (partie dynamique) ; (ii) un niveau décisionnel chargé de la prise de décision à partir des connaissances traitées.

[†]. Le mot référentiel est utilisé dans [70] et correspond à notre cadre de discernement.

3.4 Niveau Créal : partie statique

A ce niveau, la question qui se pose est comment présenter la connaissance issue d'une source d'information sous forme de fonctions de croyance. La définition du cadre de discernement dans lequel la croyance est répartie est une étape fondamentale et assez cruciale.

3.4.1 Cadre de discernement

Afin d'envisager une solution à un problème traité, Smets [66] distingue trois types d'hypothèses : hypothèses connues possibles KP (*Known as Possible*), hypothèses inconnues UH (*Unknown Hypotheses*) et hypothèses connues impossibles KI (*Known as Impossible*). La répartition des hypothèses entre KP, UH et KI dépend du problème traité et de l'information disponible. En rajoutant de l'information, deux cas peuvent se produire [66] : (i) transfert d'hypothèses de KP vers KI ce qui correspond au conditionnement permettant d'écarter des hypothèses jugées impossibles selon le nouveau état de connaissance ; (ii) transfert d'hypothèses de UH vers KP pour introduire des hypothèses manquantes jugées pertinentes.

Selon Smets [66], une méta croyance construite dans l'ensemble total d'hypothèses, sur leur appartenance à chacun des ensembles KP, UH et KI, permet de mieux les répartir. L'ensemble KP de N hypothèses ω_i mutuellement exclusives est appelé cadre de discernement noté Ω :

$$\Omega = \{\omega_1, \dots, \omega_N\}. \quad (3.1)$$

A partir de Ω , l'ensemble 2^Ω des 2^N sous-ensembles disjonctifs de Ω (par exemple $\{\omega_1, \omega_N\}$ signifie ω_1 ou ω_N) est construit. Il représente un référentiel de définition pour les fonctions de croyance utilisées par le TBM pour évaluer la véracité des sous-ensembles disjonctifs en tant que solution du problème considéré.

$$2^\Omega = \{\emptyset, \omega_1, \dots, \omega_N, \dots, \{\omega_1, \omega_N\}, \dots, \Omega\}, \quad (3.2)$$

pour alléger la notation nous ne notons pas les $\{ \}$ autour d'hypothèses de cardinal inférieur à 1.

Nous reviendrons plus loin (cf. Section 3.5.9) sur la définition du cadre de discernement pour spécifier les mécanismes qui permettent de le manipuler.

3.4.2 Fonction de masse (bba)

Une fonction de masse bba (*basic belief assignment*) m^Ω est une fonction de 2^Ω vers $[0,1]$ définie par :

$$\begin{aligned} m^\Omega : \quad 2^\Omega &\longrightarrow [0, 1], \\ \sum_{A \in 2^\Omega} m^\Omega(A) &= 1. \end{aligned} \tag{3.3}$$

$m^\Omega(A)$ est la part de croyance allouée sur A appelée bba et à aucun de ses sous-ensembles qui, en cas d'information supplémentaire, pourra être redistribuée plus précisément sur les sous-ensembles de A . Les éléments focaux de m^Ω sont les sous-ensembles A de Ω ayant une masse non nulle. Dans le cas du monde fermé [59] où Ω est supposé exhaustif, la masse sur $\emptyset$ est nulle. Smets [66] introduit la notion du monde ouvert et suppose que toute masse allouée sur un sous-ensemble A de Ω est en réalité allouée sur $\{A, UH\}$ où UH désigne l'ensemble des hypothèses inconnues mentionné dans la section 3.4.1. Ainsi, la masse $m^\Omega(\emptyset)$ n'est autre que la masse allouée sur UH [66]. Cependant, nous notons que $m^\Omega(\emptyset)$ peut s'interpréter autrement, en tant que signe d'un dysfonctionnement dû à un problème de modélisation (fiabilité de la source, paramétrisation de la bba...).

Une bba est caractérisée par l'ensemble de ses éléments focaux appelé *noyau* qui définit également son type :

Définition 1. Bba simple SSF (*Simple Support Function*)

La bba qui possède deux éléments focaux dont l'un est Ω [72].

Définition 2. Bba consonante

La bba dont les éléments focaux forment un ensemble totalement ordonné au sens de l'inclusion. Ainsi, $\forall A, B \in 2^\Omega$ si $m^\Omega(A) \neq \emptyset$ et $m^\Omega(B) \neq \emptyset$ alors $A \subseteq B$ ou $B \subseteq A$ [72].

Définition 3. Bba catégorique

La bba qui possède un seul élément focal. Deux cas limites de bba catégorique ont des noms spéciaux [72].

Définition 4. Bba vide

La bba catégorique dont Ω est le seul élément focal. Elle représente l'ignorance totale ($m^\Omega(\Omega) = 1$) [72].

Définition 5. Bba contradictoire

La bba catégorique dont $\emptyset$ est le seul élément focal ($m^\Omega(\emptyset) = 1$) [72].

Définition 6. Bba normale

La bba telle que $m^\Omega(\emptyset) = 0$. Dans le cas contraire, elle est dite sous-normale [72].

Définition 7. Bba dogmatique

La bba telle que Ω n'est pas un élément focal ($m^\Omega(\Omega) = 0$) [72].

Définition 8. Bba bayésienne

La bba dont les éléments focaux sont des singletons [72].

3.4.3 Transformations de la fonction de masse

Une croyance peut être représentée sous différentes formes. La forme basique est la fonction de masse à partir de laquelle d'autres fonctions en relation bijective sont déduites (cf. Equations (3.5), (3.9), (3.11), (3.14)), à savoir : la crédibilité bel , la plausibilité pl , la communalité q et l'implicabilité b . Ces fonctions ayant d'autres interprétations en terme de répartition de la croyance, jouent un rôle important dans la simplification des calculs lors de la combinaison des croyances. Elles sont toutes définies dans $[0,1]$ mais la somme des valeurs de leurs éléments focaux peut dépasser 1.

Définition 9. Crédibilité

Pour $A \subseteq \Omega$, la bba $m^\Omega(A)$ ne quantifie pas notre croyance en la véracité de A (que la vérité w_0 est dans A). Par exemple, pour $A = \{w_1, w_2\}$, $m^\Omega(A)$ peut se transférer vers $\{w_1\}$ si une nouvelle information le permet. Mais, étant donné l'état actuel de connaissance, la bba est uniquement allouée sur A .

Notre croyance en A notée $bel^\Omega(A)$ est quantifiée via la fonction de crédibilité bel^Ω [73]. Une bba $m^\Omega(B)$ allouée sur $B \subseteq A$ supporte que la vérité soit dans A . Ainsi, $bel^\Omega(A)$ est définie comme étant la somme de toutes les bbas $m^\Omega(B)$ pour $B \subseteq A$:

$$bel^\Omega(\emptyset) = 0 \text{ et } bel^\Omega(A) = \sum_{\emptyset \neq B \subseteq A} m^\Omega(B), \forall A \subseteq \Omega \setminus \{\emptyset\}. \quad (3.4)$$

$$\forall A \subseteq \Omega, \quad m^\Omega(A) = \sum_{C \subseteq A} (-1)^{|A|-|C|} bel^\Omega(C). \quad (3.5)$$

Par conséquent, $bel^\Omega(A)$ quantifie le support justifié et spécifique de A et représente alors, la croyance minimale en A . Notons que $\emptyset$ n'est pas inclu dans $bel(A)$ car il supporte A et aussi son complémentaire $\bar{A}$.

La fonction bel ainsi définie (cf. Equation (3.4)) vérifie l'inégalité suivante :

$$\begin{aligned} \forall n > 0 \text{ et } A_1, A_2, \dots, A_n \subseteq \Omega, \quad bel^\Omega(A_1 \cup \dots \cup A_n) &\geq \sum_i bel^\Omega(A_i), \\ \sum_{i < j} bel^\Omega(A_i \cap A_j) + \dots + \sum (-1)^{n+1} bel^\Omega(A_i \cap \dots \cap A_n). \end{aligned} \quad (3.6)$$

Comme nous l'avons déjà mentionné, l'originalité du TBM est dans l'allocation des masses sur les disjonctions d'hypothèses de Ω . Par ailleurs, si $m(A) = 0, \forall A \subseteq \Omega / |A| > 1$, l'inégalité (3.6) devient une égalité, la fonction bel se réduit à une fonction de probabilité : un cas rarement atteint dans des applications pratiques.

Définition 10. Plausibilité

C'est le pendant de la crédibilité. Elle est notée pl [73] et définie par :

$$\forall A \subseteq \Omega, \quad pl^\Omega(A) = 1 - m^\Omega(\emptyset) - bel^\Omega(\bar{A}), \quad (3.7)$$

$$pl^\Omega(A) = \sum_{B \subseteq \Omega / B \cap A \neq \emptyset} m^\Omega(B). \quad (3.8)$$

$$\forall A \subseteq \Omega, m^\Omega(A) = \sum_{C \subseteq A} (-1)^{|A|-|C|+1} pl^\Omega(\bar{C}). \quad (3.9)$$

Le degré de plausibilité $pl^\Omega(A), \forall A \subseteq \Omega$ quantifie la quantité maximale du support spécifique et potentiel qui renforce la véracité de A . Ce support est potentiel puisqu'en présence de nouvelles informations, les bbas incluses dans $pl(A)$ peuvent être transférées vers des sous-ensembles non vides de A , c'est le cas par exemple où nous apprenons que la vérité ω_0 n'est pas dans $\bar{A}$.

Définition 11. Communalité

C'est une fonction définie par :

$$\forall A \subseteq \Omega, \quad m^\Omega(A) = \sum_{A \subseteq C} (-1)^{|C|-|A|} q^\Omega(C), \quad (3.10)$$

$$\forall A \subseteq \Omega, \quad m^\Omega(A) = \sum_{A \subseteq C} (-1)^{|C|-|A|} q^\Omega(C). \quad (3.11)$$

Elle n'a pas d'interprétation intuitive. Elle sert pour simplifier les calculs lors de la combinaison conjonctive des croyances et dans la décomposition canonique des bbas.

Définition 12. Implicabilité

L'implicabilité b de A est définie comme étant la somme des masses sur les sous-ensembles B de A , y compris $\emptyset$.

$$\forall A \subseteq \Omega, \quad b^\Omega(A) = \sum_{B \in 2^\Omega / B \subseteq A} m^\Omega(B). \quad (3.12)$$

$$b^\Omega(A) = bel^\Omega(A) + m^\Omega(\emptyset). \quad (3.13)$$

$$\forall A \subseteq \Omega, \quad m^\Omega(A) = \sum_{C \subseteq A} (-1)^{|A|-|C|} b^\Omega(C). \quad (3.14)$$

De même que la communalité, l'implicabilité n'a pas d'interprétation intuitive. Mais, elle sert pour simplifier les calculs lors de la combinaison disjonctive des croyances.

3.5 Niveau Créal : partie dynamique

A ce niveau, il s'agit de combiner plusieurs sources de fonctions de croyance afin de résumer l'information et améliorer la prise de décision. Les sources délivrant une information partielle, imprécise et/ou incertaine sur les variables d'intérêt sont souvent partiellement complémentaires et partiellement redondantes et choisies selon le problème traité. Par ailleurs, le schéma de fusion de plusieurs sources doit prendre en compte leurs fiabilités et leurs dépendances pour deux raisons : (i) éviter un problème de modélisation lié à une très forte hypothèse sur la fiabilité des sources pouvant engendrer du conflit (cf. Section 3.7) entre sources ; (ii) utiliser la règle de fusion la plus adéquate.

3.5.1 Notions de redondance et de complémentarité des sources

La minimisation du temps de traitement ainsi que des ressources matérielles est un des critères pour la fusion d'information, dont la difficulté augmente avec le nombre de sources à fusionner. Ainsi, une attention particulière doit être accordée au choix des

sources. Cette sélection peut s'effectuer selon des critères dynamiques mesurant l'intérêt d'une source par rapport à un état actuel de connaissance. La mesure d'intérêt elle-même doit prendre en compte des mesures de complémentarité et de redondance de la source par rapport aux autres (ou au résultat de leur combinaison) ce qui permet d'évaluer la qualité d'information qu'elle apporte, en plus de sa quantité (liée à la multiplicité des sources d'information).

Ces deux notions de complémentarité et de redondance sont essentielles pour la gestion des sources : (i) la redondance des sources fortifie et robustifie l'information ce qui permet de résister à la défaillance de quelques sources d'information (e.g. panne). Cependant, elle n'intervient pas dans l'amélioration des performances du système de fusion et augmente le temps de traitement ; (ii) la complémentarité des sources permet de résoudre les indiscernabilités et imprécisions des sources prises individuellement ce qui donne une information globale plus complète [74–76]. En effet, nous entendons par complémentarité la différence dans les informations délivrées. Ainsi, une information peut être accessible par une source et absente pour une autre source.

3.5.2 Notion d'indépendance des sources

Il existe deux types d'indépendance : (i) indépendance statistique ; (ii) indépendance cognitive. L'indépendance statistique [77] est rarement vérifiée du fait que les sources observent souvent les mêmes phénomènes physiques et sont de même nature. Elles sont alors statistiquement modélisées de manière très proche.

Smets a introduit la notion d'indépendance cognitive des sources [78] qui est moins restrictive que l'indépendance statistique. Elle est vérifiée si les sources ne communiquent pas entre elles. Ainsi, l'information donnée par l'une ne modifie pas l'information fournie par l'autre.

L'indépendance cognitive est différente de la corrélation définie en tant que maximum d'information commune entre les sources (e.g. deux experts peuvent fournir un même diagnostic sans en avoir discuté, ni en ayant eu la même formation, etc.) [79].

3.5.3 Affaiblissement des croyances

La modélisation spécifique d'une source doit prendre en compte des connaissances quant à cette source notamment en terme de fiabilité (à donner une information sûre) [80–82].

L'opération souvent appliquée aux fonctions de croyance associées à une source, pour intégrer sa fiabilité est l'affaiblissement. L'affaiblissement le plus classique dit global a été proposé par Shafer [60] puis axiomatisé et justifié par Smets via l'application du théorème de Bayes généralisé [78]. Il est défini comme suit :

Soit m^Ω une bba de degré de fiabilité $\alpha \in [0, 1]$. La bba affaiblie notée $m^{\Omega, \alpha}$ (ou plus simplement m^α) est donnée par :

$$\begin{aligned} \forall A \in 2^\Omega, m^{\Omega, \alpha}(A) &= \alpha m^\Omega(A), \\ m^{\Omega, \alpha}(\Omega) &= \alpha m^\Omega(\Omega) + (1 - \alpha). \end{aligned} \quad (3.15)$$

Si $\alpha = 0$ alors la source est totalement non-fiable et toute la masse est transférée sur Ω (m^α est la bba vide).

L'affaiblissement de la fonction de masse m^Ω engendre l'affaiblissement de la fonction de crédibilité bel^Ω et le renforcement de la fonction de plausibilité pl^Ω [75, 80]. Dans [78], Smets présente une justification théorique de l'affaiblissement en considérant l'information de fiabilité comme une source d'information supplémentaire. Deux propriétés intéressantes de l'affaiblissement sont à noter : (i) l'affaiblissement d'une fonction de croyance vide est une fonction de croyance vide ; (ii) deux affaiblissements successifs α_1 et α_2 d'une fonction de croyance sont égaux à l'affaiblissement par le produit $\alpha_1 \alpha_2$.

D'autres types d'affaiblissement ont été proposés, par exemple l'affaiblissement contextuel [83, 84] qui estime le degré de fiabilité de la source en fonction de l'hypothèse considérée, l'affaiblissement fondé sur les fonctions de crédibilité et plausibilité [85], et l'affaiblissement généralisé fondé sur l'affaiblissement des SSF qui constituent la bba d'origine.

3.5.4 Principe du moindre engagement et ordonnancement

Le principe du moindre engagement PMI (Principe du Minimum d'Information) est central au TBM. Son principe est le suivant : si nous devons choisir une fonction de croyance parmi plusieurs fonctions de croyance satisfaisant un ensemble de contraintes, nous devrions choisir la moins engagée, autrement dit, la moins informative. Ainsi, toute confiance accordée à un sous-ensemble de Ω doit être justifiée afin d'éviter l'ajout d'une information arbitraire non contenue dans les contraintes définies.

Dans [86], Yager a étudié la mesure du taux d'imprécision d'une bba via la non spécificité

$\mathcal{N}$ [87] qui se calcule à partir des cardinaux des éléments focaux de la bba et des valeurs de leurs masses (cf. Equation (3.16)). Ainsi, plus les masses sur des sous-ensembles de Ω de grandes cardinalités sont importantes plus la bba est non-spécifique et moins elle est engagée. A titre d'exemple, la bba vide est la moins engagée tandis qu'une bba précise telle qu'une bba bayésienne est totalement engagée.

$$\mathcal{N}(m^\Omega) = \sum_{\emptyset \neq A \subseteq \Omega} m^\Omega(A) \cdot \log_2(|A|). \quad (3.16)$$

Par ailleurs, il est possible d'utiliser l'ordonnancement des fonctions de croyance lorsqu'elles sont définies dans un même cadre de discernement, pour comparer leurs engagements [88–90]. Dans ce qui suit, nous présentons les relations d'ordre les plus utilisées.

Définition 13. pl-ordonnancement

Pour deux bbas m_1^Ω et m_2^Ω , $m_1^\Omega \sqsubseteq_{pl} m_2^\Omega$ si et seulement si $\forall A \subseteq \Omega$, $pl_1(A) \leq pl_2(A)$.

Définition 14. q-ordonnancement

Pour deux bbas m_1^Ω et m_2^Ω , $m_1^\Omega \sqsubseteq_q m_2^\Omega$ si et seulement si $\forall A \subseteq \Omega$, $q_1(A) \leq q_2(A)$.

Définition 15. s-ordonnancement

Pour deux bbas m_1^Ω et m_2^Ω , $m_1^\Omega \sqsubseteq_s m_2^\Omega$ si et seulement s'il existe une fonction non-négative S de $2^\Omega \times 2^\Omega$ dans $[0,1]$ vérifiant $\forall B \subseteq \Omega$, $\sum_{A \subseteq B} S(B, A) = 1$ et telle que :

$$m_1^\Omega(A) = \sum_{B \subseteq \Omega} S(B, A) m_2^\Omega(B). \quad (3.17)$$

S est appelée fonction de spécialisation [91, 92]. Elle est représentée par une matrice stochastique en colonne[‡] qui vérifie $m_1 = S \times m_2$. Chacun des termes non nuls de cette matrice ($S(B, A) = 0, \forall A \not\subseteq B$) représente la part de la masse sur B transférée sur A . Ainsi, m_1 est reconstruite en répartissant la masse sur tous les éléments focaux de m_2 sur leurs sous-ensembles. m_1 est bien une spécialisation de m_2 . A l'inverse, m_2 est une généralisation de m_1 . Notons également que la bba vide est la bba la moins engagée au sens de pl, q et s et que $\sqsubseteq_s$ implique $\sqsubseteq_{pl}$ et $\sqsubseteq_q$.

[‡]. Une matrice stochastique en colonne est une matrice telle que ses éléments sont entre 0 et 1 et la somme de ses éléments sur une ligne est égale à 1

3.5.5 Conditionnement

Etant donnée une bba m^Ω définie sur Ω traduisant un état de connaissance initial, le conditionnement est l'opération qui s'applique sur m^Ω pour intégrer une nouvelle information qu'un sous-ensemble C de Ω est certain, signifiant que la solution au problème posé (hypothèse vraie) n'est certainement pas dans $\bar{C}$. La bba conditionnée notée $m^\Omega[C]$ (ou encore $m^\Omega(. / C)$) est telle que tous ses éléments focaux autres que $\emptyset$ sont inclus dans C induisant un transfert de masse de chaque élément focal de m^Ω sur son intersection avec C . Ainsi, une masse non nulle sur $\emptyset$ ($m^\Omega[C](\emptyset)$) reflète un conflit (cf. Section 3.7) entre la croyance initiale et la nouvelle. Les différentes grandeurs conditionnées s'obtiennent par les formules suivantes :

$$\forall A \in 2^\Omega, m^\Omega[C](A) = \sum_{B \in 2^\Omega / A=B \cap C} m^\Omega(B). \quad (3.18)$$

$$bel^\Omega[C](A) = bel^\Omega(A \cup \bar{C}) - bel^\Omega(\bar{C}). \quad (3.19)$$

$$pl^\Omega[C](A) = pl^\Omega(A \cap C). \quad (3.20)$$

$$b^\Omega[C](A) = b^\Omega(A \cup \bar{C}). \quad (3.21)$$

$$q^\Omega[C](A) = q^\Omega(A), \text{ si } A \subseteq C, \text{ 0 sinon.} \quad (3.22)$$

Notons également que $m^\Omega[C]$ se justifie à l'aide du PMI. En effet, $m^\Omega[C]$ est la spécialisation la moins engagée parmi toutes les spécialisations de m^Ω vérifiant la contrainte sur la plausibilité $pl^\Omega[C](\bar{C}) = 0$.

3.5.6 Déconditionnement

Si nous disposons d'une fonction de masse $m^\Omega[C]$ définie dans Ω et conditionnée en $C \subseteq \Omega$ alors il n'est généralement pas possible de retrouver la bba m^Ω avant conditionnement. Parmi toutes les bbas possibles, une solution consiste à choisir la bba la moins engagée $m^\Omega[C] \uparrow^\Omega$ selon laquelle la masse affectée à $B \subseteq \Omega$ avant déconditionnement est transférée sur $B \cup \bar{C}$ où $\bar{C}$ dénote le complémentaire de C dans Ω :

$$\forall B \in 2^\Omega, m^\Omega[C] \uparrow^\Omega(B \cup \bar{C}) = m^\Omega[C](B). \quad (3.23)$$

Cette solution assure que le conditionnement sur C de la bba déconditionnée via l'opérateur $\uparrow$ [78], donne de nouveau $m^\Omega[C]$ qui est aussi la solution la moins engagée.

3.5.7 Combinaison des croyances

En théorie des fonctions de croyance, plusieurs règles de combinaison ont été proposées. Nous pouvons distinguer entre les règles de type conjonctif [60, 66, 93], disjonctif [93] et hybride [89, 94]. Le type conjonctif suppose que les sources sont fiables et permet de tirer une information à la fois plus fiable et plus précise. Le type disjonctif suppose qu'au moins l'une des sources est fiable et engendre une perte de spécificité au niveau des croyances. Le type hybride a été proposé afin de gérer le conflit (cf. Section 3.7) entre les sources. Son inconvénient majeur est la non associativité de la combinaison. Dans ce qui suit, nous définissons les règles de combinaison les plus utilisées.

Définition 16. Règle orthogonale

Appelée aussi règle de Dempster notée $\oplus$, c'est la première règle de combinaison, proposée par Shafer [59], pour combiner des sources indépendantes. Elle est de type conjonctif et suppose le monde fermé. La somme orthogonale de deux bbas m_1 et m_2 définies sur un même cadre de discernement Ω s'écrit :

$$\forall C \in 2^\Omega \setminus \{\emptyset\}, m_{\oplus}(C) = \frac{1}{1 - K} \sum_{\substack{(A,B) \in 2^\Omega \times 2^\Omega \\ A \cap B = C}} m_1(A) m_2(B), \quad (3.24)$$

avec

$$K = \sum_{(A,B) \in 2^\Omega \times 2^\Omega / A \cap B = \emptyset} m_1(A) m_2(B), \quad (3.25)$$

K est le coefficient de normalisation. L'opérateur $\oplus$ est associatif, commutatif et admet la bba vide comme élément neutre.

Définition 17. Règle conjonctive TBM

Notée $\odot$ et proposée par Smets dans le cadre du TBM [95], elle est définie par l'équation (3.26). Faisant l'hypothèse du monde ouvert, la règle revient à ne pas normaliser la somme orthogonale donnée par l'équation (3.24), par la masse sur le vide.

$$\forall C \in 2^\Omega, m_{\odot}(C) = \sum_{\substack{(A,B) \in 2^\Omega \times 2^\Omega \\ A \cap B = C}} m_1(A) m_2(B). \quad (3.26)$$

De manière plus simple et en passant par la fonction de communalité, la combinaison conjonctive s'écrit $q_{\odot}(A) = q_1(A).q_2(A), \forall A \subseteq \Omega$ où q_1, q_2 et $q_{\odot}$ sont les communalités associées respectivement à m_1, m_2 et $m_{\odot}$.

Les deux règles précédentes peuvent se justifier par le PMI [91, 92, 96]. En effet, elles sont telles que le conditionnement $m^\Omega[C]$ (cf. Section. 3.5.5) equivaut à la combinaison de m^Ω avec une bba catégorique $m_C^\Omega/m_C^\Omega(C) = 1$.

Notons également que contrairement à la règle orthogonale, la règle TBM est une spécialisation des bbas combinées.

Définition 18. Règle prudente

Pour des sources dépendantes de bbas m_1^Ω et m_2^Ω , des règles prudentes notées $\oslash$ ont été définies [77, 97] pour prendre en compte la redondance entre les croyances. Suivant un ordonnancement $x \in \{q, pl, s, \dots\}$, elles sont données par :

$$m \oslash = LC_{bba}(S_x(m_1) \cap S_x(m_2)), \quad (3.27)$$

où LC_{bba} dénote la bba la moins engagée d'un ensemble de bbas, $S_x(m_1)$ et $S_x(m_2)$ désignent respectivement, l'ensemble des bbas les plus engagées que m_1 et les plus engagées que m_2 . Puisque les ordonnancements ne sont que partiels, LC_{bba} peut ne pas exister. Ainsi, les règles prudentes ne sont pas toujours applicables.

Parmi les règles prudentes, la plus répandue est celle introduite par Denoeux [98]. Nous la spécifierons plus loin (cf. Définition 22).

Définition 19. Règle disjonctive

Proposée dans [78, 99] et notée $\oplus$, elle permet de combiner des sources dont l'une au moins est fiable. Elle est donnée par :

$$\forall C \in 2^\Omega, m_\oplus(C) = \sum_{\substack{(A,B) \in 2^\Omega \times 2^\Omega \\ A \cup B = C}} m_1(A) m_2(B). \quad (3.28)$$

D'après l'équation (3.28), la règle disjonctive est une généralisation. Par ailleurs, une manière plus simplifiée de définir cette règle est de passer par les implicabilités. Ainsi, la fonction $b_\oplus$ associée à $m_\oplus$ est telle que $b_\oplus = b_1.b_2$ où b_1 et b_2 sont les implicabilités associées respectivement à m_1 et m_2 .

Définition 20. Règle de Dubois & Prade

C'est une règle de type hybride [99]. Le premier terme dans sa définition, ci-dessous, n'est autre que la combinaison conjonctive de m_1 et m_2 . Il permet de traiter la part de complémentarité non conflictuelle entre les sources. Le deuxième terme gère la part de

complémentarité conflictuelle en réallouant la masse issue de toute conjonction d'hypothèses conflictuelles (dont l'intersection est vide) sur leur disjonction.

$$\begin{aligned} \forall C &\in 2^\Omega \setminus \{\emptyset\}, \\ m(C) &= \sum_{\substack{(A,B) \in 2^\Omega \times 2^\Omega \\ A \cap B = C}} m_1(A) m_2(B) + \sum_{\substack{(A,B) \in 2^\Omega \times 2^\Omega \\ A \cup B = C, \\ A \cap B = \emptyset}} m_1(A) m_2(B). \end{aligned} \quad (3.29)$$

Cette règle est commutative mais non associative.

Définition 21. Règle de Yager

Par rapport à la règle précédente, la règle de Yager [89] définie comme suit, transfère la masse sur $\emptyset$, issue de la combinaison conjonctive des sources, sur Ω ainsi, enlève l'hypothèse qu'au moins une des sources est fiable et le conflit (cf. Section 3.7) entre les sources est interprété en tant que degré d'ignorance.

$$m(C) = \begin{cases} \sum_{\substack{(A,B) \in 2^\Omega \times 2^\Omega \\ A \cap B = C}} m_1(A) m_2(B), & \forall C \in 2^\Omega \setminus \{\emptyset, \Omega\}, \\ \sum_{\substack{(A,B) \in 2^\Omega \times 2^\Omega \\ A \cap B = \emptyset}} m_1(A) m_2(B) + m_1(\Omega).m_2(\Omega), & C = \Omega, \\ 0, & C = \emptyset. \end{cases} \quad (3.30)$$

3.5.8 Décomposition canonique

Selon la définition 1, une SSF est une bba qui admet au plus deux éléments focaux : $A \subsetneq \Omega$ et Ω . Nous présentons ici, sa notation $A^{w(A)}$ introduite par Shafer où $w(A) \in [0, 1[$ est le poids associé à l'élément A vérifiant $A^{w(A)}(A) = 1 - w(A)$ et $A^{w(A)}(\Omega) = w(A)$. A titre d'exemple, une bba catégorique sur A est notée A^0 .

Dans [60], Shafer appelle bba séparable toute bba non dogmatique m issue de la combinaison orthogonale d'un ensemble de SSFs :

$$m = \oplus_{A \subsetneq \Omega} A^{w(A)}, w(A) \in [0, 1], \forall A \subsetneq \Omega. \quad (3.31)$$

Cette représentation est alors, une décomposition canonique de la bba non dogmatique m . Elle est unique si ses éléments $\{A\}_{w(A) \neq 1}$ sont tous différents.

En introduisant la notion de fonction à support simple généralisée GSSF (*Generalized Simple Support Function*), Smets [100] a proposé la décomposition canonique de toute

bba non dogmatique m en un ensemble de GSSFs en se basant sur la règle conjonctive TBM.

$$m = \bigodot_{A \subsetneq \Omega} A^{w(A)}, w(A) \in (-\infty, +\infty), \forall A \subsetneq \Omega. \quad (3.32)$$

Une GSSF, généralisation d'une SSF, est une fonction $\mu = A^{w(A)}$ définie de 2^Ω dans $\mathbb{R}$ par : $\mu(A) = 1 - w(A)$, $\forall A \neq \Omega$, $\mu(\Omega) = w(A)$ et $\mu(B) = 0$, $\forall B \in 2^\Omega / \{A, \Omega\}$ avec $w(A) \in (-\infty, +\infty)$. Ainsi, suivant la valeur de $w(A)$, trois cas se présentent :

- (i) si $w(A) \in [0, 1]$ $\forall A \subseteq \Omega$ alors μ est une SSF qui traduit une confiance en A et la bba m est séparable.
- (ii) si $w(A) \in]1, +\infty[$ alors μ est une ISSF (*Inverse Simple Support Function*) qui exprime une défiance en A .
- (iii) si $w \in (-\infty, 0[$ alors μ est une RSSF (*a square Root Simple Support Function*) ou IRSSF (*a square Root of an Inverse Simple Support Function*) [100].

Par ailleurs, la fonction de poids w est déduite de la fonction de communalité q associée à la bba m comme suit :

$$w(A) = \prod_{B \supseteq A} q(B)^{(-1)^{|B|-|A|+1}}. \quad (3.33)$$

$$\left\{ \begin{array}{l} = \frac{\prod_{B \supseteq A, |B| \notin 2\mathbb{N}} q(B)}{\prod_{B \supseteq A, |B| \in 2\mathbb{N}} q(B)}, \text{ si } |A| \in 2\mathbb{N}, \\ = \frac{\prod_{B \supseteq A, |B| \in 2\mathbb{N}} q(B)}{\prod_{B \supseteq A, |B| \notin 2\mathbb{N}} q(B)}, \text{ sinon,} \end{array} \right. \quad (3.34)$$

où $2\mathbb{N}$ dénote l'ensemble des entiers pairs. Afin de minimiser la complexité du calcul (utiliser la multiplication matricielle [101] ou l'algorithme *Fast Mobius Transform* [102]), l'équation ci-dessous est obtenue à partir de l'équation (3.33) en appliquant le logarithme :

$$\log(w(A)) = - \sum_{B \supseteq A} (-1)^{|B|-|A|} \log(q(B)). \quad (3.35)$$

Notons également que la normalisation d'une bba sous normale $m = \bigodot_{A \subsetneq \Omega} A^{w(A)}$ revient à la combiner conjonctivement avec la GSSF $\emptyset^k$ où $k = (1 - m(\emptyset))^{-1}$. Par ailleurs, pour $m_1 = \bigodot_{A \subsetneq \Omega} A^{w_1(A)}$ et $m_2 = \bigodot_{A \subsetneq \Omega} A^{w_2(A)}$, la règle de combinaison conjonctive s'écrit :

$$m_1 \bigodot m_2 = \bigodot_{A \subsetneq \Omega} A^{w_1(A)} \bigodot_{A \subsetneq \Omega} A^{w_2(A)}, \quad (3.36)$$

$$= \bigodot_{A \subsetneq \Omega} A^{w_1(A)w_2(A)}. \quad (3.37)$$

Ainsi, $w_{1 \bigodot 2} = w_1 \cdot w_2$.

Définition 22. Règle conjonctive prudente

Le w -ordonnancement ($\sqsubseteq_w$) pour une bba non dogmatique a été introduit par Dencœux [98] : $m_1 = \bigoplus_{A \subseteq \Omega} A^{w_1(A)}$, $m_2 = \bigoplus_{A \subseteq \Omega} A^{w_2(A)}$, $m_1 \sqsubseteq_w m_2$ si et seulement si $\forall A \subseteq \Omega, w_1(A) \leq w_2(A)$. Notons que $\sqsubseteq_w$ implique $\sqsubseteq_s$. En se basant sur le w -ordonnancement, Dencœux a proposé la règle prudente pour une bba non dogmatique :

$$m_1 \bigwedge_2 m_2 = w - LC_{bba}(S_w(m_1) \cap S_w(m_2)), \quad (3.38)$$

telle que $w_1 \bigwedge_2 w_2(A) = \min\{w_1(A), w_2(A)\}, \forall A \subseteq \Omega$. L'existence et l'unicité de cette règle sont assurées.

3.5.9 Opérations sur les cadres de discernement

Nous avons considéré dans les sections précédentes que la définition des fonctions de croyance était dans un même cadre de discernement. Or, dans plusieurs applications pratiques, nous nous trouvons avec différents référentiels pour les sources : des sources complémentaires et ayant des capacités discriminatives différentes peuvent mener à des référentiels qui se complètent ou à des référentiels à différents degrés de granularité ; des sources hétérogènes c'est-à-dire qui concernent différentes variables d'intérêt, induisent des référentiels totalement indépendants. Le TBM offre encore de nombreux outils pour gérer ces différents types de sources dont la fusion nécessite trois étapes : (i) construction d'un cadre de discernement commun ; (ii) redéfinition des croyances dans ce cadre ; (iii) combinaison des croyances en se basant sur les règles de fusion définies dans la section 3.5.7.

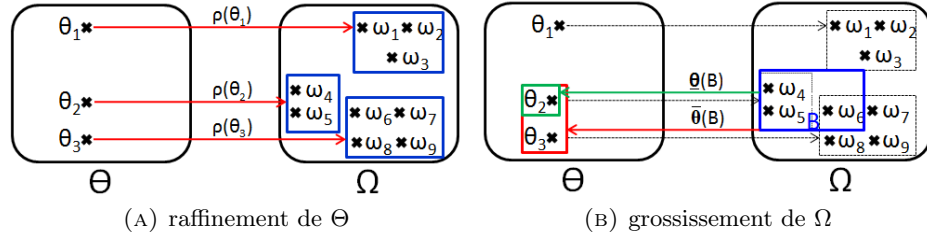
Dans ce qui suit, nous nous focalisons sur les étapes (i) et (ii) et nous détaillons certains mécanismes de modification des cadres de discernement tels que le raffinement et le grossissement. Nous présentons aussi des mécanismes de redéfinition des croyances qui obéissent au PMI.

Définition 23. Grossissement et raffinement

Soient Θ et Ω deux cadres de discernement, un raffinement de Θ vers Ω est une fonction $\rho : 2^\Theta \longrightarrow 2^\Omega$ qui vérifie : (i) l'ensemble $\{\rho(\theta), \theta \in \Theta\} \subseteq 2^\Omega$ est une partition de Ω ; (ii) $\rho(A) = \{\rho(\theta), \theta \in A\}, \forall A \subseteq \Theta$.

Ω est alors un raffinement de Θ et Θ est un grossissement de Ω .

Dans la figure 3.1a, nous présentons un exemple de raffinement où :

FIGURE 3.1: Exemple d'un grossissement Θ de Ω associé à un raffinement ρ de Θ

$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9\}$ et $\Theta = \{\theta_1, \theta_2, \theta_3\}$ avec $\theta_1 = \{\omega_1, \omega_2, \omega_3\}$, $\theta_2 = \{\omega_4, \omega_5\}$ et $\theta_3 = \{\omega_6, \omega_7, \omega_8, \omega_9\}$.

Ainsi, Θ est un grossissement de Ω fondé sur la fonction $\rho : 2^\Theta \rightarrow 2^\Omega$ définie par :

$$\begin{aligned} \rho(\theta_1) &= \{\omega_1, \omega_2, \omega_3\} & \rho(\theta_2) &= \{\omega_4, \omega_5\} \\ \rho(\theta_3) &= \{\omega_6, \omega_7, \omega_8, \omega_9\} & \rho(\{\theta_1, \theta_2\}) &= \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\} \\ \rho(\{\theta_1, \theta_3\}) &= \{\omega_1, \omega_2, \omega_3, \omega_6, \omega_7, \omega_8, \omega_9\} & \rho(\{\theta_2, \theta_3\}) &= \{\omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9\} \\ \rho(\Theta) &= \Omega & \rho(\emptyset) &= \emptyset \end{aligned}$$

La question qui se pose est comment exprimer un état de connaissance sur Θ traduit par la fonction de masse m^Θ dans le cadre de discernement plus fin Ω . Soit m^Ω la bba à définir. L'opérateur qui permet cette redéfinition et qui satisfait au PMI est notée $\uparrow$. Il est tel que :

$$\forall A \subseteq \Theta, m^{\Theta \uparrow \Omega}(\rho(A)) = m^\Theta(A). \quad (3.39)$$

D'un autre coté, $\downarrow$ est l'opérateur dual de $\uparrow$. Il redéfinit m^Ω dans le grossissement Θ selon l'équation suivante :

$$\forall A \subseteq \Theta, m^{\Omega \downarrow \Theta}(A) = \sum_{B \subseteq \Omega / \rho^{-1}(B) = A} m^\Omega(B), \quad (3.40)$$

où ρ^{-1} est l'opération inverse du raffinement. Mais, sa définition n'est pas triviale puisqu'il existe des sous-ensembles $B \subseteq \Omega$ tels que $\forall A \subseteq \Theta, \rho(A) \neq B$. Des solutions ont été proposées dans la littérature [60] afin de compléter la définition de ρ^{-1} sur ces sous-ensembles B à savoir, la réduction intérieure $\underline{\theta}$ et la réduction extérieure $\bar{\theta}$ définies par :

$$\forall B \subseteq \Omega, \underline{\theta}(B) = \{\theta \in \Theta / \rho(\theta) \subseteq B\}. \quad (3.41)$$

$$\forall B \subseteq \Omega, \bar{\theta}(B) = \{\theta \in \Theta / \rho(\theta) \cap B \neq \emptyset\}. \quad (3.42)$$

Notons également que généralement, pour un sous-ensemble arbitraire B de Ω , $\rho(\underline{\theta}(B)) \neq B$ et $\rho(\bar{\theta}(B)) \neq B$. Ainsi, il y'a une perte d'information. En effet, B est approché par un

sous-ensemble ou un sur-ensemble [103].

L'exemple de la figure 3.1b montre que pour $B = \{\omega_4, \omega_5, \omega_6\} \subseteq \Omega$, $\underline{\theta}(B) = \{\theta_2\}$, $\bar{\theta}(B) = \{\theta_2, \theta_3\}$, $\rho(\theta_2) = \{\omega_4, \omega_5\}$, $\rho(\{\theta_2, \theta_3\}) = \{\omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9\}$. Nous vérifions alors que $\rho(\underline{\theta}(B)) \subseteq B \subseteq \rho(\bar{\theta}(B))$.

Parmi les deux types de réduction, nous retiendrons la moins engagée, $\rho^{-1} = \bar{\theta}$.

Définition 24. Extension vide

Dans ce qui suit, nous considérons un cadre de discernement produit

$\Theta \times \Omega = \{(\theta_i, \omega_j), \forall i \in \{1, \dots, |\Theta|\}, \forall j \in \{1, \dots, |\Omega|\}\}$ où $|\Theta|$ et $|\Omega|$ dénotent respectivement, le cardinal de Θ et le cardinal de Ω . Etant donnée une bba m^Θ définie dans Θ , l'extension vide est un cas particulier du raffinement. Elle permet de déduire la bba la moins engagée $m^{\Theta \uparrow \Theta \times \Omega}$ définie dans $\Theta \times \Omega$ selon la règle suivante :

$$m^{\Theta \uparrow \Theta \times \Omega}(B) = \begin{cases} m^\Theta(A) & \text{si } \exists A \subseteq \Theta \text{ tel que } B = (A, \Omega) \\ 0 & \text{sinon.} \end{cases} \quad (3.43)$$

Définition 25. Marginalisation

Opération duale de l'extension vide et un cas particulier du grossissement. Considérons une bba $m^{\Theta \times \Omega}$ définie dans $\Theta \times \Omega$, $m^{\Theta \times \Omega \downarrow \Theta}$ est le résultat de la marginalisation de $m^{\Theta \times \Omega}$ sur Θ . Elle se calcule comme suit :

$$\forall A \subseteq \Theta, m^{\Theta \times \Omega \downarrow \Theta}(A) = \sum_{B \subseteq \Theta \times \Omega / B^{\downarrow \Theta} = A} m^{\Theta \times \Omega}(B), \quad (3.44)$$

où $B^{\downarrow \Theta}$ dénote la projection de B sur Θ donnée par :

$$B^{\downarrow \Theta} = \{\theta \in \Theta / \exists \omega \in \Omega, (\theta, \omega) \in B\} \quad (3.45)$$

Notons également qu'il n'est généralement pas possible de retrouver la bba initiale après marginalisation. L'extension vide et la marginalisation sont illustrées dans la figure 3.2.

3.5.9.1 Première application pratique : combinaison de bbas définies dans des cadres de discernement indépendants

L'une des utilisations pratiques de l'extension vide est la combinaison de bbas m_1^Θ et m_2^Ω définies dans des cadres de discernement indépendants. Ainsi, la combinaison conjonctive

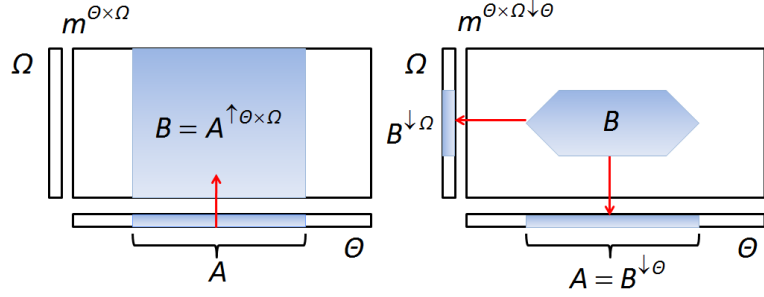


FIGURE 3.2: Extension vide et marginalisation

de leurs bbas étendues dans le cadre de discernement commun $\Theta \times \Omega$ est donnée par :

$$\begin{aligned}
 m_{1 \oplus 2}^{\Theta \times \Omega}(B) &= (m_1^{\Theta \uparrow \Theta \times \Omega} \oplus m_2^{\Omega \uparrow \Theta \times \Omega})(B) \\
 &= \begin{cases} m_1^\Theta(A) \cdot m_2^\Omega(C) & \text{si } B = (A, C), A \subseteq \Theta, C \subseteq \Omega. \\ 0 & \text{sinon.} \end{cases} \quad (3.46)
 \end{aligned}$$

3.5.9.2 Deuxième application pratique : conditionnement

Soit $m^{\Theta \times \Omega}$ une bba définie dans le produit cartésien $\Theta \times \Omega$ et soit θ_0 un sous-ensemble de Θ . La bba conditionnelle $m^\Omega[\theta_0]$ est obtenue en combinant conjonctivement la bba $m^{\Theta \times \Omega}$ avec la bba $m_{\theta_0}^{\Theta \uparrow \Theta \times \Omega}$ définie par $m^{\Theta \times \Omega}(\{\theta_0, \Omega\}) = 1$, et en marginalisant le résultat sur Ω .

$$m^\Omega[\theta_0] = \left(m^{\Theta \times \Omega} \oplus m_{\theta_0}^{\Theta \uparrow \Theta \times \Omega} \right)^{\downarrow \Omega}. \quad (3.47)$$

3.6 Niveau Décisionnel

A l'issue des fonctions de croyance établies dans le niveau crédal, une décision en faveur d'une hypothèse (nous considérons ici uniquement le cas des hypothèses simples comme structure décisionnelle) du cadre de discernement, doit être prise. Dans ce qui suit, nous mettons l'accent sur certains critères de décision dont le plus répandu est le maximum de probabilité pignistique.

3.6.1 Maximum de probabilité pignistique

D'une façon générale, prendre une décision sur un cadre de discernement Ω revient à choisir la décision d , parmi un ensemble de décisions D , qui minimise le *risque espéré* [104–106]. Ce risque noté γ utilise une mesure de probabilité $P^\Omega : \Omega \rightarrow [0, 1]$ telle que $\sum_{\omega \in \Omega} P^\Omega(\omega) = 1$ et une fonction coût $C : D \times \Omega \rightarrow \mathbb{R}$, où $C(d, \omega)$ dénote le coût de décider d sachant que la vérité est ω :

$$\gamma(d) = \sum_{\omega \in \Omega} P^\Omega(\{\omega\})C(d, \omega). \quad (3.48)$$

A ce niveau, une fonction de masse m^Ω issue du niveau crédal doit être transformée en une mesure de probabilité. Une solution est de calculer la probabilité pignistique (du latin *pignus* signifiant *un pari*) [95, 107–109] notée $BetP^\Omega$ définie par :

$$\forall \omega \in \Omega, BetP^\Omega(\{\omega\}) = \frac{1}{1 - m^\Omega(\emptyset)} \sum_{\forall A \subseteq \Omega / \omega \in A} \frac{m^\Omega(A)}{|A|}. \quad (3.49)$$

Ainsi, la décision choisie est celle qui minimise le risque pignistique donné par :

$$\gamma_{BetP}(d) = \sum_{\omega \in \Omega} BetP^\Omega(\{\omega\})C(d, \omega). \quad (3.50)$$

Notons que, selon l'équation (3.49), le $BetP$ distribue la masse sur un élément focal $A \subseteq \Omega$ de manière égale sur ses hypothèses singletons ce qui équivaut à l'hypothèse d'équiprobabilité en théorie des probabilités. Dans le cas où les coûts sont 0-1 c'est-à-dire pour $D = \Omega = \{w_1, \dots, w_N\}$, $C(w_i, w_j) = 1$ si $i = j$ et 0 sinon, la décision d qui minimise le risque pignistique (cf. Equation (3.50)) est celle qui maximise la probabilité pignistique (cf. Equation (3.51)).

$$Dec_{BetP} = \arg \max_{\omega \in \Omega} BetP(\omega). \quad (3.51)$$

Si en plus, nous considérons une décision de rejet d_0 ayant un coût de rejet $c_r = C(d_0, w_i), \forall i \in \{1, \dots, N\}$, minimiser le risque pignistique (cf. Equation (3.50)) revient alors à choisir la décision :

- (i) d_0 si $\max_{i=1, \dots, N} BetP(\{\omega_i\}) < 1 - c_r$;
- (ii) ω_j si $BetP(\omega_j) = \max_{i=1, \dots, N} BetP(\{\omega_i\}) \geq 1 - c_r$.

3.6.2 Maximum de plausibilité / Maximum de crédibilité

Alternativement, si nous adoptons une attitude optimiste ou pessimiste, nous utilisons respectivement le maximum de plausibilité (cf. Equation (3.52)) ou le maximum de crédibilité (cf. Equation (3.53)) si la décision ne peut être prise qu'en faveur des hypothèses singletons de Ω (car la croissance des fonctions Pl et Bel favorise les hypothèses composées [110]).

$$Dec_{Pl} = \arg \max_{\omega \in \Omega} Pl(\omega). \quad (3.52)$$

$$Dec_{Bel} = \arg \max_{\omega \in \Omega} Bel(\omega). \quad (3.53)$$

3.7 Conflit

Dans le cadre des fonctions de croyance, le conflit a été introduit pour quantifier le désaccord entre bbas. Comme évoqué précédemment (cf. Section 3.5.7), le conflit Dempsterien n'est autre que la masse sur le vide $m(\emptyset)$ issue de la combinaison conjonctive de bbas. Pour deux bbas m_1 et m_2 définies sur Ω ,

$$m(\emptyset) = \sum_{(A,B) \in 2^\Omega \times 2^\Omega / A \cap B = \emptyset} m_1(A) m_2(B). \quad (3.54)$$

Une forte valeur de $m(\emptyset)$ reflète un grand conflit entre sources et une décision non consensuelle. Ainsi, l'utilité du conflit apparaît dans l'estimation du désaccord entre sources s'exprimant sur la même observation, regroupement de sources selon leur consensus [111–114], détection d'erreurs de modélisation ou encore la définition de facteurs d'affaiblissement [115, 116], [84, 117, 118]. Il a même été utilisé pour définir de nouvelles mesures entre bbas telles que la distance introduite par Ristic et Smets [119]. Pour une étude profonde sur le conflit, nous référons le lecteur à [120].

3.8 Conclusion

Au cours de ce chapitre bibliographique, nous avons présenté brièvement la théorie des fonctions de croyance. Notre but n'était pas de fournir une description détaillée et exhaustive de la théorie mais d'exposer ses principes fondamentaux ainsi que les outils de base qui nous ont permis d'étudier et développer notre application. Deux approches ont été alors proposées dans ce cadre théorique. La première concerne la construction dynamique des objets présents dans une scène imagée. Elle fera l'objet du chapitre suivant. La deuxième approche détaillée dans le chapitre 5 estime l'ensemble des objets d'intérêt.

Chapitre 4

Approche crédibiliste pour la construction dynamique d'objets

4.1 Introduction

En vidéosurveillance, plusieurs applications visent à détecter les objets apparaissant dans la scène d'intérêt. Faisant l'hypothèse que ces objets sont de type quelconque (piétons, voitures, animaux...), nous avons souvent recours à des algorithmes de détection de changement (cf. Section A.3) pour détecter les objets ou des sous parties d'objets. Ces dernières doivent être rassemblées pour former ainsi les objets. Nous voulons dire par 'construction d'objets' l'étape qui permet de rassembler les fragments (les détections élémentaires de changements), les associer à l'objet adéquat et les fusionner avec les autres fragments. Cette étape est complexe pour deux raisons. La première est la détection des objets sous forme de fragments non nécessairement connexes. La deuxième raison est le déplacement des objets et leurs croisements potentiels (cf. Figure 4.1).

Dans ce travail, nous modélisons l'incertitude et l'imprécision de la localisation des fragments détectés en utilisant la théorie des fonctions de croyance. Nous illustrons l'adéquation de ce cadre théorique avec l'implémentation des deux mécanismes de construction d'objets, qui entrent en compétition : (i) l'accumulation de fragments ; (ii) la disparition des fragments. Ce formalisme s'avère aussi être bien adapté à l'association de données, notamment pour associer les nouvelles détections non labellisées avec les objets en cours

de construction. Des tests sur des données réelles nous ont permis d'évaluer qualitativement et quantitativement la méthode proposée en termes de robustesse aux occultations partielles et aux croisements d'objets.

La suite de ce chapitre est organisée comme suit : d'abord, nous définissons le problème traité, puis, nous exposons notre représentation crédibiliste d'objets, ensuite, nous présentons respectivement des généralités sur l'association de données et nous rappelons l'état de l'art des approches d'association proposées dans le cadre de la théorie des fonctions de croyance, enfin, nous proposons notre contribution sur l'association de données ainsi que l'approche de construction dynamique d'objets que nous comparons avec des approches alternatives. La conclusion sera présentée à la fin.

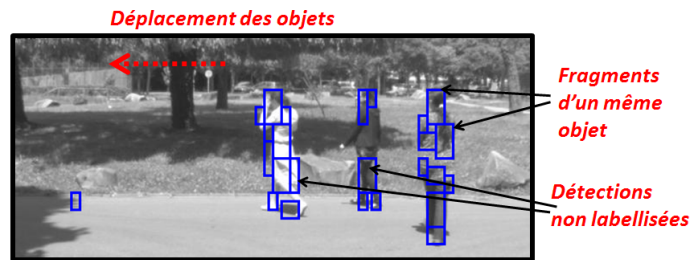


FIGURE 4.1: Illustration des verrous scientifiques au niveau de la construction d'objets : (i) déplacement d'objets ; (ii) fragmentation d'objets ; (iii) détections non labellisées.

4.2 Définition du problème

Le suivi est un problème connu en vision par ordinateur et concerne plus spécifiquement les applications de vidéosurveillance. Il a pour but l'estimation automatique de la trajectoire de chaque objet en mouvement (cf. Chapitre 2). Or, dans une scène typique, il peut y avoir plusieurs objets d'intérêt qui apparaissent et disparaissent de la scène durant le temps d'acquisition. Ainsi, les principaux défis pour le suivi multi-objets sont la fusion, la scission et le croisement des objets. Pour faire face à ces problèmes, les principaux travaux en *tracking* ont mis l'accent sur les caractéristiques (e.g. couleur, texture, flot optique, bord) les plus adéquates pour le suivi d'objets (cf. Chapitre 2). Ainsi, ils s'intéressent plutôt à la représentation adéquate d'objets (centroïde, ensemble de points, formes géométriques, contour, etc.). Cependant, la fiabilité du suivi ne dépend pas que de la représentation à partir de laquelle des caractéristiques discriminantes sont extraites mais aussi de la détection d'objets qui influe directement sur la fiabilité et la précision de ces caractéristiques. Dans ce contexte, nous distinguons deux cas : (i) cas d'un objet

dont les caractéristiques déjà connues (y compris l'utilisation d'un modèle de prédiction) sont exploitées pour sa détection ; (ii) cas d'un objet inconnu qui apparaît dans la scène. Dans ce dernier cas et en absence de tout a priori ou modèle sur les objets (nous n'évoquons pas les applications spécifiques à une classe d'objets, telles que le comptage se fondant sur la détection de sous parties d'objets de type humain (e.g. tête)), la détection est faite par contradiction avec un modèle de fond. Ainsi, pour des objets de taille non élémentaire (au contraire des applications radar où une cible (avion, bateau, intrus..) est généralement représentée par un point), les détections correspondent le plus souvent à des fragments des objets physiques. Cet aspect fragmentaire est principalement dû à des phénomènes d'occultation. Nous prenons comme exemple les techniques de soustraction de fond ou le modèle de mélanges gaussiens qui ne parviennent à éviter l'auto-occultation partielle d'un objet ou son occultation par le fond (camouflage partiel suite à la ressemblance de certaines sous parties de l'objet au fond). Or, un objet ayant des sous parties non détectées apparaît divisé en plusieurs fragments. Afin de rassembler les fragments et former les objets, les approches les plus simples sont issues d'étiquetage en composantes connexes ou considèrent plus simplement une boîte englobante. Cependant, aucune de ces techniques n'est robuste en présence d'objets multiples et proches. Ainsi, l'idée est d'accumuler les détections fragmentaires sur plusieurs images (d'une séquence vidéo) pour détecter entièrement les objets. Toutefois, plus la période d'accumulation est longue, plus deux objets différents sont susceptibles d'être fusionnés. Au lieu d'utiliser un critère ad-hoc pour rassembler les fragments détectés entre eux (e.g. seuillage d'une distance), les fragments sont utilisés pour fournir notre représentation crédibiliste pour les objets en construction. Notre problème est alors équivalent à un problème de fusion de données (fragments).

Par définition, les objets en construction n'ont pas de caractéristiques spécifiques autres que la localisation de leurs fragments. Par conséquent, nous proposons d'utiliser la localisation des détections pour la construction des objets mais de manière beaucoup plus nuancée qu'en effectuant des tests binaires sur des ensembles de pixels. Cette construction est établie prenant en considération la nature incomplète et imprécise des détections. Le cadre théorique de notre modélisation est la théorie des fonctions de croyance (cf. Chapitre 3) qui gère à la fois l'incertitude et l'imprécision.

- Notation

Dans le reste du manuscrit, l'association entre deux éléments o et d sera représentée par

la relation $\equiv$. Ainsi, $d \equiv o$ signifie que d est associé avec o ou encore d et o représentent le même élément. Par ailleurs, $\equiv$ est supposée transitive. Ainsi, pour trois éléments o_1 , o_2 et d : $o_1 \equiv d \ \& \ o_2 \equiv d \Rightarrow o_1 \equiv o_2$.

4.3 Représentation crédibiliste des objets [1, 2]

Rappelons brièvement notre problème ici : les détections étant fragmentaires, à chaque instant, les objets physiques (objets en mouvement dans la scène) sont seulement partiellement détectés en général et ne peuvent donc pas être directement énumérés. Nous cherchons alors à les construire au cours du temps en accumulant les fragments. En cas d'objets multiples, ce problème n'est pas trivial. D'une part, les détections sont non labellisées et d'autre part les objets peuvent se déplacer durant la période d'accumulation de fragments. Ainsi, nous proposons une représentation crédibiliste des objets qui tient compte de l'incertitude et l'imprécision de la localisation des détections, ce qui est la première contribution de ce chapitre.

Dans ce qui suit, le mot *objet* désignera soit un objet physique soit son approximation, à un instant donné, par l'ensemble des détections qui lui sont associées (c'est-à-dire ayant le même label).

Les observations dont nous disposons sont les détections fragmentaires. Ce sont donc ces dernières qui caractérisent les objets. Par ailleurs à ce stade, l'information directement disponible à partir de ces détections est leur localisation dans l'image (coordonnées spatiales en ligne et colonne). Nous réservons d'autres informations indirectes (dédites de traitements d'analyse subséquents) comme la radiométrie, la texture, etc. pour le *tracking* à proprement parler qui interviendra dans une étape ultérieure. Ce découpage est celui proposé dans l'introduction (cf. Chapitre 1). En résumé, ici un objet est perçu à travers des détections plus au moins fragmentaires de ses sous parties, elles mêmes caractérisées en terme de localisation spatiale dans l'image. La représentation crédibiliste des objets formalise alors la connaissance issue de cette perception des objets : les objets sont alors représentés en terme de la localisation imprécise d'une détection. Le cadre de discernement est la grille de l'image : pour une image de n_l lignes et n_c colonnes, Ω est l'espace produit $\{1, \dots, n_l\} \times \{1, \dots, n_c\}$.

Une bba qui correspond à un objet donné représente la croyance en la localisation d'une détection (un fragment) de cet objet. Par ailleurs, notons que les détections fournies par

l'algorithme de TI à un instant donné sont distinctes ce qui permet de satisfaire l'exclusivité des hypothèses du cadre de discernement. Les localisations des détections sont indexées par exemple par les coordonnées du coin gauche le plus haut.

Rappelons aussi, qu'une détection d_j peut être à l'origine d'un nouvel objet. A l'initialisation, sa représentation crédibiliste est tout simplement une bba catégorique $m\{d_j\}$ (cf. Définition 3) ayant comme seul élément focal $A(d_j)$ l'ensemble des pixels appartenant à d_j .

Lors de la construction d'une bba $m\{o_i\}$ correspondant à un objet o_i à partir des bbas des détections associées, deux mécanismes rentrent en compétition :

- le premier vise à construire spatialement l'objet à partir de ses détections ce qui permet d'étendre l'ensemble des localisations possibles des futures détections. Ce mécanisme est instancié via la règle disjonctive (cf. Equation (3.28)) pour combiner la bba d'un objet avec les bbas des détections qui lui sont associées.
- le deuxième vise, dans le cas d'une séquence vidéo et donc des détections indexées en temps, à considérer le déplacement possible d'un objet susceptible d'invalidiser certaines localisations possibles des futures détections. Ce mécanisme est instancié via les opérateurs proposés à savoir le conditionnement temporel et le conditionnement spatial.

4.3.1 Combinaison disjonctive

Selon le premier mécanisme, la bba $m\{o_i\}$ d'un objet o_i est actualisée lorsque de nouvelles sous parties de o_i sont détectées. Ces dernières ont la forme soit d'une nouvelle détection(s) ou d'un autre objet en cours de construction, reconnu comme étant une sous partie de o_i . Dans les deux cas, l'information issue d'une nouvelle sous partie de o_i est représentée par une bba. Spécifiquement, pour une détection d_j , la bba $m\{d_j\}$ est catégorique (cf. Définition 3) de seul élément focal $A(d_j)$, ensemble des pixels de d_j .

La bba actualisée de o_i est obtenue en combinant disjonctivement (cf. Equation (3.28)) $m\{o_i\}$ avec $m\{o_j\}$ ou $m\{d_j\}$.

$$m\{o_i\} \leftarrow m\{o_i\} \odot m\{d_j\}. \quad (4.1)$$

Notons également que si plusieurs sous parties de o_i sont détectées, soit simultanément soit itérativement comme nous expliquons plus tard décrivant l'étape d'association de données (cf. Section 4.6), la règle de combinaison (cf. Equation (4.1)) peut être itérée.

4.3.2 Conditionnement temporel

Le deuxième mécanisme vise à prendre en compte la dimension temporelle des détections. Pour tenir compte du déplacement éventuel des objets et/ou de leur disparition de la scène, les détections les plus récentes doivent avoir plus de poids dans la construction des bbas des objets.

Rappelons que dans cette étude nous nous concentrons sur la phase de construction des objets à partir des détections associées. Aussi acceptons-nous de détruire éventuellement des informations issues d'une construction antérieure des objets. Pratiquement, nous supposons que seuls les Δ_t derniers instants sont à considérer pour l'estimation de la localisation imprécise d'une détection. Cette première hypothèse se traduit alors par un conditionnement des bbas des objets sur la disjonction des détections sur les Δt derniers instants. Soit m_{Δ_t} la bba catégorique (cf. Définition 3) représentant ce conditionnement temporel.

Nous supposons que des détections ont d'autant plus de chance d'être proches spatialement qu'elles le sont temporellement. Ainsi, la croyance en les localisations proches des dernières observations de l'objet doit être renforcée. Cette seconde hypothèse se traduit par une combinaison conjonctive (cf. Equation (3.26)) entre la bba de l'objet (après conditionnement) et une bba simple (cf. Définition 1) ayant comme élément focal hors Ω la disjonction des détections associées à l'objet à $t - 1$, et de masse sur Ω égale à α , plus α est grand, plus les anciennes détections sont crédibles en terme de localisation des futures détections. Notons $m_{d_{o_i}, t-1}$ la bba catégorique d'élément focal la disjonction des détections associées à o_i à $t - 1$ et $m_{d_{o_i}, t-1}^\alpha$ le résultat de son affaiblissement (cf. Equation (3.15)) par α . Ainsi, la bba $m\{o_i\}$ d'un objet o_i est actualisée comme suit :

$$m\{o_i\} \leftarrow m\{o_i\} \odot m_{d_{o_i}, t-1}^\alpha \odot m_{\Delta_t}. \quad (4.2)$$

4.3.3 Conditionnement spatial

Schématiquement, nous pouvons dire que le conditionnement temporel présenté précédemment permettait de gérer les objets disparus (élimination des bba des objets non détectés durant la période de temps Δt). Le but du conditionnement spatial présenté à présent est la gestion des séparations d'objets. L'exemple typique est le cas de la séparation d'un groupe de personnes partant chacune de leur côté. Dans notre cas, ce conditionnement spatial nous permettra également de séparer deux objets malencontreusement fusionnés lors de leur croisement. L'hypothèse utilisée est celle qu'un objet est mono-composante connexe et que les composantes connexes autres que celle principale sont du bruit qui doit être éliminé. Cette première hypothèse se traduit par un conditionnement (cf. Equation (3.18)) sur la principale composante connexe de l'objet.

En pratique, les composantes connexes de l'objet ne sont pas estimées sur la grille de l'image mais après projection sur l'axe des colonnes. Ceci est issu de l'hypothèse selon laquelle les principaux objets de la scène présentent une symétrie par rapport à l'axe des colonnes et donc la fragmentation des objets se fait principalement selon l'axe des lignes. Cette hypothèse est vraie notamment pour des objets de type 'humain' imagés dans un plan environ vertical. Notons qu'un tel a priori peut être modifié selon la géométrie d'acquisition et des objets. Soit A l'élément focal correspondant à la disjonction des détections associées à l'objet, $A^{\downarrow X}$ la projection de A sur l'axe des colonnes, et $C^{\downarrow X}$ la plus grande des composantes connexes de $A^{\downarrow X}$. Le conditionnement spatial se fait sur la disjonction des détections associées à l'objet telles que leur projection sur l'axe des colonnes soit dans $C^{\downarrow X}$. Notons qu'avec ce conditionnement nous séparerons d'autant mieux les objets qu'ils se déplacent différemment selon l'axe des colonnes.

En conclusion, les opérateurs proposés de conditionnements temporel et spatial permettent de gérer le déplacement d'objets et leur changement durant la séquence vidéo (notamment en cas de séparation des objets selon l'axe des colonnes).

D'un point de vue pratique, les deux conditionnements sont effectués par combinaison de la bba d'un objet avec une bba catégorique d'élément focal estimé selon la consistance spatiale et temporelle des objets. Ces conditionnements sont intrinsèquement une renormalisation de la croyance sur les Δt derniers instants et sur la principale composante connexe de l'objet en cours de construction. Aussi, la règle de combinaison utilisée est-elle la règle orthogonale (cf. Equation (3.24)).

4.4 Généralités sur l'association de données

Soient deux ensembles $\mathcal{E}_1 = \{o_1, \dots, o_n\}$ et $\mathcal{E}_2 = \{d_1, \dots, d_l\}$. Nous avons $n = |\mathcal{E}_1|$ et $l = |\mathcal{E}_2|$ où $|\mathcal{E}|$ est le cardinal de l'ensemble $\mathcal{E}$. Le problème d'association est équivalent à un problème d'affectation de graphe biparti $G = (\mathcal{E}, \mathcal{A}_m)$. Dans cette notation, $\mathcal{E}$ est l'ensemble des sommets du graphe répartis sur les ensembles exclusifs $\mathcal{E}_1$ et $\mathcal{E}_2$ ($\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_2$; $\mathcal{E}_1 \cap \mathcal{E}_2 = \emptyset$) et $\mathcal{A}_m$ dénote une association c'est-à-dire une mise en correspondance entre $\mathcal{E}_1$ et $\mathcal{E}_2$. Elle est décrite par un ensemble d'arêtes $\{a_1, \dots, a_q\} \subseteq \mathcal{E}_1 \times \mathcal{E}_2$ de la forme $a_k = (o_i, d_j)$ liant un élément $o_i \in \mathcal{E}_1$ avec un élément $d_j \in \mathcal{E}_2$. A chaque arête (o_i, d_j) est attribué un coût local noté $c(o_i, d_j)$.

4.4.1 Fonction coût

Une fonction coût C est une fonction définie sur un ensemble $\mathcal{A}$ d'associations (ou associations acceptables en présence d'éventuelles contraintes) et à valeurs dans l'ensemble $\mathbb{R}$ des nombres réels. Elle permet de résoudre un problème d'optimisation dont le but est de trouver l'association optimale (si elle existe) $\mathcal{A}^*$ de $\mathcal{A}$ telle que $C(\mathcal{A}^*) \leq C(\mathcal{A}_m)$ pour toutes les associations $\mathcal{A}_m$ dans $\mathcal{A}$ (cf. Equation (4.3)).

$$\mathcal{A}^* = \min_{\mathcal{A}_m \in \mathcal{A}} (C(\mathcal{A}_m)). \quad (4.3)$$

Un exemple classique illustrant la notion de coût est l'optimisation d'une répartition de tâches : soient un ensemble $\mathcal{E}_1$ d'employés, un ensemble $\mathcal{E}_2$ de tâches et les contraintes d'association suivantes : (i) chaque tâche doit être affectée à un employé ; (ii) chaque employé ne doit pas effectuer plus qu'une tâche. $c(o_i, d_j)$ est le coût local d'affecter la tâche d_j à l'employé o_i . Il est d'autant plus faible que la tâche est adéquate avec l'employé. Notons également qu'en pratique, un coût est généralement additif (cf. Annexe A, e.g. 1) c'est-à-dire le coût global est la somme des coûts locaux c positifs (cf. Equation (4.4)).

$$C(\mathcal{A}_m) = \sum_{a_k \in \mathcal{A}_m, a_k \geq 0} c(a_k). \quad (4.4)$$

Par ailleurs, selon les contraintes d'association (cf. Annexe A, Exemple 2), des éléments de $\mathcal{E}_1$ et $\mathcal{E}_2$ peuvent être non associés c'est-à-dire n'appartenant à aucune arête $a_k \in \mathcal{A}_m$. Nous leur attribuons alors des coûts locaux de non association respectivement égaux à

c_1 et c_2 et dans ce cas :

$$C(\mathcal{A}_m) = \sum_{a_k \in \mathcal{A}_m} c(a_k) + n_1 c_1 + n_2 c_2, \quad (4.5)$$

où n_1 et n_2 dénotent les nombres d'éléments non associés respectivement de $\mathcal{E}_1$ et $\mathcal{E}_2$.

$$\begin{aligned} n_1 &= |o_i| \text{ tel que } \nexists a_k = (o_i, d_j) \in \mathcal{A}_m, \\ n_2 &= |d_j| \text{ tel que } \nexists a_k = (o_i, d_j) \in \mathcal{A}_m. \end{aligned} \quad (4.6)$$

Cette propriété de coût additif facilite la recherche du minimum (association optimale) ce qui est essentiel pour certaines applications. En effet, effectuer une recherche exhaustive en testant toutes les associations possibles est impossible en pratique lorsque le nombre d'éléments à associer devient important. Par exemple, en absence de contraintes, la recherche exhaustive devrait examiner $2^{n \times l}$ associations.

4.4.2 Types d'association

Les contraintes d'association se justifient par la nature du problème traité et ses hypothèses.

4.4.2.1 Type 1-1

La contrainte 1-1 signifie qu'un élément de $\mathcal{E}_1$ (respectivement de $\mathcal{E}_2$) est associé avec au maximum un élément de $\mathcal{E}_2$ (respectivement de $\mathcal{E}_1$). Cette contrainte très usuelle, réduit considérablement la taille de l'espace de solutions. En plus, pour des coûts additifs, elle permet la mise en œuvre d'algorithmes efficaces d'optimisation tels que les variantes de l'algorithme Hongrois [49], [121]. En cas de situation ambiguë, plusieurs solutions d'association peuvent éventuellement être gardées provisoirement selon une approche de type MHT pour laquelle l'algorithme de Murty [54] permet de trouver les associations de coûts minima.

– Algorithme Hongrois

Soit M_c une matrice de coûts définie par :

$$M_c(i, j) = c(o_i, d_j), \forall (i, j) \in \{1, \dots, n\} \times \{1, \dots, l\}, \quad (4.7)$$

où nous supposons que $l = n$ c'est-à-dire $|\mathcal{E}_2| = |\mathcal{E}_1|$.

Une association bijective $\mathcal{A}_m$ entre $\mathcal{E}_1$ et $\mathcal{E}_2$ est représentée par un vecteur $\mu = [\nu_1 \dots \nu_n]$ de taille n qui donne pour chaque élément d'indice $i \mid i \in \{1 \dots n\}$ l'indice de l'élément associé : $\nu_i \mid \nu_i \in \{1 \dots n\}$. Ainsi, le problème d'association se ramène à la recherche de la permutation μ^* , qui définit l'association optimale $\mathcal{A}^*$, telle que le coût total $\sum_{i=1}^n c(o_i, d_{\nu_i})$ soit minimum sur l'ensemble de toutes les permutations possibles.

L'algorithme Hongrois (ou méthode Hongroise) est un algorithme d'optimisation combinatoire permettant de résoudre ce problème d'association bijective en temps polynomial. Pour une description détaillée de l'algorithme, nous nous référons à l'annexe A.1.

4.4.2.2 Type 1-N

Dans certaines applications pratiques, nous devons relâcher la contrainte 1-1 et autoriser les associations de type M-N. Dans ce cas, une technique de type *graph cut* [122] permet la recherche d'un minimum. Cependant, ce type d'association peut induire des solutions dégénérées telles que l'association de tous les éléments de $\mathcal{E}_1$ avec tous les éléments de $\mathcal{E}_2$. C'est la raison pour laquelle, nous utilisons plutôt une contrainte de type 1-N (ou M-1) imposant à chaque association de ne pas contenir à la fois plusieurs éléments de $\mathcal{E}_1$ et plusieurs éléments de $\mathcal{E}_2$ (cf. Figure 4.2). Notons également que si le coût est

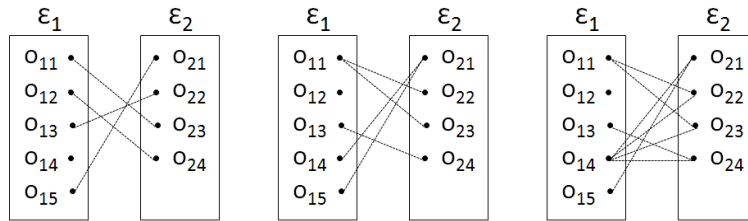


FIGURE 4.2: Associations 1-1, 1-N, M-N.

additif et que les coûts élémentaires sont positifs, alors l'association de coût minimum sera nécessairement de type 1-N. En effet, lorsqu'une arête $a_k = (o_i, d_j)$ appartient à l'association optimale $\mathcal{A}^*$, alors les deux arêtes $a_p = (o_i, d_p)$ et $a_r = (o_r, d_j)$ (avec d_p et o_r quelconques) ne peuvent pas appartenir simultanément à $\mathcal{A}^*$.

Propriété 1. [123] Soit $\mathcal{A} = \{\mathcal{A}_1, \dots, \mathcal{A}_s\}$ l'ensemble des solutions d'association de la forme $\mathcal{A}_m = \{a_1, \dots, a_q\}$ où chaque arête $a_k = (o_i, d_j)$ représente l'association entre un élément $o_i \in \mathcal{E}_1$ et un élément $d_j \in \mathcal{E}_2$. Soit $\mathcal{A}^*$ la solution d'association qui minimise

une fonction coût donnée par les équations (4.5), (4.6) alors :

$$a_k = (o_i, d_j) \in \mathcal{A}^* \Rightarrow \nexists (a_p, a_r) \in \mathcal{A}^{*2} \text{ tel que } a_p = (o_i, d_p) \text{ et } a_r = (o_r, d_j). \quad (4.8)$$

Preuve 1. En raisonnant par l'absurde, supposons que l'association $\mathcal{A}_1$ ne respecte pas cette propriété. Elle comporte dans ce cas trois arêtes de la forme : $a_k = (o_i, d_j)$, $a_p = (o_i, d_p)$ et $a_r = (o_r, d_j)$. Ainsi, l'association $\mathcal{A}_2$ formée par la suppression de l'arête a_k ne contient aucun objet isolé ni de $\mathcal{E}_1$ ni de $\mathcal{E}_2$. Par conséquent, son coût vaut $C(\mathcal{A}_2) = C(\mathcal{A}_1) - c(a_k)$ et est inférieur au coût de $\mathcal{A}_1$ puisque $c(a_k) \geq 0$. Autrement dit, si une association ne respecte pas la propriété, il sera toujours possible de supprimer une arête sans isoler des éléments. Le coût de la nouvelle association ainsi créée sera nécessairement inférieur.

Afin d'éviter la mise en place d'une méthode *graphs cuts* assez lourde, une solution approchée peut être estimée en commençant par rechercher la solution 1-1 et en tentant ensuite d'ajouter des arêtes pour lier les éléments isolés. Bien que cette approche empirique ne garantisse pas un coût total optimal, elle peut donner de bons résultats car elle favorise les associations 1-1.

4.5 Etat de l'art des méthodes d'association dans le cadre du TBM

Outre les approches probabilistes telles que le PDAF, JPDAF et le MHT (cf. Chapitre 2), plusieurs approches ont été proposées pour le problème d'association exploitant le cadre des fonctions de croyance, notamment [3, 6, 7, 124–129].

4.5.1 Approches fondées sur la définition de bbas d'associations

Les travaux [3, 4, 6, 7, 124, 130] sont des variantes fondées sur la même idée de définition de croyances associées directement à la pertinence d'une association.

Soient $\mathcal{E}_1 = \{o_1, \dots, o_n\}$ l'ensemble d'objets connus et $\mathcal{E}_2 = \{d_1, \dots, d_l\}$ l'ensemble des nouvelles détections. Les éléments de $\mathcal{E}_2$ sont à associer avec ceux de $\mathcal{E}_1$.

Deux cadres de discernement ont été proposés (cf. Equation (4.9)) : (i) Ω_2 formé par les éléments de $\mathcal{E}_1$, une hypothèse de Ω_2 est une réponse possible à la question : "A quel

objet de $\mathcal{E}_1$ est associée une détection de $\mathcal{E}_2$?" ; (ii) Ω_1 formé par les éléments de $\mathcal{E}_2$, une hypothèse de Ω_1 est une réponse possible à la question : "A quelle détection de $\mathcal{E}_2$ est associé un objet de $\mathcal{E}_1$?".

Pour gérer les problèmes d'apparition et disparition d'objets, une nouvelle hypothèse '*' est introduite dans Ω_1 et Ω_2 [130]. Si un objet de Ω_1 est associé avec l'hypothèse * de Ω_2 alors l'objet a disparu de la scène. Si une détection de Ω_2 est associée avec l'hypothèse * de Ω_1 alors la détection correspond à un nouvel objet apparu dans la scène.

$$\begin{aligned}\Omega_1 &= \{d_1 \dots d_j \dots d_l, *\}, \\ \Omega_2 &= \{o_1 \dots o_i \dots o_n, *\}.\end{aligned}\tag{4.9}$$

Ce monde ainsi proposé est appelé *monde ouvert étendu*. De ce fait, toute masse non nulle sur $\emptyset$ pouvant provenir d'une combinaison conjonctive de bbas définies sur Ω_1 (ou Ω_2) reflète un conflit entre sources et non pas un problème de non exhaustivité de Ω_1 (ou Ω_2).

– Bbas d'associations locales

Un cadre de discernement élémentaire $\Omega_{ij} = \{y_{ij}, n_{ij}\}$ est défini pour chaque association potentielle $(o_i, d_j)_{i \in \{1, \dots, n\}, j \in \{1, \dots, l\}}$. Une bba $m_{ij}^{\Omega_{ij}}$ définie dans ce cadre modélise une information concernant la réponse à la question : "oui (représentée par l'hypothèse y_{ij}) ou non (représentée par l'hypothèse n_{ij}) d_j est associée avec o_i ?". Ainsi, une masse non nulle sur Ω_{ij} modélise l'ignorance partielle concernant l'association.

$$\begin{aligned}m_{ij}^{\Omega_{ij}}(y_{ij}) &: \text{degré de croyance sur l'hypothèse 'd}_j \text{ est associée avec } o_i', \\ m_{ij}^{\Omega_{ij}}(n_{ij}) &: \text{degré de croyance sur l'hypothèse 'd}_j \text{ n'est pas associée avec } o_i', \\ m_{ij}^{\Omega_{ij}}(\Omega_{ij}) &: \text{degré de croyance sur l'ignorance que 'd}_j \text{ est associée avec } o_i'.\end{aligned}\tag{4.10}$$

Notons également que m_{ij} est symétrique puisque le degré de croyance sur l'hypothèse ' d_j est associée avec o_i ' est égal au degré de croyance sur l'hypothèse ' o_i est associé avec d_j '. Elle est généralement déduite d'une mesure de similarité entre o_i et d_j .

– Combinaison de croyances sur une association locale

L'idée de définition de bbas d'association paraît intéressante si nous disposons de plusieurs informations hétérogènes. En effet, l'association peut se faire selon plusieurs critères dont la fusion permet d'enlever les ambiguïtés entre différentes associations possibles. Ces critères peuvent être des caractéristiques d'objets telles que la position, vitesse, classe, couleur, forme, texture etc. [7].

Soit k un critère. Nous désignons par $m_{k,ij}$ la bba qui représente la croyance sur l'association de d_j avec o_i selon le critère k :

$$\begin{aligned} m_{k,ij}^{\Omega_{ij}}(y_{ij}) &: \text{degré de croyance sur l'hypothèse '}'d_j \text{ est associée avec } o_i \text{ selon le critère } k', \\ m_{k,ij}^{\Omega_{ij}}(n_{ij}) &: \text{degré de croyance sur l'hypothèse '}'d_j \text{ n'est pas associée avec } o_i \text{ selon } k', \\ m_{k,ij}^{\Omega_{ij}}(\Omega_{ij}) &: \text{degré de croyance sur l'ignorance que '}'d_j \text{ soit associée avec } o_i \text{ selon } k'. \end{aligned} \quad (4.11)$$

Si nous disposons de L critères indépendants $\{1...k...L\}$ à partir desquels L bbas $\{m_{k,ij}\}_{k=1}^L$ sont définies pour un même couple (o_i, d_j) alors il est possible de combiner ces bbas par règle conjonctive (cf. Equation (4.12)) pour tirer une information à la fois plus complète et précise sur l'association locale entre o_i et d_j . Dans ce cas, la masse sur $\emptyset$ représente le conflit entre les L critères au sujet de l'association locale.

$$m_{ij} = \bigodot_{k=1}^L m_{k,ij}. \quad (4.12)$$

Pour se décider de l'association locale entre o_i et d_j , nous pouvons appliquer une fonction de décision ($BetP, Pl, Bel$) sur la bba m_{ij} (cf. Equation (4.12)) définissant la croyance. En absence d'ambiguïté, de telles décisions locales peuvent donner de bons résultats. Dans le cas contraire, une décision globale est nécessaire.

– Bbas d'associations globales

Une bba d'association globale $m_i^{\Omega_1}$ (respectivement $m_j^{\Omega_2}$) modélise la croyance sur l'association d'un objet $o_i \in \mathcal{E}_1$ (respectivement une détection $d_j \in \mathcal{E}_2$) avec les différentes détections de $\mathcal{E}_2 \cup \{*\}$ (respectivement $\mathcal{E}_1 \cup \{*\}$). Elle est définie dans Ω_1 (respectivement Ω_2) comme suit : chaque bba de l'ensemble $\{m_{ij}^{\Omega_{ij}}\}_{j=1}^l$ faisant intervenir l'objet o_i est redéfinie de Ω_{ij} dans Ω_1 par extension vide selon l'équation (4.13). La bba $m_i^{\Omega_1}$ est obtenue par combinaison conjonctive des l bbas $m_{ij}^{\Omega_1}$ (cf. Equation (4.14)).

$$\begin{aligned} m_{ij}^{\Omega_{ij} \uparrow \Omega_1}(d_j) &= m_{ij}(y_{ij}), \\ m_{ij}^{\Omega_{ij} \uparrow \Omega_1}(\bar{d}_j) &= m_{ij}(n_{ij}), \\ m_{ij}^{\Omega_{ij} \uparrow \Omega_1}(\Omega_1) &= m_{ij}(\Omega_{ij}). \end{aligned} \quad (4.13)$$

$$m_i^{\Omega_1} = \bigodot_{j=1}^l m_{ij}^{\Omega_{ij} \uparrow \Omega_1}. \quad (4.14)$$

De même, chaque bba de l'ensemble $\left\{m_{ij}^{\Omega_{ij}}\right\}_{i=1}^n$ relative à la détection d_j est redéfinie dans le cadre raffiné Ω_2 selon l'équation (4.15). Les bbas sont ensuite conjonctivement combinées selon l'équation (4.16) pour donner la bba $m_j^{\Omega_2}$.

$$\begin{aligned} m_{ij}^{\Omega_{ij} \uparrow \Omega_2}(o_i) &= m_{ij}(y_{ij}), \\ m_{ij}^{\Omega_{ij} \uparrow \Omega_2}(\bar{o}_i) &= m_{ij}(n_{ij}), \\ m_{ij}^{\Omega_{ij} \uparrow \Omega_2}(\Omega_2) &= m_{ij}(\Omega_{ij}). \end{aligned} \quad (4.15)$$

$$m_j^{\Omega_2} = \bigoplus_{i=1}^n m_{ij}^{\Omega_{ij} \uparrow \Omega_2}. \quad (4.16)$$

4.5.1.1 Approche de Gruyer et al : bbas d'association spécialisées [3–5]

Selon le contexte applicatif et la réalité physique modélisée, les bbas d'associations locales peuvent être définies selon le postulat suivant : *une source ne peut pas donner en même temps de la croyance sur l'association entre deux éléments o_i et d_j et sur la non association entre ces deux mêmes éléments.*

Ainsi, les éléments focaux d'une bba $m_{ij}^{\Omega_{ij}}$, dite spécialisée, sont $\{\Omega_{ij}, y_{ij}\}$ ou $\{\Omega_{ij}, n_{ij}\}$. Dans ce cas, la bba issue de la combinaison conjonctive des l bbas (respectivement n bbas) spécialisées concernant un même objet o_i (respectivement une même détection d_j) selon l'équation (4.14) (respectivement (4.16)) est définie selon l'équation (4.17). Selon cette bba, la masse est répartie sur les sous-ensembles de Ω_1 . Notons que la masse sur l'hypothèse $*$ provient de la combinaison des croyances sur les hypothèses (d_j n'est pas associée avec o_i , $\forall j \in \{1, \dots, l\}$). La masse sur $\emptyset$ est le conflit réel entre les sources élémentaires d'associations locales.

La bba donnée par l'équation (4.17) est à définir pour chaque objet $o_i \in \mathcal{E}_1$. Chaque bba $m_i^{\Omega_1}$ nous donne une distribution de croyance sur l'association multiple de type 1-N (1 objet de $\mathcal{E}_1$ - N objets de $\mathcal{E}_2$).

Définir une association multiple de type $N-M$ est possible si les bbas $m_i^{\Omega_1} \forall i \in \{1, \dots, n\}$, sont combinées dans un cadre discernement commun et plus général faisant abstraction de o_i .

Cependant, vu la complexité combinatoire, un raisonnement selon l'autre point de vue a été adopté permettant de calculer les bbas $m_j^{\Omega_2} \forall j \in \{1, \dots, l\}$. Chacune des bbas définit la croyance en l'association duale de type $N-1$ (N objets de $\mathcal{E}_1$ -1 détection de $\mathcal{E}_2$).

L'application d'une fonction de décision sur les bbas calculées selon les deux points de

vue permet de décider sur l'association en ayant une bonne vue de la situation.

$$\begin{aligned}
 m_i(d_j) &= m_{ij}(d_j) \prod_{a=1, a \neq j}^l (1 - m_{ia}(d_a)), \\
 m_i(\{d_j, *\}) &= m_{ij}(\Omega_1) \prod_{a=1, a \neq j}^l (m_{ia}(\bar{d}_a)), \\
 &\dots \\
 m_i(\{d_j, d_k \dots *\}) &= m_{ij}(\Omega_1) \cdot m_{ik}(\Omega_1) \dots \prod_{a=1, a \neq j, k, \dots}^l (m_{ia}(\bar{d}_a)), \\
 m_i(\bar{d}_j) &= m_{ij}(\bar{d}_j) \prod_{a=1, a \neq j}^l m_{ia}(\Omega_1), \\
 m_i(*) &= \prod_{a=1}^l (m_{ia}(\bar{d}_a)), \\
 m_i(\Omega_1) &= \prod_{a=1}^l (m_{ia}(\Omega_1)), \\
 m_i(\emptyset) &= 1 - \left[\prod_{a=1}^l (1 - m_{ia}(d_a)) + \sum_{a=1}^l m_{ia}(d_a) \prod_{b=1, b \neq a}^l (1 - m_{ib}(d_b)) \right].
 \end{aligned} \tag{4.17}$$

– Prise de décision sur l'association

Pour une association globale, dans [5], il a été proposé d'utiliser les deux tables de probabilités pignistiques (cf. Table 4.1 et Table 4.2) calculées selon chacun des points de vue (du côté de $\mathcal{E}_1$ et du côté de $\mathcal{E}_2$). Tout d'abord, pour chaque colonne, le maximum ($\max_A(BetP_i^{\Omega_1}(.)), A \in \Omega_1$ et $\max_B(BetP_j^{\Omega_2}(.)), B \in \Omega_2$) est calculé. Différents cas sont traités par la suite :

- si pour un ou plusieurs $o_i, A = \emptyset$ ou pour un ou plusieurs $d_j, B = \emptyset$ et si le conflit est assez important suite à un problème de modélisation ou à des critères d'association non discriminants, il est nécessaire de gérer le conflit avant de prendre une décision.
- si pour une valeur de $o_i, A = *$ ou pour une valeur de $d_j, B = *$ alors nous déduisons d'une part la perte de l'objet o_i et d'autre part l'apparition du nouvel objet d_j .
- il se peut que les deux matrices de probabilités pignistiques soient ambiguës ou contradictoires : l'ambiguïté apparaît lorsque dans la même matrice et sur la même colonne, deux probabilités ou plus sont égales, e.g. à partir de la table 4.1, $BetP_1^{\Omega_1}(d_1) = BetP_1^{\Omega_1}(d_l)$. Ainsi, il est difficile de décider laquelle des détections est associée avec o_1 . La contradiction apparaît lorsque les décisions à partir des tables 4.1 et 4.2 sont différentes, e.g. d'une part, dans la table 4.1, $BetP_1^{\Omega_1}(d_1)$ est maximale sur la première

colonne c'est-à-dire d_1 est associée avec o_1 , d'autre part, dans la table 4.2, $BetP_1^{\Omega_2}(o_n)$ est maximale sur la première colonne c'est-à-dire o_n est associé avec d_1 .

Dans de tels cas (ambiguïté, contradiction), [5] propose de construire une nouvelle table M où $M_{i,j} = \frac{BetP_i^{\Omega_1}(d_j) + BetP_j^{\Omega_2}(o_i)}{2}$.

 TABLE 4.1: Probabilités pignistiques calculées pour chaque objet de $\mathcal{E}_1$

$BetP_i^{\Omega_1}$	o_1	o_2	...	o_n
d_1	$BetP_1^{\Omega_1}(d_1)$	$BetP_2^{\Omega_1}(d_1)$	...	$BetP_n^{\Omega_1}(d_1)$
...	...	...	...	...
d_l	$BetP_1^{\Omega_1}(d_l)$	$BetP_2^{\Omega_1}(d_l)$	...	$BetP_n^{\Omega_1}(d_l)$
*	$BetP_1^{\Omega_1}(*)$	$BetP_2^{\Omega_1}(*)$	...	$BetP_n^{\Omega_1}(*)$

 TABLE 4.2: Probabilités pignistiques calculées pour chaque détection de $\mathcal{E}_2$

$BetP_j^{\Omega_2}$	d_1	d_2	...	d_l
o_1	$BetP_1^{\Omega_2}(o_1)$	$BetP_2^{\Omega_2}(o_1)$	...	$BetP_l^{\Omega_2}(o_1)$
...	...	...	...	...
o_n	$BetP_1^{\Omega_2}(o_n)$	$BetP_2^{\Omega_2}(o_n)$	...	$BetP_l^{\Omega_2}(o_n)$
*	$BetP_1^{\Omega_2}(*)$	$BetP_2^{\Omega_2}(*)$	...	$BetP_l^{\Omega_2}(*)$

TABLE 4.3: Matrice pignistique globale

$M_{i,j}$	d_1	d_2	...	d_l	$*_1$	...	$*_n$
o_1	$M_{1,1}$	$M_{1,2}$	...	$M_{1,l}$	$BetP_1^{\Omega_1}(*)$	...	0
...	...	...	...	...	$*_1$	...	$*_n$
o_n	$M_{n,1}$	$M_{n,2}$	...	$M_{n,l}$	0	...	$BetP_n^{\Omega_1}(*)$
$*_1$	$BetP_1^{\Omega_2}(*)$	0	...	0	0	...	0
...	...	...	...	...	0	...	0
$*_l$	0	0	...	$BetP_l^{\Omega_2}(*)$	0	...	0

Dans [3, 4], la matrice pignistique globale (cf. Table 4.3) est considérée comme une matrice de coûts. Faisant l'hypothèse qu'un objet de $\mathcal{E}_1$ peut être associé avec au plus une détection de $\mathcal{E}_2$ et qu'une détection de $\mathcal{E}_2$ peut être associée avec au plus un objet de $\mathcal{E}_1$, l'algorithme Hongrois a été utilisé pour avoir l'association globale optimale.

4.5.1.2 Approche de Mercier et al [6]

Nous présentons ici la méthode développée dans [6] qui se fonde sur les travaux [4, 5] mais se distingue par son étape décisionnelle. Dans [6], les contraintes d'association sont les suivantes : (i) chaque objet o_i de $\mathcal{E}_1$ est associé avec au plus une détection de $\mathcal{E}_2$; (ii) chaque détection d_j de $\mathcal{E}_2$ est associée avec au plus un objet de $\mathcal{E}_1$; (iii) l'élément $*$ peut être associé avec n'importe quel objet.

A partir des bbas $m_i^{\Omega_1}$ et $m_j^{\Omega_2}$ calculées respectivement selon l'équation (4.14) et l'équation (4.16), les probabilités pignistiques $BetP_i^{\Omega_1}$ et $BetP_j^{\Omega_2}$ sont déduites. Elles définissent les coûts d'associations élémentaires d'une part entre un objet o_i et une détection d_j (coût noté $c_1(o_i, d_j)$) et d'autre part entre une détection d_j et un objet o_i (coût noté $c_2(d_j, o_i)$) :

$$\begin{aligned} \forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, l\}, c_1(o_i, d_j) &= BetP_i^{\Omega_1}(d_j), \\ \forall j \in \{1, \dots, l\}, \forall i \in \{1, \dots, n\}, c_2(d_j, o_i) &= BetP_j^{\Omega_2}(o_i). \end{aligned} \quad (4.18)$$

Deux points de vue sont distingués : (i) solution estimée du côté de $\mathcal{E}_1$. Dans ce cas, l'association optimale représentée par le vecteur de permutation μ_1^* (cf. Equation (4.19)) maximise la probabilité jointe $BetP^{\overbrace{\Omega_1 \times \dots \times \Omega_1}^{n \text{ fois}}}$; (ii) solution estimée du côté de $\mathcal{E}_2$. L'association optimale représentée par le vecteur de permutation μ_2^* (cf. Equation (4.20)) maximise la probabilité jointe $BetP^{\overbrace{\Omega_2 \times \dots \times \Omega_2}^{l \text{ fois}}}$.

$$\mu_1^* = \underset{\mu_1}{argmax} \left[\prod_{i=1}^n BetP_i^{\Omega_1}(d_{\nu_i}) \right], \quad (4.19)$$

$$\mu_2^* = \underset{\mu_2}{argmax} \left[\prod_{j=1}^l BetP_j^{\Omega_2}(o_{\nu_j}) \right], \quad (4.20)$$

où $\mu_1 = [\nu_1 \dots \nu_i \dots \nu_n]$ et $\mu_2 = [\nu_1 \dots \nu_j \dots \nu_l]$ sont les vecteurs de permutation qui représentent les deux associations de type 1-1 (cf. Section 4.4.2.1).

Pour le problème d'optimisation, les auteurs ont proposé un algorithme d'arborescence [6]. Notons cependant qu'il devrait être possible d'utiliser l'algorithme Hongrois dès lors que le logarithme appliqué au produit des coûts élémentaires nous donne une somme de coûts. Pour conclure, l'inconvénient majeur de l'approche de Mercier et al. est qu'elle ne garantit pas l'équivalence des deux solutions (μ_1^* et μ_2^*). Des travaux plus récents [7] ont donc réétudié cette approche pour pallier cet inconvénient.

4.5.1.3 Approche de Dencœux et al. : association optimale [7, 8]

L'approche proposée dans [7, 8] considère un problème d'association 1-1. Elle est optimale puisqu'elle traite les objets de $\mathcal{E}_1$ et les détections de $\mathcal{E}_2$ de manière symétrique. Par ailleurs, elle permet de dépasser le problème de la complexité numérique rencontré

dans les travaux précédents. En effet, l'algorithme d'arborescence proposé dans [6] pour trouver l'association maximisant la probabilité pignistique, nécessite l'énumération des $(n+1)^l$, $n = |\mathcal{E}_1|$, $l = |\mathcal{E}_2|$ associations possibles. Il devient difficile à résoudre lorsque n ou l dépasse la dizaine.

Dans [7, 8], les $n \times l$ bbas $m_{ij}^{\Omega_{ij}}$ (cf. Equation (4.10)) concernant les associations locales entre les objets $o_i \in \mathcal{E}_1$ et les détections $d_j \in \mathcal{E}_2$ sont combinées dans un cadre de discernement commun $\mathcal{R}$ qui est l'ensemble des relations $R \subseteq \mathcal{E}_1 \times \mathcal{E}_2$ vérifiant : $\forall i, j, k$,

(i) $(o_i, d_j) \in R$ et $(o_i, d_k) \in R \Rightarrow j = k$;

(ii) $(o_i, d_k) \in R$ et $(o_j, d_k) \in R \Rightarrow i = j$.

Une relation R est représentée par une matrice R telle que $R_{ij} = 1$ si $(o_i, d_j) \in R$ et $R_{ij} = 0$ sinon.

Les $n \times l$ bbas $m_{ij}^{\Omega_{ij}}$ sont exprimées dans $\mathcal{R}$ comme suit :

$$\begin{aligned} m_{ij}^{\mathcal{R}}(\mathcal{R}_{ij}) &= m_{ij}^{\Omega_{ij}}(y_{ij}), \\ m_{ij}^{\mathcal{R}}(\bar{\mathcal{R}}_{ij}) &= m_{ij}^{\Omega_{ij}}(n_{ij}), \\ m_{ij}^{\mathcal{R}}(\mathcal{R}) &= m_{ij}^{\Omega_{ij}}(\Omega_{ij}), \end{aligned} \quad (4.21)$$

où $\mathcal{R}_{ij} = \{R \in \mathcal{R} / R_{ij} = 1\}$ est l'ensemble des relations associant o_i et d_j .

Les fonctions de plausibilité $pl_{ij}^{\mathcal{R}}$ associées aux bbas $m_{ij}^{\mathcal{R}}$ sont définies selon l'équation (4.22) où $\beta_{ij} = m_{ij}^{\Omega_{ij}}(n_{ij})$ et $\alpha_{ij} = m_{ij}^{\Omega_{ij}}(y_{ij})$.

$$\forall R \in \mathcal{R}, pl_{ij}^{\mathcal{R}}(R) = (1 - \beta_{ij})^{R_{ij}} (1 - \alpha_{ij})^{1-R_{ij}}. \quad (4.22)$$

Supposant l'indépendance des sources élémentaires d'associations locales, le résultat de la combinaison conjonctive des bbas $m_{ij}^{\mathcal{R}}$ est la bba de fonction de plausibilité $pl^{\mathcal{R}}$ définie par l'équation (4.23) dont le logarithme est donné par l'équation (4.24)

$$\forall R \in \mathcal{R}, pl^{\mathcal{R}}(R) = \prod_{i \in \{1, \dots, n\}, j \in \{1, \dots, l\}} (1 - \beta_{ij})^{R_{ij}} (1 - \alpha_{ij})^{1-R_{ij}}, \quad (4.23)$$

$$\forall R \in \mathcal{R}, \ln pl^{\mathcal{R}}(R) = \sum_{i \in \{1, \dots, n\}, j \in \{1, \dots, l\}} [R_{ij} \ln (1 - \beta_{ij}) + (1 - R_{ij}) \ln (1 - \alpha_{ij})] + C, \quad (4.24)$$

où C est une constante. Notons que $\forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, l\}, \alpha_{ij} < 1$ et $\beta_{ij} < 1$.

Pour des situations où $\beta_{ij} = 1$ et $\alpha_{ij} = 1$ traduisant une croyance certaine en l'association ou non de deux éléments, les croyances sont affaiblies légèrement. Ainsi, $\beta_{ij} = 1 - \epsilon$

et $\alpha_{ij} = 1 - \epsilon$, $\epsilon > 0$.

Finalement, à partir de l'équation (4.24), la relation de correspondance la plus plausible R^* entre $\mathcal{E}_1$ et $\mathcal{E}_2$ est obtenue en résolvant un problème de programmation linéaire (cf. Equation (4.25)) sous la contrainte d'association 1-1.

$$R^* = \max_R \left(\sum_{i \in \{1, \dots, n\}, j \in \{1, \dots, l\}} w_{i,j} R_{i,j} \right), \quad (4.25)$$

où

$$w_{i,j} = \ln \frac{1 - \beta_{ij}}{1 - \alpha_{ij}}, \forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, l\}. \quad (4.26)$$

Selon [8], l'algorithme Hongrois (cf. Section (4.4.2.1)) permet d'avoir la solution d'association optimale.

4.5.2 Bbas de description des objets

Dans certaines applications telles que l'identification d'objets, les objets sont modélisés par des bbas. Le cadre de discernement de ces bbas est donc celui de la caractérisation des objets (en terme de classes, localisation, etc.). Ainsi, pour résoudre le problème d'association, [125] a proposé d'exploiter les croyances (les bbas) concernant les caractéristiques des objets sans introduire de nouvelles bbas définies dans des cadres de discernement spécifiques aux croyances sur les pertinences des associations. Il s'agit alors de déduire un coût d'association local à partir de la similarité ou d'une mesure du conflit entre les bbas des objets à associer. Notons qu'hormis le fait de ne pas avoir à définir de bbas d'associations, cette approche offre un critère crédibiliste global consistant du point de vue théorique.

4.5.2.1 Coûts crédibilistes

Etant donnés un objet o_i et une nouvelle détection d_j modélisés respectivement par les bbas m_i et m_j alors le coût de leur association peut être déduit d'une mesure entre m_i et m_j . Dans le cadre de la théorie des fonctions de croyance, plusieurs mesures ont été définies afin de comparer des bbas entre elles. Certains auteurs ont généralisé des mesures définies dans le cadre probabiliste telles que la divergence de Kullback-Liebler [131] et

la distance de Bhattacharyya [126], d'autres se sont inspirés d'une interprétation géométrique des fonctions de croyance d'où la distance L_1 sur les crédibilités [132], la distance L_∞ sur les probabilités pignistiques [133] et la distance L_2 sur les fonctions de masse. Cependant, la mesure la plus répandue est la distance introduite par Jousselme et al. [134] qui tient compte du cardinal des éléments focaux. Par ailleurs, elle a été utilisée par plusieurs auteurs, par exemple Chen et al. [117], Yong et al. [135] et Denœux [136]. Cette mesure crédibiliste quantifie la similarité entre bbas. Ainsi, plus la similarité entre m_i et m_j est grande, plus l'association de l'objet o_i et la détection d_j est pertinente et plus faible doit être leur coût d'association $c(o_i, d_j)$ (cf. Equation (4.27)).

$$\begin{aligned} c(o_i, d_j) &= \sqrt{\frac{1}{2} (\langle m_i, m_i \rangle + \langle m_j, m_j \rangle - 2 \langle m_i, m_j \rangle)}, \\ \langle m_i, m_j \rangle &= \sum_{A \subseteq \Omega} \sum_{B \subseteq \Omega} m_i(A) m_j(B) \frac{|A \cap B|}{|A \cup B|}. \end{aligned} \quad (4.27)$$

La pertinence d'une association peut aussi être définie à partir du conflit entre bbas. En effet, il paraît logique et intuitif de juger la pertinence de la combinaison de deux bbas selon l'absence ou la présence d'éléments contradictoires. Ainsi, le conflit comme critère d'association a été investigué par certains auteurs. Par exemple, [125] a présenté une méthode qui détermine la présence de différents objets observés par différents capteurs en utilisant le degré de conflit pour l'association dans le cadre d'un problème de détection de sous-marins. Précisément, dans [125], le coût local d'association de deux objets peut être directement déduit de la masse sur le vide issue de la combinaison conjonctive de leurs bbas (cf. Equation (4.28)). Ainsi, plus le conflit est important, moins l'association est pertinente.

$$\begin{aligned} c(o_i, d_j) &= \sum_{A \cap B = \emptyset} m_i(A) m_j(B), \\ &= m_{i \odot j}(\emptyset). \end{aligned} \quad (4.28)$$

Notons que, dans [125], le coût d'association globale est ensuite pris comme étant la somme de coûts locaux. Ce choix est arbitraire et peut être revu [126].

4.5.2.2 Plausibilité globale

Dans un contexte un peu différent, à savoir celui de l'identification d'objets (cibles) (en anglais *combat ID declaration*), Ristic a proposé [126] de déduire le critère d'association

des bbas modélisant les classes (ID) des objets à identifier. Nous considérons le fait de s'appuyer sur des bbas représentant les objets ou la croyance dans certaines de leurs caractéristiques comme un avantage majeur puisqu'ainsi la définition des bbas spécifiques pour le sous-problème d'association est évitée.

L'idée fondamentale est d'introduire, de manière formelle, un cadre de discernement qui soit un produit cartésien afin de comparer les bbas des cibles à associer et exprimer ainsi, dans cet espace de façon simple le critère d'association. Dans [126], ce critère est l'identité des identifiants ID des cibles.

Soit $\Omega = \{\omega_1, \dots, \omega_N\}$ l'ensemble des ID possibles pour une cible. Ainsi, étant donnés deux ensembles de cibles $\mathcal{E}_1$ et $\mathcal{E}_2$ et en supposant que les cibles sont identiques si elles ont le même ID, le critère d'association crédibiliste défini par Ristic [126] est :

– Association d'une seule cible

Soient les bbas $m\{o_i\}$ et $m\{d_j\}$ définies sur Ω et qui représentent respectivement la croyance en l'ID de o_i et la croyance en l'ID de d_j , la bba sur le produit cartésien $\Omega^2 = \Omega \times \Omega$ est le résultat de la combinaison conjonctive des extensions vides des bbas $m\{o_i\}$ et $m\{d_j\}$ sur Ω^2 :

$$m^{\Omega^2}\{o_i, d_j\} = m\{o_i\}^{\Omega \uparrow \Omega^2} \odot m\{d_j\}^{\Omega \uparrow \Omega^2}, \quad (4.29)$$

de sorte que, $\forall C \in 2^{\Omega^2}$,

$$m^{\Omega^2}\{o_i, d_j\}(C) = \begin{cases} m^{\Omega}\{o_i\}(A)m^{\Omega}\{d_j\}(B) & \text{si } C = (A, B) \\ 0 & \text{sinon.} \end{cases} \quad (4.30)$$

Cette bba représente la croyance jointe relative aux IDs des deux cibles à associer. En faisant l'hypothèse que deux cibles sont les mêmes ($o_i \equiv d_j$) si elles ont le même ID, la plausibilité relative à leur association est donnée par la somme des produits $m\{o_i\}(A)m\{d_j\}(B)$ pour lesquels $A \cap B \neq \emptyset$:

$$\begin{aligned} pl^{\Omega^2}\{o_i, d_j\}(o_i \equiv d_j) &= \sum_{A \cap B \neq \emptyset} m^{\Omega}\{o_i\}(A)m^{\Omega}\{d_j\}(B), \\ &= 1 - \sum_{A \cap B = \emptyset} m_i^{\Omega}(A)m_j^{\Omega}(B), \end{aligned} \quad (4.31)$$

où les notations $m^\Omega\{o_i\}$ et $m^\Omega\{d_j\}$ ont été abrégées en m_i^Ω et m_j^Ω . De même, puisque $bel(A) = 1 - pl(\bar{A})$:

$$bel^{\Omega^2}\{o_i, d_j\} (o_i \neq d_j) = \sum_{A \cap B = \emptyset} m_i^\Omega(A) m_j^\Omega(B), \quad (4.32)$$

où $\neq$ signifie que les deux cibles sont différentes c'est-à-dire n'ayant pas le même ID.

Or, le deuxième terme n'est autre que la masse sur $\emptyset$ issue de la combinaison conjonctive des bbas m_i^Ω et m_j^Ω . D'où, l'équation (4.31) s'écrit :

$$pl^{\Omega^2}\{o_i, d_j\} (o_i \equiv d_j) = 1 - m_{i \odot j}^\Omega(\emptyset). \quad (4.33)$$

– Association multi-cibles

La généralisation de cette approche à l'association de n cibles o_i avec n cibles d_j est directe. Les bbas des cibles définies sur Ω sont projetées par extension vide sur le produit cartésien Ω^{2n} . La plausibilité de chaque association est ensuite évaluée sur Ω^{2n} comme suit :

$$pl^{\Omega^{2n}}(o_i \equiv d_{\nu_i} : i = 1, \dots, n) = \prod_{i=1}^n (1 - m_{i \odot \nu_i}^\Omega(\emptyset)), \quad (4.34)$$

où $\mu = [\nu_1 \dots \nu_i \dots \nu_n]$ est le vecteur qui représente l'association : $o_i \equiv d_{\nu_i}$ pour $i = 1 \dots n$. [126] définit alors le coût global à optimiser à partir de la plausibilité globale (cf. Equation (4.34)) ce qui conduit à faire apparaître des coûts locaux fondés sur les conflits locaux mais combinés de façon non additive comme dans [125]. En prenant l'opposé du logarithme de l'expression de $pl^{\Omega^{2n}}(o_i \equiv d_{\nu_i} : i = 1, \dots, n)$, [126] écrit finalement le coût global d'une association de manière additive à partir des coûts élémentaires :

$$c_0(\nu_i = j) = -\log(1 - m_{i \odot j}^\Omega(\emptyset)). \quad (4.35)$$

Une fois encore des algorithmes classiques d'optimisation (e.g. [49]) fournissent la solution de manière efficace. Enfin, si le nombre n de cibles dans $\mathcal{E}_1$ est différent du nombre l de cibles dans $\mathcal{E}_2$, [126] propose de compléter l'ensemble le plus petit par des éléments * représentés par des bbas vides. Les cibles de $\mathcal{E}_1$ ou les cibles de $\mathcal{E}_2$ mises en correspondance avec ces éléments dans l'association optimale resteront alors isolées.

4.6 Contribution sur l'association de données

Lors de notre représentation crédibiliste des objets (cf. Section 4.3), nous avons supposé connues les relations entre les détections et les objets. Cependant, les détections sont non labellisées. Ainsi, pour avoir ces relations ‘détecton-objet’, une étape d’association de données est nécessaire.

Formellement, nous désignons par $\mathcal{E}_2 = \{d_1, \dots, d_l\}$ l’ensemble des, $l = |\mathcal{E}_2|$, détections à l’instant t et par $\mathcal{E}_1 = \{o_1, \dots, o_n\}$ l’ensemble des, $n = |\mathcal{E}_1|$, objets au même instant. Comme déjà introduit, une association de type 1-1 (cf. Section 4.4.2.1) est représentée par un vecteur $\mu = [\nu_1 \dots \nu_n]$ de taille n qui donne pour chaque objet d’indice $i \in \{1 \dots n\}$, l’indice ν_i | $\nu_i \in \{0 \dots l\}$ de l’élément associé (détecton ou objet comme expliqué plus loin) s’il existe ; $\nu_i = 0$ signifie l’absence de l’élément associé.

Afin d’associer plusieurs détections à un même objet, il faudrait alors autoriser des associations multiples, i.e. de type 1-N (cf. Section 4.4.2.2). Cependant, ce type d’association nécessite des paramètres sensibles notamment seuils [3] et son comportement est alors difficile à contrôler. Dans ce travail, nous proposons plutôt de considérer itérativement deux types d’association 1-1 :

- une association dite ‘détecton-objet’ entre chaque détection et un objet (notons que la non association est possible s’il n’existe aucun objet compatible avec une détection). Les objets sont alors actualisés pour prendre en compte les nouvelles détections associées tandis que les détections non associées initient de nouveaux objets.
- une association dite ‘objet-objet’ entre les différents objets de façon à les fusionner s’il s’avère qu’ils représentent des sous parties d’un même objet physique.

Chacune des deux associations repose sur la minimisation d’une fonction coût (cf. Section 4.4.1) calculée à partir des bbas qui représentent les objets (cf. Section 4.3). La figure 4.3 décrit l’enchaînement des étapes de l’algorithme proposé.

4.6.1 Fonctions coût

Voyons à présent comment les fonctions coût sont calculées à partir de la représentation même des objets (cf. Section 4.3).

Rappelons tout d’abord que pour chaque objet, la bba représente la croyance en la localisation imprécise d’une prochaine détection. Nous proposons deux fonctions coût : une pour l’association détecton-objet et l’autre pour l’association objet-objet. Pour les

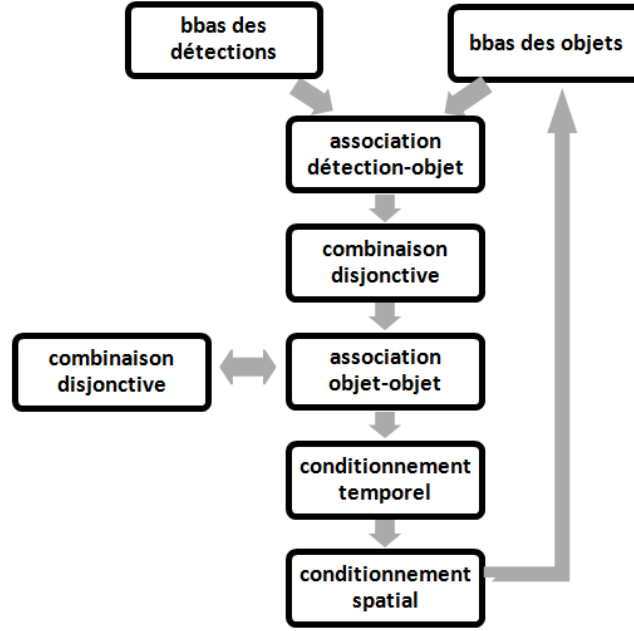


FIGURE 4.3: Algorithme de la construction dynamique d'objets.

définir, nous avons choisi de nous appuyer sur les travaux de [126] (cf. Section 4.5.2.2) bien qu'ils concernent un problème différent, ce qui nous permettra d'exploiter les bbas des objets et les bbas des détections déjà calculées. Dans [126], l'association se fonde sur l'identité des identifiants des objets (le problème considéré étant l'identification de cibles), dans notre cas, elle est fondée sur la localisation des détections dans l'image.

– Association détection-objet

En faisant l'hypothèse qu'une détection d_j et un objet o_i sont des sous régions d'un même objet physique ($o_i \equiv d_j$) s'ils se trouvent à des positions spatiales compatibles, nous avons déduit le critère d'association détection-objet directement à partir de [126] (cf. Section 4.5.2.2). Ainsi, le coût d'une association élémentaire c_0 entre une détection et un objet est le conflit issu de la combinaison conjonctive de leurs bbas (cf. Equation (4.35)). Pour la recherche de l'association optimale et de manière similaire à [126], nous nous sommes fondés sur l'algorithme Hongrois [49]. Par ailleurs, si $|\mathcal{E}_2| = l \neq n$, [126] autorise un nombre de non-associations égal à $|n - l|$. Cependant, dans notre application, souhaitant ne forcer potentiellement aucune association, la matrice de coût est étendue à une matrice carrée de taille $2 \times \max\{l, n\}$ telle que :

$$\begin{aligned}
 \forall (i, j) \in \{1 \dots n\} \times \{1 \dots l\}, M_c(i, j) &= c(\nu_i = j), \\
 \forall (i, j) \notin \{1 \dots n\} \times \{1 \dots l\}, M_c(i, j) &= c_{na},
 \end{aligned}
 \tag{4.36}$$

où c_{na} désigne le coût de non association.

– Association objet-objet

Pour ce second type d'association, nous devons veiller à ce que deux objets distincts puissent se croiser sans pour autant être fusionnés. Afin de limiter les fusions d'objets se croisant, nous utilisons les a priori sur ces derniers et les détections associées.

Tout d'abord, nous prenons en compte la taille des éléments focaux des bbas sous l'hypothèse que plus ces éléments focaux sont de grande taille, plus ils ont de chance de correspondre à des objets déjà construits (et indépendants). Aussi, les compléter par de nouvelles détections (ou un ensemble de détections) doit être fait de manière prudente notamment en tenant compte des tailles des intersections des éléments focaux.

Ensuite, comme déjà mentionné lors du conditionnement spatial, nous supposons que la fragmentation d'un objet a lieu principalement selon l'axe des lignes (notamment s'il présente une symétrie par rapport à l'axe des colonnes). De façon ad hoc, nous proposons alors de prendre en compte non la taille des éléments focaux dans Ω mais celle de leur projection sur l'axe des colonnes. Le critère pour l'association objet-objet est alors le suivant :

$$c_1(\nu_i = j) = -\log \sum_{A \cap B \neq \emptyset} m_i(A) m_j(B) \frac{|A_{\downarrow X} \cap B_{\downarrow X}|}{|A_{\downarrow X}| |B_{\downarrow X}|}, \quad (4.37)$$

où $D_{\downarrow X}$ est la projection de $D \subseteq \Omega$ sur l'axe des colonnes.

Comme pour le cas détection-objet, l'obtention de la solution passe par la définition d'une matrice de coût. Cette dernière est ici de taille $n \times n$ et ses termes sont soit égaux à $c_1(\nu_i = j)$ pour les termes non diagonaux, soit égaux à c'_{na} pour les termes diagonaux (la sélection d'un terme de la diagonale correspond en effet à l'association d'un objet avec lui-même donc sa non-association avec un autre objet). Cette matrice est donc symétrique et la solution obtenue doit également l'être : si le terme de coordonnées (i, j) est sélectionné comme participant à la minimisation, le terme (j, i) doit l'être également. Nous avons donc modifié l'algorithme [49] de façon à préserver la symétrie de la matrice au cours de ses transformations en matrices équivalentes. Dans ce cas, le temps de convergence n'est malheureusement plus garanti. Cependant, expérimentalement, nous avons observé que l'algorithme converge rapidement dans la plupart des cas. Dans les rares autres cas, nous fournissons une solution, mais son optimalité n'est plus garantie (cf. Section A.2).

Finalement, pour gérer la fragmentation des objets en termes de détections en sortie des algorithmes de TI, nous proposons d'associer itérativement les objets en cours de construction, n_{it} fois, en contrôlant le coût de non-association. Cela permet de gérer

les associations multiples de façon séquentielle. Les résultats vont illustrer qu'une telle approche est plus robuste qu'une approche classique 1-N.

4.7 Exemple jouet

Nous illustrons à présent à travers un exemple jouet, les étapes de la méthode proposée : (i) association 1-1 entre les détections à t et les objets en construction en utilisant les coûts définis par l'équation (4.35) ; (ii) actualisation des bbas par combinaison disjonctive des bbas des éléments associés selon l'équation (4.1) ; (iii) associations itératives type 1-1 entre les objets en utilisant les coûts définis par l'équation (4.37) ; (iv) actualisation des bbas en combinant les bbas des éléments associés selon l'équation (4.1) ; (v) conditionnement temporel et combinaison via l'équation (4.2) ; (vi) conditionnement spatial.

Pour cet exemple jouet, les paramètres de l'algorithme sont $\Delta t = 2$, $c_{na} = -\log(0.1) = 2.3$, $c'_{na} = -\log(0.001) = 6.9$ et $\alpha = 0.8$.

A $t = 0$ (cf. Figure 4.4a), nous considérons un objet o_1 tel que $m\{o_1\}^{t=0}(o_1) = 1$, et deux détections d_1 et d_2 . Supposons que d_1 est associée avec o_1 , o_1 est fusionné avec d_1 tandis que d_2 étant non associée initie un nouvel objet o_2 . Après combinaison disjonctive, les bbas de o_1 et o_2 sont catégoriques (cf. Figure 4.4b). Puisque les éléments focaux de ces bbas sont disjoints, les objets ne peuvent pas être associés (un coût d'association infini) ainsi les étapes (iii) et (iv) ne modifient pas les bbas des objets. Le résultat de la combinaison conjonctive et du conditionnement temporel (étape v) est la bba $m\{o_1\}^{t=0}$ ayant deux éléments focaux : $m\{o_1\}^{t=0}(A_1) = 0.2$ et $m\{o_1\}^{t=0}(A_2) = 0.8$ (cf. Figure 4.4c).

A $t = 1$ (cf. Figure 4.4d), une nouvelle détection d_3 apparaît. Puisque $c_0(d_3 \equiv o_1) = -\log(1 \times 0.2) = 1.6$, $c_0(d_3 \equiv o_2) = -\log(1 \times 1) = 0$ et $c_{na} = 2.3$, d_3 est associée avec o_2 . Après combinaison disjonctive, $m\{o_2\}^{t=1}$ est catégorique, la bba de o_1 est unchangede (cf. Figure 4.4e). Suite à l'étape d'association objet-objet, o_2 est associé avec o_1 puisque $c_1(o_1 \equiv o_2) = -\log(0.2 \times 1 \times \frac{15}{45 \times 60}) = 6.7 < c'_{na}$. Après combinaison disjonctive, $m\{o_1\}^{t=1}$ a deux éléments focaux A_1 et A_2 tels que $m\{o_1\}^{t=1}(A_1) = 0.2$ et $m\{o_1\}^{t=1}(A_2) = 0.8$ (cf. Figure 4.4f). Après conditionnement temporel, $m\{o_1\}^{t=1}$ est actualisée et dispose d'un nouveau élément focal A_3 : $m\{o_1\}^{t=1}(A_3) = 0.8$, $m\{o_1\}^{t=1}(A_2) = 0.2 \times 0.8 = 0.16$ et $m\{o_1\}^{t=1}(A_1) = 0.2 \times 0.2 = 0.04$ (cf. Figure 4.4g).

A $t = 2$ (cf. Figure 4.4h), d_4 est détectée. Puisque $c_0(d_4 \equiv o_1) = -\log(0.04 \times 1) = 3.2 >$

c_{na} , d_4 est non associée et initie un nouveau objet o_2 . Après l'étape d'association objet-objet, les deux objets demeurent séparés puisque $c_1(o_1 \equiv o_2) = -\log(0.04 \times 1 \times \frac{15}{75 \times 70}) = 3.2 > c'_{na}$.

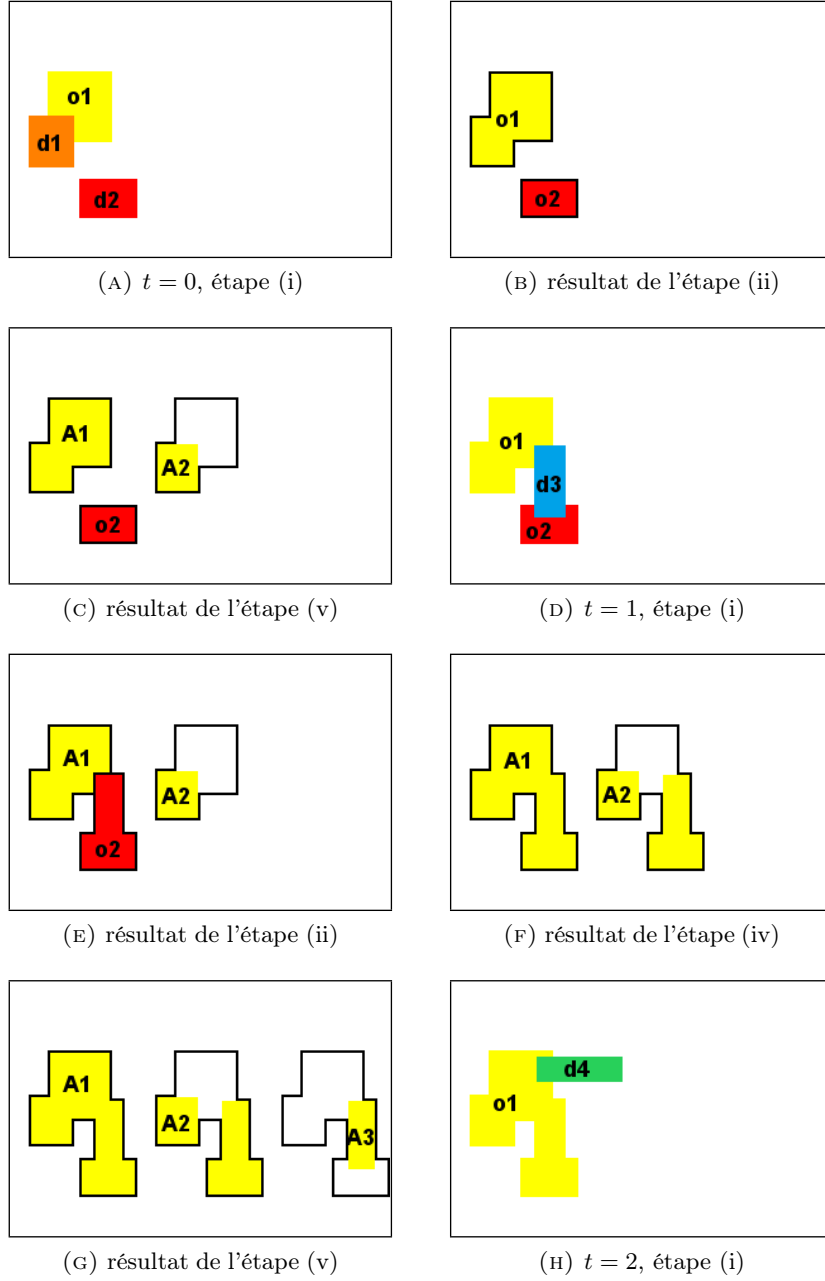


FIGURE 4.4: Cas d'école (les objets sont représentés sans traits; les éléments focaux des bbas sont dessinés en traits gras) : (A,D,H) localisation spatiale des objets et des détections à (A) $t = 0$, (D) $t = 1$, (H) $t = 2$; (B-C,E-G) éléments focaux des bbas d'objets issues des différentes étapes de la méthode proposée.

4.8 Résultats

Afin d'évaluer la méthode proposée, nous l'avons testée sur la séquence de données réelles décrite dans la section 2.5.1. Nous avons utilisé deux algorithmes différents de TI qui génèrent les détections présentées en traits gras sur la figure 4.5 pour un algorithme dérivé de [137] (cf. Section A.3) et sur la figure 4.6, pour l'algorithme [15] (cf. Section 6.2.1).

4.8.1 Validation qualitative

Quatre points sont à vérifier pour valider qualitativement l'algorithme de construction d'objets : (i) faire face aux déplacements d'objets ; (ii) gérer les objets multiples ; (iii) permettre la séparation d'objets ; et (iv) limiter la fusion d'objets se croisant. Ces quatre points sont illustrés dans la figure 4.6. Sur cette figure, les détections à t sont représentées en traits gras et les détections associées à un même objet prennent la même couleur :

- le premier point est vérifié. En effet, les objets (personnes) sont en mouvement. Ceci est visible à partir du décalage spatial entre les détections à t et les détections antérieures présentées en traits fins.
- le second point est aussi vérifié puisque la méthode parvient à construire plusieurs objets simultanément présents.
- le troisième point est illustré sur la figure 4.6a où les deux personnes en blanc et noir viennent d'un groupe de personnes distinctes.
- le dernier point est illustré sur la figure 4.6b où les deux personnes à gauche de l'image demeurent séparées (présentées en couleurs différentes) malgré leur chevauchement spatial.

Ainsi, qualitativement, nous remarquons l'efficacité de la méthode de construction d'objets à partir des détections antérieures.

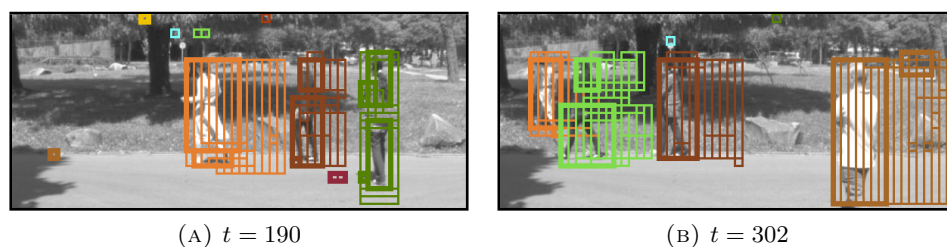


FIGURE 4.5: Exemples de construction d'objets à 2 instants. A un instant donné t , les détections de même couleur sont associées à un même objet. En traits gras, sont représentées les détections à t et, en traits fins, les détections précédentes.

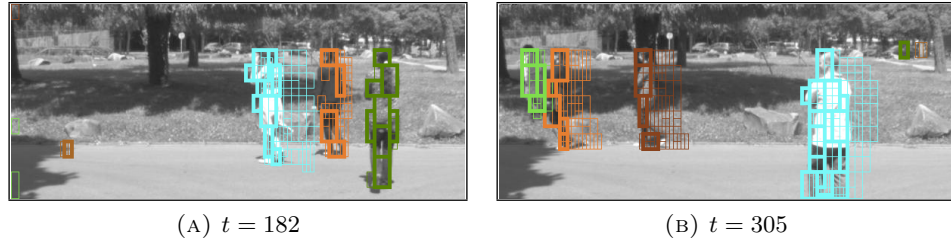


FIGURE 4.6: Exemples de construction d'objets à 2 instants. A un instant donné t , les détections de même couleur sont associées à un même objet. En traits gras, sont représentées les détections à t et, en traits fins, les détections précédentes.

4.8.2 Validation quantitative

Pour l'évaluation quantitative de l'approche nous nous référons à la section 2.6.1.

– Notion d'objets et de fragments

Nous calculons les mesures 'précision (PPV)' et 'rappel (S_e)' (c.f. Section 2.6) soit en terme d'objets soit en terme de fragments. Ceci nous permet de vérifier si les erreurs de labellisation concernent de 'grands' ou 'petits' objets, en terme de nombre de fragments associés. Le premier cas est présenté sur la figure 2.8 où $n_{tp} = 4$, $n_{fp} = 2$ et $n_{fn} = 1$. D'après l'exemple d'association de la figure 2.8, les nombres n_{tp} , n_{fp} et n_{fn} calculés en terme de fragments sont les suivants : $n_{tp} = 8 + 2 + 9 + 1 = 20$ (fragments bien labellisés), $n_{fp} = 1$ et $n_{fn} = 2$.

– Notion de retard à la décision

Nous introduisons aussi la notion de 'retard à la décision' définie comme suit : soit δ_t un retard à la décision (délai temporel) ainsi, les détections fragmentaires à t sont labellisées à partir du résultat d'association à $t + \delta_t$. En variant δ_t , nous évaluons la sensibilité des résultats au nombre n_{it} d'itérations de l'association objet-objet, ainsi que l'ordonnancement temporel des détections. En effet, pour n_{it} fixe, il est possible que des ensembles de fragments (des objets ou des sous parties d'objets) demeurent isolés. Par ailleurs, par transitivité, l'association de deux ensembles de fragments peut se produire par l'intermédiaire d'un nouveau fragment (ou ensemble de fragments) qui les connecte. Dans ce cas, cette association dépend de l'itération à laquelle se produit cette liaison. Finalement, notons que des non associations peuvent induire des FP (Faux Positifs) ou éviter certains FN (Faux Négatifs).

– Approches alternatives

Outre l'évaluation intrinsèque de la méthode proposée, nous avons évalué son intérêt par rapport à des approches alternatives. Rappelons que notre première contribution est de

proposer une représentation imprécise pour les objets en construction (cf. Section 4.3). Ceci est réalisé en définissant une bba par objet qui représente la localisation imprécise d'une détection de cet objet. Cette représentation nous a permis de mettre en place une méthode d'association entre les nouvelles détectons et les objets en cours de construction, en se basant sur leurs localisations dans l'image (deuxième contribution). Une approche alternative est d'ignorer cette représentation crédibiliste et utiliser la théorie des fonctions de croyance directement pour l'association. Ceci a été le sujet de plusieurs travaux tels que [3, 6–8, 124] (cf. Section 4.5.1).

Nous nous focalisons sur la méthode d'association [7, 8] (cf. Section 4.5.1.3) comme référence pour les approches d'association de données exploitant les fonctions de croyance spécifiquement pour l'association. En effet, elle présente les meilleurs résultats selon les auteurs ainsi que la meilleure formulation du problème d'association. Dans les résultats présentés sur la figure 4.7, cette méthode a été introduite dans le schéma de construction d'objets (cf. Figure 4.3) comme suit. Les deux étapes d'association 'détection-objet' et 'objet-objet' sont effectuées selon cette méthode [7, 8] faisant intervenir une distance spatiale entre les éléments à associer. Nous considérons deux distances : (i) si les détectons ou les objets sont représentés par leurs centres alors la distance utilisée, appelée par la suite *center-to-center distance*, est la distance euclidienne entre les centres ; (ii) lorsque les détectons ou les objets sont interprétés en tant qu'ensembles de pixels, la distance appelée '*surface-to-surface distance*' est la distance minimale séparant deux de leurs pixels.

L'algorithme [7, 8] comporte deux paramètres δ et ρ que nous avons fixés respectivement, à 0.001 et 0.9 de manière à obtenir les meilleurs résultats. Par ailleurs, nous avons calculé les coûts de non association c_{na} à partir des distances de non association d_{na} choisies égales à 150 pixels (rayon approximatif d'un objet) pour le cas *center-to-center distance* ($c_{na} = \frac{0.1+0.9\exp(-0.001 \times 150)}{1-0.9 \times \exp(-0.001 \times 150)} = 1.3561$) et 0 pour l'autre cas *surface-to-surface distance* signifiant qu'une association entre deux éléments n'est possible que s'ils s'intersectent ($c_{na} = \frac{0.1+0.9\exp(-0.001 \times 0)}{1-0.9 \times \exp(-0.001 \times 0)} = 10$) ou 100.

Finalement, notons qu'en utilisant les fonctions de croyance seulement pour l'association [7, 8], les objets sont considérés comme étant des sous-ensembles de pixels. Par conséquent, les étapes de combinaison disjonctive, conditionnement temporel et conditionnement spatial sont instanciées à partir de l'équation (4.38) où $o^{\Delta t}$ dénote l'ensemble de pixels appartenant aux détectons produites durant la période de temps Δt et C_i l'ensemble disjonctif des détectons associées à o_i dont la projection sur X appartient à $C^{\downarrow X}$

(cf. Section 4.3.2) :

$$\begin{aligned}
 \forall i \in \{1, \dots, n\} \mid \nu_i \in \{1, \dots, l\}, o_i &\leftarrow o_i \cup d_{\nu_i}, \\
 \forall i \in \{1, \dots, n\}, o_i &\leftarrow o_i \cap o^{\Delta t}, \\
 \forall i \in \{1, \dots, n\}, o_i &\leftarrow o_i \cap C_i.
 \end{aligned} \tag{4.38}$$

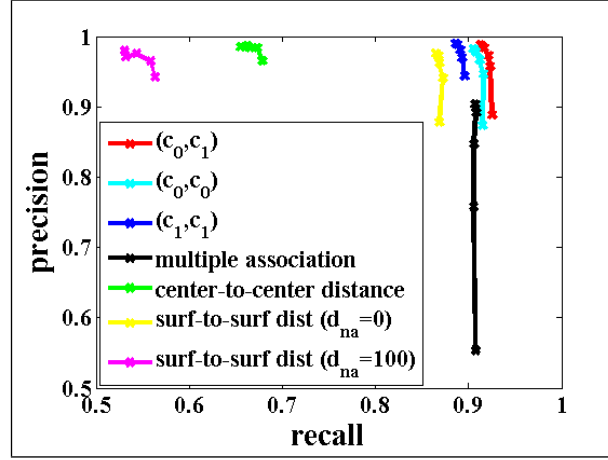
Finalement, les approches par rapport auxquelles notre méthode a été comparée sont trois de ses variantes ainsi que trois instanciations de la méthode fondée sur [7, 8] et adaptée pour le problème de construction d'objets. Ces approches y compris celle proposée sont les suivantes :

- (i) la méthode proposée fondée sur la représentation crédibiliste des objets et utilisant les coûts c_0 et c_1 respectivement pour l'association 'détection-objet' et l'association 'objet-objet'. Elle est appelée (c_0, c_1) ;
- (ii) une variante de notre méthode appelée (c_0, c_0) utilisant le coût c_0 pour les deux types d'associations 'détection-objet' et 'objet-objet' ;
- (iii) une variante de notre méthode appelée (c_1, c_1) utilisant le coût c_1 pour les deux types d'associations 'détection-objet' et 'objet-objet' ;
- (iv) une variante de notre méthode appelée *association multiple* telle que l'association 'détection-objet' est multiple (de type 1 – N c'est-à-dire plusieurs détections peuvent être associées simultanément à un même objet) suivie d'une association 'objet-objet' ;
- (v) une instanciation de la méthode exploitant les fonctions de croyance directement pour l'association selon [7, 8] utilisant la distance *center-to-center* avec $d_{na} = 150$;
- (vi) une instanciation de la méthode exploitant les fonctions de croyance directement pour l'association selon [7, 8] utilisant la distance *surface-to-surface* avec $d_{na} = 0$;
- (vii) une instanciation de la méthode exploitant les fonctions de croyance directement pour l'association selon [7, 8] utilisant la distance *surface-to-surface* avec $d_{na} = 100$.

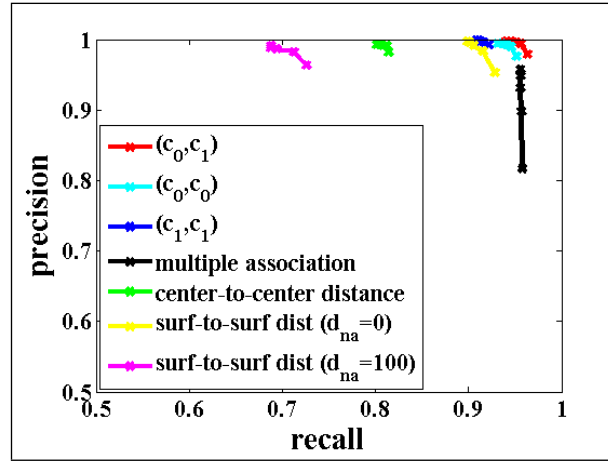
– Résultats et interprétations

La figure 4.7 montre les performances des différentes approches en terme de précision (l'axe des y) en fonction du rappel (l'axe des x) estimées en terme d'objets (cf. Figure 4.7a) et de fragments (cf. Figure 4.7b). Les différentes courbes sont les résultats des sept méthodes y compris la proposée, calculés sur la séquence de 450 images, déjà utilisée pour la validation qualitative (cf. Section 4.8.1).

Tout au long de la séquence, ces méthodes opèrent par constructions et destructions d'objets. En effet, une destruction peut se produire lorsque deux personnes se croisent.



(A)



(B)

FIGURE 4.7: Evaluation quantitative de la méthode proposée en fonction du retard à la décision : taux de rappel et précision calculés en termes de label (A) et fragments (B).

Après séparation ultérieure, l'une des deux personnes se détruit et se reconstruit ultérieurement en cas de réapparition. De manière similaire, si des personnes se séparent d'un groupe (représenté par un seul objet), elles seront détruites et reconstruites après réapparition.

Les points sur une courbe correspondent aux différents retard à la décision variant dans $[0,4]$. Nous constatons que l'augmentation de δ_t implique la décroissance de n_{fp} et la croissance de n_{fn} et par conséquent, l'augmentation de la 'précision' et la diminution légère du 'rappel'. Remarquons aussi, que notre méthode est très performante (de l'ordre de 0.95 pour le rappel et la précision) et elle reste meilleure que les autres approches.

Parmi les trois instanciations de la méthode utilisant les fonctions de croyance directement pour l'association selon [7, 8], (appelées : *center-to-center distance*, *surf-to-surf dist*

($d_{na} = 0$) et *surf-to-surf dist* ($d_{na} = 100$)), nous remarquons que la méthode *surface-to-surface* avec $d_{na} = 0$ fournit les résultats les plus proches de ceux de notre méthode. En effet, cette approche est la plus proche de point de vue critère d'association par rapport à celui de non association (c'est-à-dire accepter ou refuter l'association de deux éléments). Cependant, elle a l'inconvénient de ne pas associer deux objets que s'ils présentent une intersection vide, ce qui a induit la différence au niveau des résultats. En effet, en utilisant la représentation crédibiliste d'objets, il est possible de ne pas associer des objets se croisant, ou au moins d'en retarder l'association pour ceux présentant une intersection non vide.

Parmi les variantes de notre méthode (appelées (c_0, c_0) , (c_1, c_1) et *multiple association*), l'association multiple donne le plus mauvais résultat : en variant δ_t , la précision diminue considérablement sans augmenter notablement le rappel. Ceci montre l'intérêt des associations itératives. En effet, en utilisant des associations itératives, il est possible d'associer des fragments par l'intermédiaire d'autres (la propriété de transitivité) ce qui permet d'avoir des taux importants de rappel et précision. Notons également que l'association objet-objet est essentielle pour connecter des sous parties d'un objet qui ont été construites séparément (e.g. suite à une occultation partielle et temporaire). Finalement, nous observons qu'experimentalement, l'utilisation des deux critères légèrement différents c_0 et c_1 respectivement pour l'association détection-objet et l'association objet-objet fournit les meilleurs résultats : faciliter l'association des détections aux objets (ne pas introduire dans c_0 , la taille des éléments focaux) tout en contrôlant l'association des objets (introduire dans c_1 , la taille de l'intersection des éléments focaux ainsi que les valeurs de leurs masses).

4.9 Conclusion

Ce chapitre avait pour but de définir une méthode de construction des objets présents dans une séquence vidéo à partir de détections fragmentaires en vue de les énumérer et éventuellement les suivre. La dimension temporelle étant alors essentielle pour cumuler les observations. Techniquement, nous avons été confrontés à un problème d'association multiple de données incertaines et imprécises (en terme d'existence et de localisation), que nous avons traité dans le cadre des fonctions de croyance. Ceci nous a mené à deux contributions. La première concerne la représentation crédibiliste des objets qui nous a

permis de proposer un critère à la fois simple et efficace pour résoudre le sous-problème d'association, qui est notre deuxième contribution.

A partir des objets ainsi construits, notre approche d'estimation du cadre de discernement qui sera détaillée dans le chapitre suivant, estime dans un cadre évidentiel l'ensemble des objets d'intérêt et fournit à l'opérateur les menaces les plus pertinentes.

Chapitre 5

Estimation dynamique du cadre de discernement

5.1 Introduction

Notre application réelle de vidéosurveillance consiste à définir un système de protection sur un site dont le but est de détecter les intrus et remonter une alarme ou une menace. Par ailleurs, seules les menaces pertinentes sont présentées à l'utilisateur ce qui permet de diminuer le nombre des fausses alarmes. Dans ce chapitre, nous entendons par cadre de discernement, l'ensemble des objets d'intérêt présents dans une scène imagée. L'estimation de ce cadre n'est autre qu'un filtrage d'objets qui élimine les objets correspondant à des fausses alarmes. Cet ensemble est actualisé de manière séquentielle c'est-à-dire à chaque prise en compte d'une nouvelle source d'information (nouvel ensemble de détections). A un instant donné, l'ensemble estimé des objets d'intérêt peut contenir deux types d'erreurs. Premièrement, il peut contenir des objets dupliqués c'est-à-dire des détections que nous croyons distinctes, correspondent en réalité au même objet physique, qui devrait n'être représenté que par un seul objet. Deuxièmement, certains objets peuvent être fictifs de part la présence des faux positifs dans les résultats de détections. Le premier type d'erreur est reconnu grâce à l'étape d'association étudiée dans le chapitre 4, qui évite d'identifier des sous parties d'un même objet physique en tant qu'objets différents (sur la figure 5.1, les fragments d'un même objet physique tel que l'objet englobé par la boîte jaune doivent avoir le même identifiant). Le deuxième type

d'erreur nécessite un processus de décision regardant la pertinence des différents objets afin de sélectionner ceux les plus pertinents.

Notre approche est développée dans le cadre de la théorie des fonctions de croyance qui gère à la fois l'imprécision et l'incertitude en fusion de données et systèmes de décisions. Dans ce travail, nous proposons de nouveaux mécanismes d'actualisation du cadre de discernement que nous validons sur des données réelles de vidéosurveillance. Notons également que notre approche est beaucoup plus générale et peut concerner d'autres applications telles que la classification et les problèmes d'estimation prenant une décision en faveur d'une hypothèse d'un cadre de discernement qui peut varier d'une façon régulière en fonction du temps ou du contexte et qui doit être actualisé.

Ce chapitre est organisé comme suit. Nous rappelons d'abord l'état de l'art des méthodes d'estimation du cadre de discernement, avant de spécifier le problème considéré. L'approche proposée et la méthodologie qui en découle sont alors présentées. Des résultats obtenus sur des données simulées ainsi que des données réelles illustrent ces travaux avant de conclure.

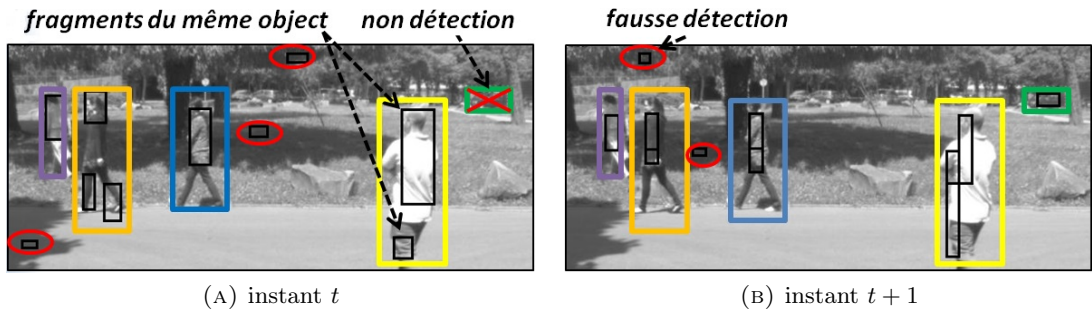


FIGURE 5.1: Exemples de résultats de détection à deux instants successifs t et $t+1$ qui illustrent l'aspect fragmentaire des détections, les fausses détections ainsi que les non détections. Chaque objet physique est englobé par une boîte de couleur différente. Les fausses détections sont entourées en rouge.

5.2 Etat de l'art des méthodes d'estimation du cadre de discernement

Comme exposé au chapitre 3 (cf. Section 3.4.2), sous l'hypothèse du *monde fermé*, un cadre de discernement Ω est un ensemble exhaustif d'hypothèses exclusives, $\Omega = \{\omega_1, \dots, \omega_N\}$ [60]. L'exclusivité des hypothèses signifie qu'elles sont mutuellement disjointes ($\omega_i \cap \omega_j = \emptyset, \forall i \neq j$). L'exhaustivité de Ω implique qu'il contient toutes les

hypothèses possibles pour la variable d'intérêt, et donc nécessairement la solution du problème ω_i appartient à Ω . Sous l'hypothèse du *monde ouvert* introduit par Smets [66], la contrainte d'exhaustivité est écartée. Dans ce cas, l'hypothèse $\emptyset$ représente l'ensemble des hypothèses manquantes. Cependant, l'interprétation d'une masse non nulle sur $\emptyset$ est ambiguë car elle peut traduire soit un conflit entre les sources fusionnées soit le fait que la solution est hors de Ω . Pour lever l'ambiguïté, certains auteurs ont proposé [4] d'ajouter une nouvelle hypothèse spécifique $\{*\}$ à Ω autre que $\emptyset$, qui représente de manière explicite l'ensemble des hypothèses manquantes $\Omega = \{\omega_1, \dots, \omega_N, *\}$. Le cadre ainsi proposé est appelé *monde ouvert étendu*. Cependant, cette approche ne permet pas de distinguer les différentes hypothèses manquantes rassemblées sous l'hypothèse $\{*\}$. Ainsi, une approche alternative est d'estimer l'ensemble actuel d'hypothèses pour déduire le vrai cadre de discernement. C'est dans ce contexte que plusieurs travaux ont été menés afin d'estimer de manière automatique le cadre de discernement, parmi lesquels nous citons [70], [138] et [139].

5.2.1 Estimation du cadre de discernement à partir de sources homogènes

Dans [70], Janez et al. considèrent le cas de sources homogènes (cf. Section 3.5.9) et complémentaires (cf. Section 3.5.1) S_i . Chacune est définie dans un cadre de discernement non exhaustif Ω_i via une bba $m_i^{\Omega_i}$. Ω_i est un sous-ensemble du cadre de discernement commun $\Omega = \bigcup_i \Omega_i$ qui inclut les hypothèses discernées conjointement par toutes les sources dont les bbas doivent être définies dans Ω pour être fusionnées. Ainsi, les auteurs ont proposé deux nouvelles méthodes de définition de bbas :

- Déconditionnements

Deux déconditionnements ont été proposés pour définir une bba $m_i^{\Omega_i}$ dans Ω . Le premier se fonde sur le principe du minimum d'information (cf. Section 3.5.4). Il est tel que :

$$\forall A \subseteq \Omega_i, m_i^{\Omega_i}(A \cup (\Omega \setminus \Omega_i)) = m_i^{\Omega_i}(A), \quad (5.1)$$

où $\Omega \setminus \Omega_i$ dénote l'ensemble d'hypothèses non discernées par S_i . Le deuxième déconditionnement est sous contraintes (cf. Equation (5.3)). Il nécessite des informations supplémentaires sur la compatibilité entre les hypothèses de Ω_i et celles de $\Omega \setminus \Omega_i$. Ainsi, au lieu de rajouter à un élément focal A de $m_i^{\Omega_i}$ toutes les hypothèses manquantes, on lui rajoute

uniquement, l'ensemble des hypothèses manquantes $R(A)$, qui lui sont compatibles où R est la fonction définie par :

$$\begin{aligned} R : 2^{\Omega_i} &\rightarrow 2^{\Omega \setminus \Omega_i} \\ A &\rightarrow \bigcup_{\omega_i \in A} R(\omega_i), \end{aligned} \quad (5.2)$$

avec $R(\omega_i)$ l'ensemble d'hypothèses compatibles avec ω_i .

$$\forall A \subseteq \Omega_i, m_i^\Omega(A \cup R(A)) = m_i^{\Omega_i}(A). \quad (5.3)$$

– Correction des plausibilités

Une autre façon de définir les bbas sans passer par le déconditionnement est la méthode de correction des plausibilités [70]. Pour pouvoir l'appliquer, les sources doivent être définies sur des cadres de discernement connectés : $\forall i, \exists j \neq i, \Omega_i \cap \Omega_j = \Omega_C \neq \emptyset$. En effet, la méthode fusionne toujours des sources disposant d'une partie commune de leurs cadres de discernement et comme le montre l'équation (5.4), la fusion n'aura lieu réellement que sur la partie commune Ω_C . Etant données deux sources S_1 et S_2 de fonctions de plausibilités respectives $Pl_1^{\Omega_1}$ et $Pl_2^{\Omega_2}$, la fonction de plausibilité après fusion $Pl_{1,2}^\Omega$ est donnée par :

$$\begin{aligned} \forall \omega \in (\Omega_1 \setminus \Omega_C), \quad Pl_{1,2}^\Omega(\omega) &= Pl_1^{\Omega_1}(\omega). \\ \forall \omega \in \Omega_C, \quad Pl_{1,2}^\Omega(\omega) &= Pl_1^{\Omega_1}(\omega).Pl_2^{\Omega_2}(\omega). \\ \forall \omega \in (\Omega_2 \setminus \Omega_C), \quad Pl_{1,2}^\Omega(\omega) &= \frac{Pl_1^{\Omega_1}(\Omega_C)}{Pl_2^{\Omega_2}(\Omega_C)}.Pl_2^{\Omega_2}(\omega). \end{aligned} \quad (5.4)$$

– Interprétation

Ce travail traite la possibilité d'exploiter une source d'information même si elle est définie dans un cadre de discernement non exhaustif, et propose deux manières de corriger son information étant données d'autres sources complémentaires. L'intérêt de ce travail est d'autant plus important que les sources fusionnées sont complémentaires c'est-à-dire aptes à discerner les hypothèses manquantes à une source. Ici, le cadre de discernement à estimer n'est autre que l'ensemble des hypothèses discernées conjointement par toutes les sources et décrivant mieux le problème compte tenu de l'information disponible. Par conséquent, l'incomplétude du cadre de discernement peut être corrigée. Finalement, certains inconvénients des deux méthodes de définition des bbas sont notés : (i) l'estimation

des liens de compatibilité entre les hypothèses est une étape critique dans le deuxième déconditionnement, qui affecte fortement le résultat ; (ii) la méthode de correction des plausibilités a l'inconvénient de fournir après fusion, les plausibilités sur les hypothèses simples de l'espace commun sans pouvoir accéder aux fonctions de masse et de crédibilité. Par la suite, elle n'est applicable que si le critère de décision est le maximum de plausibilité ce qui limite son utilisation.

5.2.2 Estimation automatique du cadre de discernement dans le cas d'une classification non supervisée

Dans [138], Le Hégarat et al. traitent un problème de classification non supervisée où l'ensemble des hypothèses à considérer est inconnu. Pour résoudre ce problème d'estimation du cadre de discernement, les auteurs supposent que les classes monosources sont des unions d'hypothèses et que les hypothèses singletons découlent de l'information plus précise multisource, c'est-à-dire fournie par des sources multiples. Par conséquent, dans [138] on propose que le cadre de discernement soit automatiquement estimé à partir des intersections non vides entre les classes discernées par les sources prises individuellement et reconnues à partir des résultats de classifications monosources non supervisées. Plus explicitement, l'idée est la suivante : si deux classes (ou plus) sont non distinguées par une source alors elles forment une seule classe dans le résultat de la classification monosource alors qu'elles représentent une classe composée dans le résultat de la classification multisource. A titre d'exemple, étant données deux sources S_1 et S_2 distinguant respectivement, les classes A_1, A_2 ($\Omega_1 = \{A_1, A_2\}$) et B_1, B_2 ($\Omega_2 = \{B_1, B_2\}$), alors Ω peut contenir jusqu'à 2×2 hypothèses à savoir $\{A_1 \cap B_1, A_1 \cap B_2, A_2 \cap B_1, A_2 \cap B_2\}$. Ω est donc un raffinement (cf. Définition 23) de Ω_1 et Ω_2 de cardinal minimal. Cependant, l'imperfection des résultats de classification monosource peut engendrer des classes multisources fictives c'est-à-dire des hypothèses n'ayant aucune réalité physique, qui correspondent à des 'fausses' intersections de classes. L'élimination des classes fictives passe par une estimation itérative évaluant statistiquement les classes en terme de nombre de pixels. Autrement dit, les classes ayant un nombre de pixels inférieur à un certain seuil sont éliminées. Pour robustifier cette décision, le nombre de pixels par classe est calculé après une classification multisource robuste qui permet de corriger les erreurs en monosource.

– Interprétation

Ce travail est proche de notre problème dans la mesure où il traite l'estimation automatique du cadre de discernement et l'élimination des hypothèses fictives. Toutefois, deux différences sont à noter : (i) dans [138], le cadre de discernement commun est un raffinement des cadres de discernement monosources. Dans notre cas, il est plus grossier. En effet, nous représentons des hypothèses dupliquées par une seule hypothèse ; (ii) pour l'élimination des hypothèses fictives, dans [138], seul un seuil sur le nombre de pixels par classe est utilisé. Dans notre cas, nous passons par un processus de décision dans le cadre de la théorie des fonctions de croyance.

5.2.3 Estimation automatique du cadre de discernement à partir de sources hétérogènes

Rappelons d'abord les notions d'homogénéité et hétérogénéité des sources : des sources sont dites homogènes si elles délivrent de l'information concernant une même variable d'intérêt. Dans le cas contraire où l'information concerne différentes variables d'intérêt, les sources sont dites hétérogènes (cf. Chapitre 3).

Schubert [139] traite le cas de sources potentiellement hétérogènes. Le cadre de discernement commun ainsi défini est un produit cartésien de cadres de discernement associés à des sous-ensembles de sources homogènes en soi et hétérogènes par rapport aux autres. En effet, selon [139], un produit cartésien $\Omega = \Omega_1 \times \dots \times \Omega_N$ ne peut être un cadre de discernement valide si et seulement si il vérifie les deux conditions suivantes :

1. les axes Ω_i de Ω sont : (i) des ensembles homogènes $\Rightarrow \forall i \in \{1, \dots, N\}$, les éléments de Ω_i sont des hypothèses possibles pour une même variable d'intérêt ; (ii) des ensembles exclusifs $\Rightarrow \forall i \in \{1, \dots, N\}, \forall (\omega_j, \omega_k) \in \Omega_i^2, j \neq k \Rightarrow \omega_j \cap \omega_k = \emptyset$.
2. les axes Ω_i de Ω sont : (i) des ensembles mutuellement hétérogènes $\Rightarrow \forall (i, j) \in \{1, \dots, N\}^2, i \neq j$, les éléments de Ω_i et ceux de Ω_j concernent deux variables d'intérêt différentes ; (ii) des ensembles mutuellement indépendants $\Rightarrow$ une hypothèse de Ω est un N -uplet d'éléments prenant des valeurs indépendantes.

Définition 26. Abridgment

A partir d'un produit cartésien Ω vérifiant les conditions mentionnées ci-dessus, il est possible de construire un espace de cardinal plus réduit appelé *abridgment* [139] en réduisant au moins un axe de Ω à l'un de ses sous-ensembles non vides. Par conséquent, il peut y avoir des *abridgments* multiples pour un même Ω . Ainsi, pour juger la qualité des

abridgments et choisir l'optimal, une mesure appelée *Frame Appropriateness* (cf. Définition 27) a été proposée dans [139].

Définition 27. Frame Appropriateness (FA)

La fonction FA est une fonction de H , *abridgment*, et à valeur dans $[0,1]$. Elle dépend d'un ensemble de bbas $\{m_j^H\}$. La valeur de FA en H notée $FA(H/\{m_j^H\})$ est définie par :

$$FA(H/\{m_j^H\}) = \begin{cases} \left[1 - m_{\odot\{m_j^H\}}(\emptyset)\right] \left[1 - \frac{AU(\oplus\{m_j^H\})}{\log_2(|H|)}\right] & \text{si } |H| > 1 \\ 1 - m_{\odot\{m_j^H\}}(\emptyset) & \text{si } |H| = 1. \end{cases} \quad (5.5)$$

où $m_{\odot\{m_j^H\}}(\emptyset) \in [0,1]$ dénote le conflit issu de la combinaison conjonctive (cf. Définition 17) de sources partiellement conflictuelles et $AU \in [0, \log_2 |H|]$ est l'*aggregated uncertainty* qui est une mesure généralisée de l'entropie de Shannon [140] et la non spécificité de Hartley [141] et tient compte alors des deux types d'incertitude. Ainsi, plus le cardinal de H est faible, plus la valeur de AU est faible et plus important sera le conflit. Spécifiquement, si l'*abridgment* H est une hypothèse simple du produit cartésien Ω ($|H| = 1$) obtenue en réduisant chacun des axes de Ω à l'un de ses éléments, $FA(H/\{m_j^H\}) = 1 - m_{\odot\{m_j^H\}}$.

Dans ce travail, nous supposons que de nouvelles informations sont acquises séquentiellement ce qui engendre non seulement l'actualisation du cadre de discernement mais aussi des fonctions de masse définies dans ce cadre. Citons deux applications :

- la première application est la classification. En effet, plus les données sont acquises plus l'information devient complète et précise ce qui peut faire apparaître de nouvelles classes et/ou découvrir de nouvelles relations entre les classes susceptibles d'être fusionnées en une seule. A titre d'exemple, au XIX^e siècle, le naturaliste Huxley a démontré la parenté profonde entre les reptiles et les oiseaux qui les rassemble dans un groupe appelé Sauropsides.
- la deuxième application est notre détection/suivi d'objets. En effet, plus les données sont acquises, plus l'information devient complète et précise de sorte, comme déjà mentionné, que des objets supposés différents peuvent être reconnus comme étant des sous-parties d'un même objet (partiellement occulté) et de nouveaux objets peuvent être détectés.

Dans ce qui suit, nous nous concentrons sur notre application de vidéosurveillance pour illustrer l'estimation dynamique du cadre de discernement dans le cadre de la théorie des fonctions de croyance. Notre approche est plus générale que les travaux cités car elle traite trois types de modifications au niveau du cadre de discernement : (i) addition de nouvelles hypothèses ; (ii) élimination d'hypothèses ; (iii) fusion d'hypothèses.

5.3 Définition du problème (cas général)

Soit $\Theta = \{\theta_1, \dots, \theta_M\}$ un cadre de discernement conçu pour décider entre M hypothèses θ_i . Chaque fois qu'une nouvelle source est utilisée, Θ peut être sujet à des modifications soit par ajout, suppression ou encore fusion d'hypothèses. Ainsi, nous traitons dans ce chapitre, le problème de l'actualisation automatique et dynamique de Θ . Pour cela, c'est intuitif de se poser des questions sur la pertinence de chacune des hypothèses de Θ pour mieux l'estimer, e.g. 'l'hypothèse $\theta_i, i \in \{1, \dots, M\}$ est elle suffisamment pertinente pour qu'elle soit dans Θ ?' ou ' $\theta_j, j > M$ est elle une nouvelle hypothèse qui doit être ajoutée?'. Par conséquent, nous proposons de modéliser la confiance en chacune des hypothèses potentielles θ_k de Θ (potentielles car elles peuvent appartenir ou non à Θ). La raison pour laquelle, nous introduisons les hypothèses o_k et $\bar{o}_k$ associées à l'existence ou non de θ_k dans Θ . Par ailleurs, nous supposons que la confiance en chaque hypothèse dépend uniquement des caractéristiques de ses instances et quantifiée indépendamment des autres. Même s'il pourrait y avoir dans certaines applications, des corrélations entre les différentes hypothèses de Θ , nous supposons dans ce travail l'indépendance. Un exemple typique est le problème de classification : Θ représente l'ensemble des classes. Le deuxième exemple sera donné par notre application de détection d'objets. Ceci nous a mené à construire un cadre de discernement *intermédiaire* $\Omega = \times_{k=1}^N \{o_k, \bar{o}_k\}$ qui est le produit cartésien des éléments $\{o_k, \bar{o}_k\}$.

5.3.1 Produit cartésien

L'utilisation du produit cartésien pour des hypothèses non exclusives a été proposée par Haenni [142] pour contrer une critique [65] de la théorie des fonctions de croyance due en fait à un problème de modélisation : Zadeh [65] a montré que la règle de combinaison

orthogonale est une limite de la théorie des fonctions de croyance. Selon [65], la normalisation masque un conflit éventuel entre les sources. L'exemple présenté est un cas d'experts donnant des avis conflictuels concernant la maladie parmi M : Méningite, C : Commotion cérébrale et T : Tumeur, pouvant affecter un patient et le cadre de discernement considéré est $\Theta = \{M, C, T\}$ qui est un problème de modélisation. En effet, les hypothèses sont non exclusives car le patient peut avoir plusieurs maladies. D'après Haenni [142], le cadre de discernement doit être un produit cartésien tel que les hypothèses non exclusives appartiennent à différents axes. Ainsi, $\Theta = \{M, \bar{M}\} \times \{C, \bar{C}\} \times \{T, \bar{T}\}$ où $\bar{M}$, $\bar{C}$ et $\bar{T}$ dénotent respectivement l'absence de M , C et T .

Pour illustrer ceci sur un exemple très simple, nous supposons quatre hypothèses A , B , C et D telles que A et B d'un côté, et C et D de l'autre sont des ensembles exclusifs et exhaustifs ($A = o_1$, $B = \bar{o}_1$, $C = o_2$, $D = \bar{o}_2$). Une croyance dans ce cas concerne soit une hypothèse (A , B , C ou D) ou une paire d'hypothèses compatibles ((A, C) , (A, D) , (B, C) ou (B, D)). Notons également que le produit cartésien $\{A, B\} \times \{C, D\}$ nous permet de modéliser n'importe quelle croyance, e.g. l'élément focal (A, D) représente A et D simultanément et l'élément focal $(A, \{C, D\})$ va porter la croyance uniquement en A .

5.3.2 Hypothèse d'indépendance

Selon Schubert [139], les axes du produit cartésien sont des ensembles exclusifs, mutuellement indépendants et mutuellement hétérogènes et une hypothèse du produit cartésien est un N-uplet dont les éléments prennent des valeurs indépendantes (cf. Section 5.2.3). Ainsi, si quatre hypothèses A , B , C et D sont telles que A et B d'un côté, et C et D de l'autre sont exclusives mais A et B sont indépendantes de C et D alors le cadre de discernement est $\{A, B\} \times \{C, D\}$. Un exemple typique est la classification multi-label de scènes, e.g. A est le label 'nuit', B est le label 'jour', C est le label 'forêt' et D est le label 'urbain'. Dans notre cas, $A = o_i$ et $B = \bar{o}_i$. Ainsi, la modélisation du problème doit distinguer entre les hypothèses appartenant à un même axe du produit cartésien et celles à différents axes. Dans certaines applications, une telle distinction est évidente dès lors que certaines interactions 'faibles' (par opposition avec les interactions fortes telles que l'incompatibilité) entre les hypothèses sont négligées. Par exemple, si nous considérons notre application de détection d'objets, la présence d'un objet (homme, voiture, chien, etc.) en vidéosurveillance, ne permet pas généralement de prédire la présence ou l'absence d'autres objets dans la scène. Dans d'autres cas, certaines hypothèses sont susceptibles

de se produire simultanément et nous parlons d’une corrélation positive, e.g. les labels ‘voiture’ et ‘route’ dans la classification multi-label de scène. Inversement certaines hypothèses ne peuvent pas se produire simultanément et il s’agit dans ce cas d’une corrélation négative entre les hypothèses, e.g. les labels ‘urbain’ et ‘tigre’. Toutefois, l’information a priori de corrélation entre les hypothèses est très spécifique et généralement négligée même en classification multi-label.

Dans ce contexte, l’idée qui nous paraît la plus prometteuse pour considérer telles corrélations entre les hypothèses a été proposée par Janez et al. [70] qui ont introduit un nouvel opérateur de déconditionnement plus engagé que l’opérateur de déconditionnement classique pour prendre en compte les relations de compatibilité entre les hypothèses (cf. Section 5.2.1). Ainsi, une perspective de ce travail peut être l’intégration de ce nouvel opérateur de déconditionnement à la place de l’extension vide au niveau du produit cartésien.

5.3.3 Cadre de discernement commun

D’une façon générale, si nous disposons d’un cadre de discernement Ω_1 associé à une bba m_1 , et d’une nouvelle bba m_2 définie dans un cadre de discernement Ω_2 (Ω_1 et Ω_2 peuvent être des produits cartésiens), nous cherchons à construire le cadre de discernement commun $\Omega_{1,2}$ et redéfinir les bbas dans $\Omega_{1,2}$ afin de les combiner, par exemple dans [138], $\Omega_{1,2}$ est un raffinement de Ω_1 et Ω_2 (cf. Section 5.2.2). Ici, nous considérons un cas plus général décrit ci après. Ainsi, pour construire $\Omega_{1,2}$, il faut supposer les éventuelles égalités entre certaines hypothèses de Ω_1 et celles de Ω_2 . Lorsque Ω_1 est supposé correct même si non exhaustif [70], ces relations nous conduisent à reconnaître des hypothèses de Ω_2 en tant que des hypothèses de Ω_1 ou de nouvelles hypothèses.

Dans notre cas, nous supposons que Ω_1 peut être erroné. En particulier, il peut être sujet à deux types d’erreurs : (i) des hypothèses dupliquées (c’est-à-dire des hypothèses que nous croyons différentes sont réellement les mêmes) ; (ii) des hypothèses fictives (c’est-à-dire des hypothèses qui ne correspondent à aucune réalisation). Si le premier type d’erreur peut être reconnu via les relations entre les hypothèses de Ω_1 et Ω_2 , l’identification du deuxième nécessite un processus de décision.

5.3.4 Application 1 : classification au niveau pixel

Nous considérons un premier exemple d'application, à savoir la classification au niveau pixel où un pixel de l'image doit être labellisé entre le fond θ_0 et les M_o objets $\theta_i, i \in \{1, \dots, M_o\}$ présents dans la scène ($M_o = M - 1$ où M a été introduit dans la section 5.3 et dénote le cardinal de Θ). Ainsi, le cadre de discernement est $\Theta = \{\theta_0, \theta_1, \dots, \theta_{M_o}\}$. Les M_o objets sont à estimer. Chaque source S_i est un résultat d'un algorithme de détection qui fournit un ensemble E_i de n_i objets différents $\{o_j\}_{j=1, \dots, n_i}$ (Notons que o_j désigne aussi l'existence de l'objet o_j associé à une hypothèse du produit cartésien). Lorsqu'une nouvelle source S_j fournit un nouveau ensemble d'objets E_j , trois cas peuvent se produire. Soit un objet appartient à E_i ou à E_j , soit il appartient aux deux. Dans ce cas, l'association est unique ou multiple. Ainsi, Θ peut être modifié en (i) ajoutant de nouvelles hypothèses (objets appartenant uniquement à E_j) ; (ii) fusionnant les hypothèses multiplement associées ; (iii) supprimant les hypothèses faiblement supportées. Cependant, les décisions quant à ces modifications (ajout d'hypothèses, fusion d'hypothèses, suppression d'hypothèses) doivent être soigneusement prises. C'est la raison pour laquelle, nous utilisons les fonctions de croyance définies sur le produit cartésien Ω . Notons également que Ω peut concerner une sous partie de Θ . Dans cette application de classification, Ω ne concerne pas l'hypothèse fond θ_0 supposée toujours pertinente.

5.3.5 Application 2 : détection d'objets d'intérêt

L'application qui nous intéresse en *tracking* est le filtrage des détections issues d'algorithmes de TI. Par filtrage, nous entendons distinction entre les objets d'intérêt et les fausses alarmes qui doivent donc être éliminées au moment du *tracking*. Le cadre de discernement à estimer est $\Theta = \{\theta_1, \dots, \theta_{M_o}\}$ ($M = M_o$ où M a été introduit dans la section 5.3 et dénote le cardinal de Θ). Ses hypothèses sont les M_o objets d'intérêt θ_j présents dans la scène.

5.4 Méthode proposée [9–11]

Dans cette section, pour une meilleure illustration, nous utilisons le vocabulaire associé à l'application de détection d'objets. Pour généraliser, il suffira de changer les termes

‘détection d’objets’ par ‘distinction d’hypothèses’. Comme nous l’avons déjà noté, nous manipulons deux cadres de discernement : le produit cartésien $\Omega = \times_{j=1}^N \{o_j, \bar{o}_j\}$ où o_j et $\bar{o}_j$ dénotent respectivement la présence du $j^{\text{ème}}$ objet et son absence et le cadre de discernement Θ qui en découle. Lorsque de nouvelles sources fournissent de l’information, non seulement Ω et Θ sont actualisés en ajoutant, fusionnant et/ou écartant des hypothèses, mais aussi les bbas qui leurs sont associées.

5.4.1 Actualisation du produit cartésien Ω

Le cadre de discernement commun qui est un produit cartésien doit vérifier les deux conditions [139] mentionnées dans la section 5.2.3. Ainsi, ses axes sont des ensembles exclusifs (première condition) et mutuellement indépendants (deuxième condition), en particulier, il ne doit pas contenir de duplications entre ses différents axes.

Soient S_1 et S_2 deux sources distinguant respectivement, n_1 objets o'_j et n_2 objets o''_j . Les bbas associées à S_1 et S_2 sont initialement définies dans $\Omega_1 = \times_{j=1}^{n_1} \{o'_j, \bar{o}'_j\}$ et $\Omega_2 = \times_{j=1}^{n_2} \{o''_j, \bar{o}''_j\}$. Les relations entre o'_j et o''_j définissent les correspondances entre les objets. Idéalement, tous les objets sont détectés par les deux sources ce qui implique une bijection entre Ω_1 et Ω_2 . Cependant, en pratique, les objets sont susceptibles d’être détectés uniquement, par l’une des deux sources. Ainsi, nous définissons le cadre de discernement commun Ω comme suit :

$$\Omega = \times_{\substack{j \in \{1 \dots n_1\} / \\ \nexists (k,l) \in \{1 \dots n_2\} \times \{j+1 \dots n_1\} \\ / o'_j \equiv o''_k \equiv o'_l}} \{o'_j, \bar{o}'_j\} \times_{\substack{k \in \{1 \dots n_2\} / \\ \nexists j \in \{1 \dots n_1\} \\ / o'_j \equiv o''_k}} \{o''_k, \bar{o}''_k\} \quad (5.6)$$

où $\equiv$ est la relation transitive d’équivalence définie dans le chapitre 4, liant les éléments de Ω_1 avec ceux de Ω_2 . Selon l’équation (5.6), la définition de Ω se base sur Ω_1 en ajoutant les objets de Ω_2 qui n’existent pas dans Ω_1 et en éliminant les objets de Ω_1 qui sont dupliqués selon S_2 . Spécifiquement, dans l’équation (5.6), le premier terme est le produit de tous les éléments o'_j qui ne sont pas en relation $\equiv$ (par transitivité) avec un autre élément o'_l via un nouvel élément o''_k fourni par S_2 . Le second terme est le produit de tous les nouveaux éléments o''_k n’ayant pas des relations $\equiv$ avec des éléments précédents o'_j .

5.4.2 Actualisation des bbas définies sur Ω

Ayant actualisé le cadre de discernement Ω , actualisons ses bbas. Prenons d'abord le cas le plus simple où $\Omega_{1,2}$ est déduit de Ω_1 en ajoutant un ou plusieurs axes. Ainsi, selon le PMI (cf. section 3.5.4), $m_1^{\Omega_{1,2}}$ est déduite de $m_1^{\Omega_1}$ par extension vide (cf. Equation (3.43)). Le cas où $\Omega_{1,2}$ est dérivé de Ω_1 par grossissement (cf. Définition 23) est plus complexe. Par ailleurs, dans le cas général, la fonction de grossissement n'est pas nécessairement surjective et le nombre d'hypothèses à manipuler peut être important. Ici, nous nous focalisons sur le cas où $\Omega = \times_{k=1}^N \{o_k, \bar{o}_k\}$ et des bbas obtenues par combinaison conjonctive (cf. Equation (3.26)) des extensions vides dans Ω de 'bbas partielles' définies dans $\{o_k, \bar{o}_k\}$. Ce cas est beaucoup plus simple. Par exemple, pour $N = 2$, les éléments faisant intervenir des disjonctions appartenant à

$\{\{(o_1, o_2), (o_1, \bar{o}_2), (\bar{o}_1, o_2)\}, \{(o_1, o_2), (o_1, \bar{o}_2), (\bar{o}_1, \bar{o}_2)\}, \{(o_1, o_2), (\bar{o}_1, o_2), (\bar{o}_1, \bar{o}_2)\}, \{(o_1, \bar{o}_2), (\bar{o}_1, o_2), (\bar{o}_1, \bar{o}_2)\}\}$ ne peuvent pas être des éléments focaux d'une bba définie dans Ω . Ce cas spécifique nous permet de proposer une formulation efficace pour la définition des bbas dans $\Omega_{1,2}$ en se basant sur la propriété suivante.

Propriété 2. Soit m_1^Ω et m_2^Ω deux bbas définies dans le produit cartésien $\Omega = \{o_1, \bar{o}_1\} \times \{o_2, \bar{o}_2\}$ ainsi, la bba issue de la combinaison conjonctive, $\odot$, des bbas m_1^Ω et m_2^Ω s'écrit :

$$m_1^\Omega \odot m_2^\Omega = \left(m_1^{\Omega \downarrow \{o_1, \bar{o}_1\}} \odot m_2^{\Omega \downarrow \{o_1, \bar{o}_1\}} \right)^{\uparrow \Omega} \odot \left(m_1^{\Omega \downarrow \{o_2, \bar{o}_2\}} \odot m_2^{\Omega \downarrow \{o_2, \bar{o}_2\}} \right)^{\uparrow \Omega}, \quad (5.7)$$

où les indices $\Omega \downarrow \{o_1, \bar{o}_1\}$ et $\Omega \downarrow \{o_2, \bar{o}_2\}$ dénotent la marginalisation (cf. Définition 25) des bbas définies dans le produit cartésien Ω , respectivement sur les cadres de discernement élémentaires $\{o_1, \bar{o}_1\}$ et $\{o_2, \bar{o}_2\}$ et l'indice $\uparrow \Omega$ dénote l'extension vide des bbas définies dans $\{o_1, \bar{o}_1\}$ ou $\{o_2, \bar{o}_2\}$, dans le cadre de discernement Ω . D'après l'équation (5.7), la combinaison conjonctive de bbas définies dans un produit cartésien est équivalente à combiner les bbas marginalisées sur chaque axe du produit cartésien, étendre le résultat de la combinaison 1D et combiner conjonctivement les bbas étendues. L'intérêt de l'équation (5.7) est qu'elle ne nécessite que des combinaisons de bbas 1D (résultats des marginalisation) et combinaison de bbas 1D étendues à partir de différents axes du produit cartésien. Par conséquent, elle fait intervenir qu'un terme dans la somme dans l'équation (3.26). Notons également que l'équation (5.7) n'est vraie que si $\emptyset$ est traité similairement aux autres hypothèses : soit $2^{\Omega_i} = \{\emptyset, o_i, \bar{o}_i, \{o_i, \bar{o}_i\}\}$, $\forall i \in \{1, 2\}$ ainsi $\Omega = \{(\omega_1, \omega_2), \omega_1 \in 2^{\Omega_1}, \omega_2 \in 2^{\Omega_2}\}$ et la marginalisation de la bba m^Ω sur $\{o_i, \bar{o}_i\}$

est telle que $\forall \omega_i \in 2^{\Omega_i}, m^{\Omega \downarrow \{o_i, \bar{o}_i\}}(\omega_i) = \sum_{\omega_j \in 2^{\Omega_j}} m^{\Omega}((\omega_i, \omega_j))$. L'Annexe B fournit la démonstration de l'équation (5.7).

Dans cette étude, nous proposons de redéfinir une bba dans le cadre de discernement $\Omega_{1,2}$ de manière plus simple que le grossissement (cf. Définition 23). Pour cela, introduisons la partition $\Pi_1^{\equiv_2}$ de l'ensemble des n_1 détections de S_1 . Un élément $\pi_{1,j}$ de $\Pi_1^{\equiv_2}$ est un sous-ensemble de détections qui, selon S_2 , sont liées deux à deux par $\equiv$. Ainsi, $\Pi_1^{\equiv_2} = \{\pi_{1,1}, \dots, \pi_{1,\nu_1^{\equiv_2}}\}$ ayant $\nu_1^{\equiv_2}$ éléments est tel que :

$$\forall (j, k) \in \{1 \dots n_1\}^2, \exists (i, l) \in \{1 \dots \nu_1^{\equiv_2}\}^2 / \begin{cases} o'_j \in \pi_{1,i}, o'_k \in \pi_{1,l}, \\ i = l \Rightarrow o'_j \equiv o'_k. \end{cases} \quad (5.8)$$

Notons que $\nu_1^{\equiv_2} \leq n_1$ et que l'exposant $\equiv_2$ sur $\Pi_1^{\equiv_2}$ et $\nu_1^{\equiv_2}$ rappelle que les relations entre les objets sont données par la source S_2 . En utilisant les notations précédentes, nous désignons par $m_1^{\Omega_{1,2}}$ la bba m_1 redéfinie de Ω_1 vers $\Omega_{1,2}$.

$$m_1^{\Omega_{1,2}} = \bigoplus_{i \in \{1 \dots \nu_1^{\equiv_2}\}} \left(\bigoplus_{j \in \pi_{1,i}} m_1^{\Omega_1 \downarrow \{o'_j, \bar{o}'_j\}} \right)^{\uparrow \Omega_{1,2}}, \quad (5.9)$$

Selon l'équation (5.9), la marginalisation de la bba de S_1 sur chaque axe du produit cartésien Ω_1 est suivie par la combinaison conjonctive des bbas associées aux éléments dupliqués. Ensuite, les bbas associées aux éléments de la partition c'est-à-dire aux différents objets selon l'état actuel de connaissance sont redéfinies sur $\Omega_{1,2}$ par extension vide. Finalement, les $\nu_1^{\equiv_2}$ bbas définies sur $\Omega_{1,2}$ sont combinées conjonctivement ce qui donne la bba sur $\Omega_{1,2}$ associée à S_1 . Notons également, que si chaque élément de la partition $\Pi_1^{\equiv_2}$ est de cardinal 1 alors $\nu_1^{\equiv_2} = n_1$ (c'est-à-dire il n'existe pas de détections à combiner) et l'équation (5.9) équivaut à une extension vide.

En appliquant l'équation (5.9) sur $m_2^{\Omega_2}$, les bbas $m_1^{\Omega_{1,2}}$ et $m_2^{\Omega_{1,2}}$ peuvent être combinées pour avoir la bba $m_{1,2}^{\Omega_{1,2}} = m_1^{\Omega_{1,2}} \odot m_2^{\Omega_{1,2}}$. Dans plusieurs applications, les sources sont séquentiellement intégrées. Etant données trois sources S_1 , S_2 et S_3 , grâce à la propriété 2 et la transitivité de $\equiv$, nous déduisons l'égalité $m_{1,2,3}^{\Omega_{1,2,3}} = m_{1,2}^{\Omega_{1,2,3}} \odot m_3^{\Omega_{1,2,3}}$, autrement dit, la redéfinition des combinaisons intermédiaires des sources est égale à la combinaison après redéfinition des bbas associées aux sources prises individuellement. L'intérêt d'un tel cas est important dans la mesure où, alors, non seulement l'actualisation de Ω est associative mais aussi la mise à jour des bbas définies sur Ω , et ainsi, pour un grand

nombre d'actualisations (si nous considérons comme exemple, l'acquisition séquentielle des images dans une séquence vidéo), les calculs de la mémoire sont gérés plus efficacement.

5.4.3 Actualisation du cadre de discernement Θ

La dernière étape du processus concerne la décision à partir des bbas actualisées de l'ensemble des objets actuels. Une première approche classique est de décider selon le critère de la probabilité pignistique (cf. Section 3.6.1) qui dans le cas d'un cadre de discernement ayant la forme d'un produit cartésien, s'écrit comme suit :

$$\forall H \in \Omega/H = (H_1, \dots, H_N), H_i \in \{o_i, \bar{o}_i\}, \text{Bet}P(H) = \prod_{i=1}^N \text{Bet}P^{\Omega \downarrow \{o_i, \bar{o}_i\}}(H_i). \quad (5.10)$$

D'après l'équation (5.10), maximiser la probabilité pignistique revient donc à choisir, pour chaque objet o_i , l'hypothèse parmi $\{o_i, \bar{o}_i\}$ maximisant la bba marginalisée $m^{\downarrow \{o_i, \bar{o}_i\}}$.

Le deuxième type de décision est de choisir l'*abridgment* optimal (cf. Définition 26) selon [139], c'est-à-dire minimisant le conflit représenté par la masse sur $\emptyset$. Par ailleurs, dans notre cas, l'*abridgment* est une hypothèse de Ω ($|H| = 1$).

Selon Schubert [139], lorsqu'un *abridgment* H est adopté, la bba m^Ω est conditionnée en H . Le résultat est la bba notée $m^\Omega[H]$. D'une manière générale, rappelons que le conditionnement (cf. Section 3.5.5) d'une bba m^Ω en un sous-ensemble C de Ω selon l'équation (3.18) considère une information selon laquelle C est certain. Il consiste à combiner conjonctivement m^Ω avec une bba catégorique ayant comme seul élément focal C . Dans le cas de l'*abridgment*, le but est de mesurer le degré de conflit en supposant que l'hypothèse H est certaine :

$$\begin{aligned} \forall H \in \Omega/H = (H_1, \dots, H_N), H_i \in \{o_i, \bar{o}_i\}, \\ m[H](\emptyset) &= \sum_{A \in 2^\Omega / A \cap H = \emptyset} m^\Omega(A) \\ &= 1 - \sum_{A \in 2^\Omega / A \cap H \neq \emptyset} m^\Omega(A) \\ &= 1 - \sum_{A \in 2^\Omega / H \in A} m^\Omega(A). \end{aligned} \quad (5.11)$$

Selon l'équation (5.11), minimiser $m[H](\emptyset)$ revient à maximiser la somme des masses sur les sous-ensembles de Ω contenant H . Ce critère est différent du critère $BetP$ qui maximise la somme des masses sur les sous-ensembles de Ω , contenant H , divisés par leurs cardinaux.

5.4.3.1 Contrainte sur le cardinal de Θ

Dans les cas pratiques, nous ajoutons une contrainte sur le cardinal de Θ . Soit M_{max} le nombre maximal d'objets d'intérêt M_o ($M_o \leq M_{max}$) i.e. un cardinal maximum pour Θ qui dépend de l'application. Par exemple, pour notre système de vidéosurveillance qui fournit des alarmes à un opérateur, l'attention de l'opérateur étant une ressource non infinie, l'opérateur ne peut se focaliser que sur un nombre limité d'alarmes. Ainsi, même si le nombre d'objets détectés dépasse M_{max} , seuls les objets les plus importants sont signalés. Une pratique courante est alors d'ordonner les alarmes selon l'importance (ou encore le degré de dangerosité qui peut être déduit de caractéristiques haut niveau telles que les trajectoires des objets, le regroupement, etc.), de présenter les $M_{max} - 1$ alarmes et de regrouper les autres alarmes en une seule (d'indice M_{max}).

Avec une telle contrainte, la maximisation du $BetP$ et la minimisation de la masse sur $\emptyset$ sont réalisées non pas sur toutes les hypothèses de Ω mais seulement sur les hypothèses de Ω ayant au plus M_{max} éléments o_i . Enfin Θ est directement déduit à partir de l'hypothèse maximisant l'équation (5.10) (ou minimisant l'équation (5.11)). Dans notre application à savoir le filtrage des détections, Θ est l'ensemble des éléments $o_i \in H$ (auquel nous lui ajoutons l'hypothèse fond θ_0 si l'application considérée est la classification au niveau pixel).

5.4.4 Exemple jouet

Nous illustrons à présent à travers un exemple jouet la méthode proposée ainsi que les deux critères de décision. Soient trois sources S_1 , S_2 et S_3 qui détectent respectivement, les objets $\{o'_1\}$, $\{o''_1, o''_2\}$ et $\{o'''_1, o'''_2\}$. Ainsi, $\Omega_1 = \{o'_1, \bar{o}'_1\}$, $\Omega_2 = \times_{j=1}^2 \{o''_j, \bar{o}''_j\}$ et $\Omega_3 = \times_{j=1}^2 \{o'''_j, \bar{o}'''_j\}$. Supposons que $o'_1 \equiv o''_1$. Par suite, selon l'équation (5.6), le cadre de discernement commun à S_1 et S_2 est $\Omega_{1,2} = \{o'_1, \bar{o}'_1\} \times \{o''_2, \bar{o}''_2\}$.

En outre, les informations élémentaires concernant les différents objets sont supposées indépendantes (cf. Section 5.3.2). Par exemple, les deux objets o''_1 et o''_2 détectés par S_2

correspondent à deux sources d'information élémentaires qui induisent deux bbas 'partielles' notées $m_2^{\Omega_2 \downarrow \{o_1'', \bar{o}_1''\}}$ et $m_2^{\Omega_2 \downarrow \{o_2'', \bar{o}_2''\}}$, à partir desquelles m_2 est calculée.

La table 5.1 présente les bbas 'partielles' associées aux différents objets détectés par S_1 et S_2 : l'objet o_1' est détecté par S_1 et S_2 tandis que o_2'' est détecté uniquement par S_2 .

TABLE 5.1: bbas 'partielles' correspondant aux objets détectés par S_1 et S_2

Hypothèses	$\emptyset$	o_1'	$\bar{o}_1'$	$\{o_1', \bar{o}_1'\}$	o_2''	$\bar{o}_2''$	$\{o_2'', \bar{o}_2''\}$
$m_1^{\Omega_1 \downarrow \{o_1', \bar{o}_1'\}}$	0	0.5	0.4	0.1	/	/	/
$m_2^{\Omega_2 \downarrow \{o_1'', \bar{o}_1''\}}$	0	0.5	0.4	0.1	/	/	/
$m_2^{\Omega_2 \downarrow \{o_2'', \bar{o}_2''\}}$	0	/	/	/	0.35	0.15	0.5

TABLE 5.2: Actualisation des bbas dans $\Omega_{1,2}$ et estimation de $\Theta_{1,2}$. Pour plus de lisibilité, les hypothèses disjonctives $\{o_1', \bar{o}_1'\}$ et $\{o_2'', \bar{o}_2''\}$ ont été notées respectivement, $\mathcal{O}'_1$ et $\mathcal{O}''_2$, et $(\emptyset, \emptyset), (o_1', \emptyset), (\bar{o}_1', \emptyset)$ et $(\mathcal{O}'_1, \emptyset)$ n'étant pas des éléments focaux ni pour $m_1^{\Omega_{1,2}}$, ni pour $m_2^{\Omega_{1,2}}$ ni pour $m_{1,2}^{\Omega_{1,2}}$, ils ne figurent pas dans la table.

hypothèses de $\Omega_{1,2}$	$(\emptyset, o_2'')$	(o_1', o_2'')	$(\bar{o}_1', o_2'')$	$(\mathcal{O}'_1, o_2'')$
$m_1^{\Omega_{1,2}}$	0	0	0	0
$m_2^{\Omega_{1,2}}$	0	0.175	0.14	0.035
$m_{1,2}^{\Omega_{1,2}}$	0.14	0.1225	0.0840	0.0035
$BetP_{1,2}^{\Omega_{1,2}}(H)$	/	0.355	0.245	/
$m_{1,2}^{\Omega_{1,2}}(\emptyset/H)$	/	0.694	0.7875	/
hypothèses de $\Omega_{1,2}$	$(\emptyset, \bar{o}_2'')$	$(o_1', \bar{o}_2'')$	$(\bar{o}_1', \bar{o}_2'')$	$(\mathcal{O}'_1, \bar{o}_2'')$
$m_1^{\Omega_{1,2}}$	0	0	0	0
$m_2^{\Omega_{1,2}}$	0	0.075	0.06	0.015
$m_{1,2}^{\Omega_{1,2}}$	0.06	0.0525	0.036	0.0015
$BetP_{1,2}^{\Omega_{1,2}}(H)$	/	0.2367	0.1633	/
$m_{1,2}^{\Omega_{1,2}}(\emptyset/H)$	/	0.766	0.8375	/
hypothèses de $\Omega_{1,2}$	$(\emptyset, \mathcal{O}''_2)$	$(o_1', \mathcal{O}''_2)$	$(\bar{o}_1', \mathcal{O}''_2)$	$(\mathcal{O}'_1, \mathcal{O}''_2)$
$m_1^{\Omega_{1,2}}$	0	0.5	0.4	0.1
$m_2^{\Omega_{1,2}}$	0	0.25	0.2	0.05
$m_{1,2}^{\Omega_{1,2}}$	0.2	0.175	0.12	0.005
$BetP_{1,2}^{\Omega_{1,2}}(H)$	/	/	/	/
$m_{1,2}^{\Omega_{1,2}}(\emptyset/H)$	/	/	/	/

Les résultats numériques des différentes étapes de l'estimation de $\Theta_{1,2}$ sont décrites dans la table 5.2 : les deux premières lignes présentent les bbas initiales associées à S_1 et S_2 , définies dans le cadre de discernement commun $\Omega_{1,2}$. Ces bbas sont les résultats d'une extension vide pour la première source et d'une extension vide suivie d'une combinaison conjonctive des bbas 'partielles' pour la deuxième. La troisième ligne de la table concerne

la bba issue de la combinaison conjonctive de $m_1^{\Omega_{1,2}}$ et $m_2^{\Omega_{1,2}}$. Finalement, les valeurs des deux critères proposés sont présentées dans les deux dernières lignes.

Les hypothèses considérées dans la table 5.2, sont des paires d'hypothèses concernant les objets individuels, le premier élément concerne o'_1 et le deuxième concerne o''_2 . Chaque élément d'une paire peut avoir quatre valeurs ($\in \{\emptyset, o_j, \bar{o}_j, \{o_j, \bar{o}_j\}\}$ avec $o_j \in \{o'_1, o''_2\}$), cependant, seuls les éléments focaux pour au moins une bba figurent dans la table 5.2. $m_1^{\Omega_{1,2}}$ est le résultat de l'extension vide (cf. Equation (3.43)) de $m_1^{\Omega_1}$ (égale à $m_1^{\Omega_1 \downarrow \{o'_1, \bar{o}'_1\}}$) : les éléments focaux de $m_1^{\Omega_{1,2}}$ sont de la forme $(., \{o''_2, \bar{o}''_2\})$.

Pour calculer $m_2^{\Omega_{1,2}}$, les bbas partielles $m_2^{\Omega_2 \downarrow \{o''_i, \bar{o}''_i\}}$, $i \in \{1, 2\}$ sont d'abord étendues similairement à $m_1^{\Omega_1 \downarrow \{o'_1, \bar{o}'_1\}}$ puis combinées : $m_2^{\Omega_{1,2}} = (m_2^{\Omega_2 \downarrow \{o''_1, \bar{o}''_1\}})^{\uparrow \Omega_{1,2}} \odot (m_2^{\Omega_2 \downarrow \{o''_2, \bar{o}''_2\}})^{\uparrow \Omega_{1,2}}$. $m_2^{\Omega_{1,2}}$ peut être calculée de manière plus simple à partir des bbas partielles en utilisant l'équation (3.46). Ainsi, la masse sur une hypothèse (x, y) , $x \in \{\emptyset, o'_1, \bar{o}'_1, \{o'_1, \bar{o}'_1\}\}$, $y \in \{\emptyset, o''_2, \bar{o}''_2, \{o''_2, \bar{o}''_2\}\}$ est le produit des valeurs des bbas partielles respectivement, sur x et y . Par exemple, $m_2^{\Omega_{1,2}}((o'_1, o''_2)) = m_2^{\Omega_2 \downarrow \{o''_1, \bar{o}''_1\}}(o'_1) \times m_2^{\Omega_2 \downarrow \{o''_2, \bar{o}''_2\}}(o''_2) = 0.5 \times 0.35 = 0.175$. La bba $m_1^{\Omega_{1,2}}$ est égale à $m_1^{\Omega_{1,2}} \odot m_2^{\Omega_{1,2}}$ (e.g. $m_1^{\Omega_{1,2}}((o'_1, o''_2)) = 0.5 \times 0.175 + 0.5 \times 0.035 + 0.1 \times 0.175 = 0.1225$).

Les valeurs du *BetP* sont déduites à partir de $m_{1,2}^{\Omega_{1,2}}$
(e.g. $BetP_{1,2}^{\Omega_{1,2}}((o'_1, o''_2)) = \frac{0.1225 + 0.0035/2 + 0.175/2 + 0.005/4}{1 - 0.14 - 0.06 - 0.2} = 0.355$).

Finalement, les valeurs du conflit dû à un *abridgment* sont calculées à partir de l'équation (5.11) (e.g. $m_{1,2}(\emptyset/(o'_1, o''_2)) = 0.084 + 0.0525 + 0.036 + 0.0015 + 0.12 + 0.14 + 0.06 + 0.2 = 0.694$).

A partir de la table 5.2, $BetP_{1,2}^{\Omega_{1,2}}(H)$ est maximale et $m_{1,2}^{\Omega_{1,2}}(\emptyset/H)$ est minimale pour $H = (o'_1, o''_2)$ ainsi $\Theta_{1,2} = \{o'_1, o''_2\}$. Comme dit à la section 5.4.3.1 et selon l'application, le nombre des éléments dans Θ peut être limité. Supposons dans cet exemple jouet, que nous nous focalisons sur l'objet le plus important. Par suite, $M_o = 1$. Ainsi, notons que les résultats selon le premier critère et le second critère de décision sont différents : $\Theta_{1,2}$ est égal soit $\{o''_2\}$ soit à $\{o'_1\}$.

De manière un peu plus générale, nous varions $m_2^{\Omega_2 \downarrow \{o''_2, \bar{o}''_2\}}$ comme suit : soit m_{22} une bba définie dans $\{o''_2, \bar{o}''_2\}$ par $m_{22}(\emptyset) = 0$, $m_{22}(o''_2) = 0.7$, $m_{22}(\bar{o}''_2) = 0.3$. $m_2^{\Omega_2 \downarrow \{o''_2, \bar{o}''_2\}}$ est le résultat de l'affaiblissement de m_{22} par un facteur α variant dans $[0.01, 1]$ (par exemple, la bba $m_2^{\Omega_2 \downarrow \{o''_2, \bar{o}''_2\}}$ dans la table 5.1 correspond à $\alpha = 0.5$).

La figure 5.2 présente deux sous figures 5.2a et 5.2b qui décrivent respectivement, l'évolution du *BetP* et de la masse sur $\emptyset$ pour une hypothèse H de l'ensemble $\{(o'_1, o''_2), (\bar{o}'_1, o''_2), (o'_1, \bar{o}''_2), (\bar{o}'_1, \bar{o}''_2)\}$, et ce en fonction de l'affaiblissement α variant dans

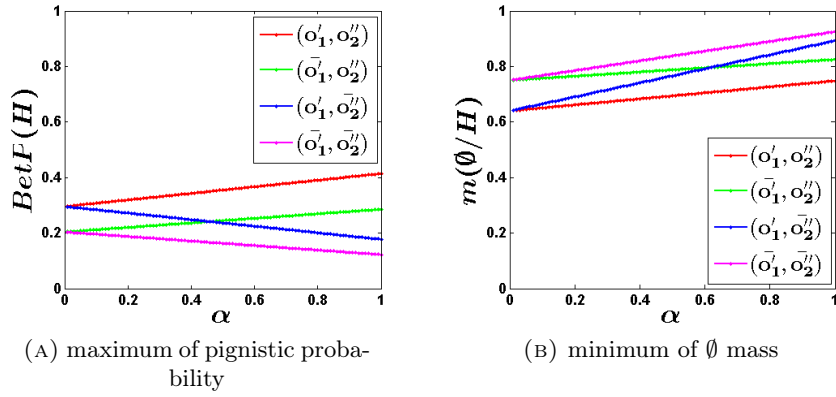


FIGURE 5.2: Comparaison des deux critères de décision.

[0, 1].

A partir de la figure 5.2, nous pouvons constater que quelle que soit la valeur de α , $\Theta_{1,2} = \{o'_1, o''_2\}$ est le cadre de discernement obtenu pour $M_o = 2$. Cependant, si nous supposons qu'il existe au plus un objet alors les résultats de l'estimation de $\Theta_{1,2}$ diffèrent selon la valeur de α : si $\alpha = 1$ ($m_2^{\Omega_{2\downarrow}\{o''_2, \bar{o}''_2\}} = m_{22}$) alors $\Theta_{1,2} = \{o''_2\}$, si $\alpha = 0$ ($m_2^{\Omega_{2\downarrow}\{o''_2, \bar{o}''_2\}}$ est la bba vide) alors $\Theta_{1,2} = \{o'_1\}$. Par ailleurs, $m_2^{\Omega_{2\downarrow}\{o''_2, \bar{o}''_2\}}$ représente la seule croyance en o''_2 . Notons aussi que la valeur α pour laquelle $\Theta_{1,2} = \{o'_1\}$ est 0.52 si le critère de décision est le $BetP_{1,2}^{\Omega_{1,2}}(H)$ et 0.37 s'il est $m_{1,2}^{\Omega_{1,2}}(\emptyset/H)$. Ainsi, pour α variant entre 0.37 et 0.52, les deux critères donnent des résultats différents.

Dans ce qui suit, nous supposons une troisième source S_3 telle que $o'_1 \equiv o''_2 \equiv o'''_1$. Ainsi, o'''_1 est un nouvel objet non détecté par S_1 et S_2 . Les objets o'_1 et o''_2 sont reconnus comme étant des sous parties d'un même objet vu quatre fois : o'_1, o''_1, o''_2 et o'''_1 . En effet, puisque $o'_1 \equiv o''_1$ et $o'_1 \equiv o''_2 \equiv o'''_1$ alors par transitivité de la relation $\equiv$, o'_1, o''_1, o''_2 et o'''_1 sont des détections fragmentaires d'un même objet. D'où, $\Omega_{1,2,3} = \{o'_1, \bar{o}'_1\} \times \{o'''_1, \bar{o}'''_1\}$. La table 5.3 définit initialement la bba m_3 et actualise la bba $m_{1,2}$ suite à la fusion des hypothèses o'_1 et o''_2 . La table 5.4 décrit le résultat de la combinaison conjonctive des bbas $m_{1,2}$ et m_3 dans $\Omega_{1,2,3}$ et nous constatons que les résultats de l'estimation de $\Theta_{1,2,3}$ issus des deux critères de décision sont les mêmes : $\Theta_{1,2,3} = \{o'_1, o'''_1\}$ pour $M_o = 2$ et $\Theta_{1,2,3} = \{o'_1\}$ pour $M_o = 1$.

TABLE 5.3: Actualisation de $m_{1,2}$ suite à la fusion de o'_1 et o''_2 et illustration des bbas ‘partielles’ correspondant aux objets détectés par S_3 . Les hypothèses disjonctives $\{o'_1, \bar{o}'_1\}$ et $\{o''_2, \bar{o}''_2\}$ ont été notées $\mathcal{O}'_1$ et $\mathcal{O}''_2$, respectivement.

hypothèses	$\emptyset$	o'_1	$\bar{o}'_1$	$\mathcal{O}'_1$	o''_2	$\bar{o}''_2$	$\mathcal{O}''_2$
$m_{1,2}^{\Omega_{1,2} \downarrow \{o'_1, \bar{o}'_1\}}$	0.4	0.35	0.24	0.01	/	/	/
$m_{1,2}^{\Omega_{1,2} \downarrow \{o''_2, \bar{o}''_2\}}$	0	0.35	0.15	0.5	/	/	/
$m_{1,2}^{\Omega_{1,2} \downarrow \{o'_1, \bar{o}'_1\}} \oplus m_{1,2}^{\Omega_{1,2} \downarrow \{o''_2, \bar{o}''_2\}}$	0.5365	0.301	0.1575	0.005	/	/	/
$m_3^{\Omega_3 \downarrow \{o'''_1, \bar{o}'''_1\}}$	0	0.7	0.1	0.2	/	/	/
$m_3^{\Omega_3 \downarrow \{o'''_2, \bar{o}'''_2\}}$	0	/	/	/	0.5	0.2	0.3

TABLE 5.4: Actualisation des bbas dans $\Omega_{1,2,3}$ et estimation de $\Theta_{1,2,3}$. Les hypothèses disjonctives $\{o'_1, \bar{o}'_1\}$ et $\{o'''_2, \bar{o}'''_2\}$ ont été notées $\mathcal{O}'_1$ et $\mathcal{O}'''_2$, respectivement. $(\emptyset, \emptyset)$, $(o'_1, \emptyset)$, $(\bar{o}'_1, \emptyset)$ et $(\mathcal{O}'_1, \emptyset)$ n'étant pas des éléments focaux ni pour $m_{1,2}^{\Omega_{1,2,3}}$ ni pour $m_3^{\Omega_{1,2,3}}$ ni pour $m_{1,2,3}^{\Omega_{1,2,3}}$, ils ne figurent pas dans la table.

hypothèses de $\Omega_{1,2,3}$	$(\emptyset, o'''_2)$	(o'_1, o'''_2)	$(\bar{o}'_1, o'''_2)$	$(\mathcal{O}'_1, o'''_2)$
$m_{1,2}^{\Omega_{1,2,3}}$	0	0	0	0
$m_3^{\Omega_{1,2,3}}$	0	0.35	0.05	0.1
$m_{1,2,3}^{\Omega_{1,2,3}}$	0.3384	0.1372	0.0239	0.0005
$BetP_{1,2,3}^{\Omega_{1,2,3}}(H)$	/	0.5528	0.0971	/
$m_{1,2,3}^{\Omega_{1,2,3}}(\emptyset/H)$	/	0.7797	0.9610	/
hypothèses de $\Omega_{1,2,3}$	$(\emptyset, \bar{o}'''_2)$	$(o'_1, \bar{o}'''_2)$	$(\bar{o}'_1, \bar{o}'''_2)$	$(\mathcal{O}'_1, \bar{o}'''_2)$
$m_{1,2}^{\Omega_{1,2,3}}$	0	0	0	0
$m_3^{\Omega_{1,2,3}}$	0	0.14	0.02	0.04
$m_{1,2,3}^{\Omega_{1,2,3}}$	0.1354	0.0549	0.0096	0.0002
$BetP_{1,2,3}^{\Omega_{1,2,3}}(H)$	/	0.2977	0.0524	/
$m_{1,2,3}^{\Omega_{1,2,3}}(\emptyset/H)$	/	0.8623	0.9756	/
hypothèses de $\Omega_{1,2,3}$	$(\emptyset, \mathcal{O}'''_2)$	$(o'_1, \mathcal{O}'''_2)$	$(\bar{o}'_1, \mathcal{O}'''_2)$	$(\mathcal{O}'_1, \mathcal{O}'''_2)$
$m_{1,2}^{\Omega_{1,2,3}}$	0.5365	0.301	0.1575	0.005
$m_3^{\Omega_{1,2,3}}$	0	0.21	0.03	0.06
$m_{1,2,3}^{\Omega_{1,2,3}}$	0.2030	0.0823	0.0143	0.0003
$BetP_{1,2,3}^{\Omega_{1,2,3}}(H)$	/	/	/	/
$m_{1,2,3}^{\Omega_{1,2,3}}(\emptyset/H)$	/	/	/	/

5.5 Application

Nous évaluons à présent l'intérêt de l'approche proposée dans le cas de notre application de détection/ énumération d'objets dans une séquence vidéo.

5.5.1 Données en entrée et relation entre objets $\equiv$

Nos données (sources d'information) sont les détections fournies à chaque instant, à partir d'algorithmes [15] (cf. Section 6.2.1) détectant des fenêtres de changements qui correspondent à des fragments des objets physiques présents dans la scène imagée (cf. Chapitres 2, 4). Etant donné un premier ensemble $\mathcal{E}_1 = \{o_1, \dots, o_n\}$ de $n = |\mathcal{E}_1|$ objets connus à l'instant t et un deuxième ensemble $\mathcal{E}_2 = \{d_1, \dots, d_l\}$ de $l = |\mathcal{E}_2|$ nouvelles détections au même instant t , notre approche crédibiliste d'association de données décrite dans le chapitre 4 nous permet d'associer les éléments de $\mathcal{E}_2$ avec ceux de $\mathcal{E}_1$ et définit ainsi la relation $\equiv$ entre les détections et les objets : $o_i \equiv d_j$ signifie que o_i et d_j sont des sous parties d'un même objet physique.

Notons que dans notre approche d'estimation du cadre de discernement, l'association (cf. Chapitre 4) est supposée suffisamment fiable. Dans le cas contraire, une association fausse augmentera à tort la croyance en un objet (l'objet auquel est associé une nouvelle détection) au détriment de l'objet correct. Par conséquent, pour pallier cette déficience, il nous faut considérer la fiabilité de chaque association élémentaire. Etant donnée une séquence vidéo, une première idée est d'affaiblir la bba de la détection avant de la combiner avec la bba de l'objet auquel elle s'associe.

Notons également que lorsque deux éléments (un objet et une détection) sont mis en relation $\equiv$, cette relation n'est jamais remise en question. Cependant, dans notre approche, deux mécanismes remédient indirectement aux erreurs potentielles d'association de données : (i) les aspects temporels (cf. Section 5.5.3) permettent d'oublier progressivement un objet (diminuer sa croyance avec le temps) même si sa croyance a augmenté suite à l'erreur d'association ; (ii) si deux objets très proches ont été regroupés en un seul, le conditionnement spatial (cf. Section 4.3.3) permet d'initialiser un nouvel objet après séparation des objets malencontreusement fusionnés. Ceci permet aussi de corriger l'erreur d'association.

5.5.2 Définition de la bba élémentaire

Nous désignons par m_i la bba définie dans Ω_i et associée à la source S_i . Elle a pour but de formaliser l'information délivrée par S_i . Nous supposons aussi qu'une détection, de l'objet o_j , fournie par S_i est une source d'information qui renseigne sur l'existence de o_j . Nous lui associons une bba m_{ij} définie dans le cadre de discernement $\Omega_{ij} = \{o_j, \bar{o}_j\}$. La

bba m_i est alors obtenue en combinant les bbas m_{ij} étendues dans Ω_i :

$$m_i^{\Omega_i} = \bigoplus_{j=1}^{n_i} m_{ij}^{\Omega_{ij} \uparrow \Omega_i}, \quad (5.12)$$

où n_i est le nombre de détections issues de S_i générant autant de bbas m_{ij} . Focalisons nous d'abord, sur la définition de la bba m_{ij} .

Dans notre cas, le nombre de pixels de o_j est la seule information dont nous disposons concernant cet objet. Ceci est toujours dans l'esprit de ne considérer que des informations géométriques issues des détections à ce niveau du traitement et de réserver les informations radiométriques pour le *tracking* effectué à plus haut niveau. Par ailleurs, une hypothèse très courante en TI est la décorrélation spatiale et temporelle des fausses alarmes. Par conséquent, notre confiance en la pertinence d'un objet augmente avec sa taille spatio-temporelle : si une détection est de taille significative ou si elle est temporellement persistante (c'est-à-dire détectée plusieurs fois) alors elle est susceptible d'appartenir à un objet pertinent.

Par suite, nous supposons qu'à un instant donné, plus le nombre de pixels d'une détection de o_j est important plus nous sommes confiants que l'objet o_j existe réellement dans la scène. La masse $m_{ij}(o_j)$ est ainsi définie en se basant sur ce nombre noté $|A(o_j)|$. Donc, si à un même instant, deux objets sont détectés pour la première fois alors le petit est moins pertinent que le grand pour la simple raison que le petit objet a la taille d'une fausse alarme. Cependant, si le petit objet persiste suffisamment, sa croyance augmente et il devient autant pertinent que le grand objet.

Supposons également que nos sources présentent plus de faux négatifs que de faux positifs. Spécifiquement, parmi les nombreux algorithmes de détection d'objets, nous nous focalisons sur les méthodes qui permettent de contrôler le nombre de fausses alarmes. En utilisant des données redondantes telles que fournies par une séquence vidéo, le fait de ne pas détecter un fragment d'un objet peut causer un retard dans la détection de l'objet alors qu'une fausse détection introduira un objet potentiel, ce qui nous paraît plus critique [15]. Par conséquent, lorsqu'une détection de o_j présente un faible nombre de pixels, une faible masse est affectée à o_j alors qu'une masse importante est assignée à Ω_{ij} .

Spécifiquement, nous proposons de définir $m_{ij}^{\Omega_{ij}}$ de sorte que l'actualisation des bbas sur Ω et l'estimation de Θ sont associatives. Nous considérons alors qu'un pixel s d'une détection de o_j est une source d'information modélisée via la bba $m_s^{\{o_j, \bar{o}_j\}}$ définie sur

$\{o_j, \bar{o}_j\}$. $m_s^{\{o_j, \bar{o}_j\}}$ est une fonction à support simple (SSF) (cf. Section 3.5.8) qui formalise l'information sur l'hypothèse o_j : $m_s^{\{o_j, \bar{o}_j\}} = \{o_j\}^{1-\beta}$, avec β un paramètre a priori qui représente la masse élémentaire correspondant à la détection d'un pixel. Notons qu'une telle bba présente une dissymétrie entre les hypothèses o_j et $\bar{o}_j$ suite à l'a priori sur les caractéristiques des sources selon lequel le nombre de faux négatifs est plus important que celui de faux positifs.

Par suite, la bba m_{ij} concernant l'existence d'un objet o_j selon la source S_i , est le résultat de la combinaison conjonctive des bbas par pixel : $m_{ij}^{\Omega_{ij}} = \bigodot_{s \in A(o_j)} \{o_j\}^{1-\beta}$.

$$\begin{cases} m_{ij}(o_j) &= 1 - (1 - \beta)^{|A(o_j)|}, \\ m_{ij}(\Omega_{ij}) &= (1 - \beta)^{|A(o_j)|}. \end{cases} \quad (5.13)$$

Le paramètre β est estimé de sorte qu'un objet de taille τ_a supposé '*presque certain*' a une masse égale à η (dans notre cas $\eta = 0.98$).

$$\beta = 1 - (1 - \eta)^{\frac{1}{\tau_a}}. \quad (5.14)$$

τ_a représente le seuil en nombre de pixels au dessus duquel l'objet est considéré certain. C'est un paramètre d'échelle lié à la résolution spatiale de l'image.

La figure 5.3a montre l'évolution du paramètre β (le poids intervenant dans la SSF) par rapport à τ_a (la taille minimale d'un objet '*presque certain*') variant de 5×5 à 50×50 pixels. Plus τ_a est important, plus β est faible. La figure 5.3b montre, la croissance non linéaire de la masse sur un objet o_j en fonction de sa taille $|A(o_j)|$ (cf. Equation (5.13)) pour différentes valeurs de τ_a . Etant donné τ_a , $m_{ij}(o_j)$ croît de manière monotone avec $|A(o_j)|$ et atteint la valeur $\eta = 0.98$ à τ_a .

5.5.3 Aspects temporels

Notre application traite une séquence vidéo. Ainsi, un objet peut être présent uniquement, durant une partie de la séquence. Par conséquent, un processus d'oubli est nécessaire pour diminuer la croyance sur un objet disparu c'est-à-dire n'est plus détecté. Nous proposons de définir ce processus en utilisant l'opérateur d'affaiblissement (cf. Equation (3.15)). Nous désignons alors par α le facteur d'affaiblissement qui permet d'affaiblir une bba $m^{\Omega_{\{o_j, \bar{o}_j\}}}$ associée à un objet o_j observé dans le passé. α est estimé de sorte

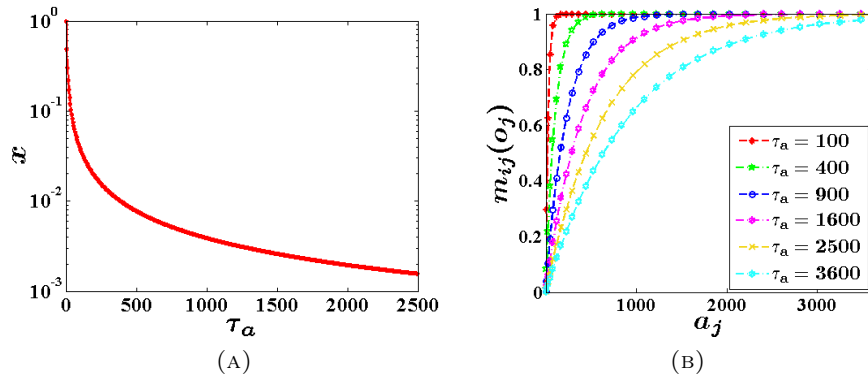


FIGURE 5.3: Définition de la *bba* élémentaire : (a) évolution de la masse élémentaire β (notée x dans la figure) par rapport à la taille τ_a d'un objet certain ; (b) croissance monotone de la masse $m_{ij}(o_j)$ sur un objet o_j en fonction de sa taille $|A(o_j)|$ (notée a_j dans la figure) et du paramètre τ_a .

qu'un objet o_j 'presque certain' (c'est-à-dire de masse η) et qui n'est plus détecté, soit 'presque oublié' après τ_{it} itérations. Ainsi, $\alpha^{\tau_{it}} \times \eta = \epsilon$ et α est donné par :

$$\alpha = \eta^{-\frac{1}{\tau_{it}} \epsilon^{\frac{1}{\tau_{it}}}}, \quad (5.15)$$

où $\epsilon=0.1$ est un paramètre égal à la masse sur une fausse alarme. La figure 5.4 montre

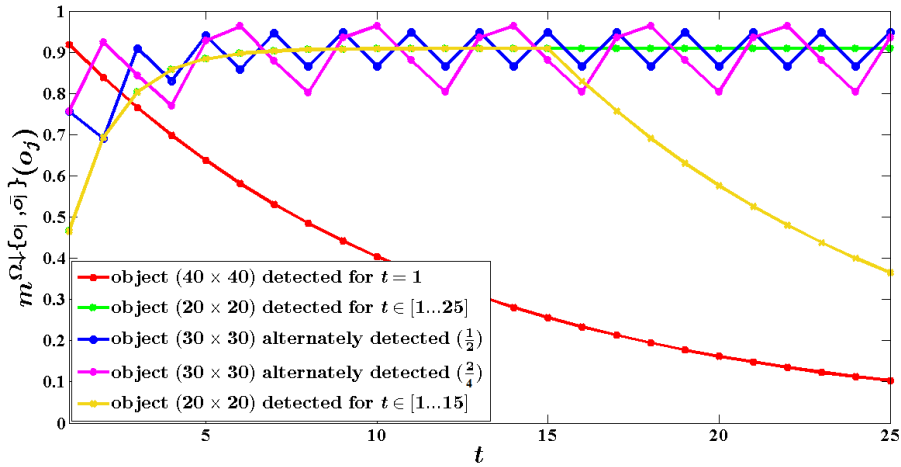


FIGURE 5.4: Exemples de l'évolution de la masse sur un objet par rapport au temps. Les paramètres utilisés sont : $\tau_a = 2500$, $\eta = 0.98$, $\tau_{it} = 25$, $\epsilon = 0.1$.

différents cas d'évolution de la masse sur un objet durant 25 itérations, durée de la persistance d'une croyance. Dans cet exemple, τ_a est égal à 2500 (50×50). Notons également que la masse sur l'objet supposé 'grand', de taille 40×40 et détecté une seule fois, décroît pour atteindre ϵ après 25 itérations. La masse sur le 'petit' objet de taille 20×20 détecté à chaque itération, croît jusqu'à atteindre une valeur asymptotique égale à

$0.9088 = \alpha + (1 - \alpha) \left(1 - (1 - \eta) \frac{|A(o_j)|}{\tau_a} \right)^{-1}$ où α donné par l'équation (5.15) est fonction de η et τ_{it} . Elle dépend de la taille initiale de l'objet et du facteur d'affaiblissement appliqué à chaque itération. Finalement, la masse sur un objet alternativement détecté croît à chaque détection et décroît via le processus d'oubli, à chaque non détection.

5.5.4 Aspects numériques

Comme dit dans la section 5.4.3, le nombre M_o d'hypothèses dans Θ est déduit à partir de l'hypothèse du produit cartésien Ω maximisant la probabilité pignistique ou minimisant la masse sur $\emptyset$. Or, une bba $m_i^{\Omega_i}$ est la combinaison conjonctive de bbas m_{ij} (cf. Equation (5.12)), elles mêmes définies à partir de SSFs $\{o_j\}^{1-\beta}$ (cf. Equation (5.13)). Par conséquent, M_o est toujours égal au nombre d'objets potentiels N qui correspond au nombre d'axes du produit cartésien Ω . Dans la section 5.4.3, il est noté que l'optimisation (maximisation de $BetP(H)$ ou minimisation de $m(\emptyset/H)$) est possible considérant une valeur maximale M_{max} pour M_o . Ainsi, les hypothèses de Ω ayant $M_o \leq M_{max}$ éléments o_i sont des solutions possibles pour le problème d'optimisation.

Cependant, pour une même valeur M_o , plusieurs optimums peuvent exister. Ceci nous a conduit à définir une distance d_{opt} séparant deux premiers optimums notés $H^{*(1)}$ et $H^{*(2)}$. Ainsi, d_{opt} est égale soit à $BetP(H^{*(1)}) - BetP(H^{*(2)})$ soit à $m(\emptyset/H^{*(2)}) - m(\emptyset/H^{*(1)})$. En outre, une solution est supposée non ambiguë si et seulement si $d_{opt} > d_{min}$. En termes de la précision numérique, d_{min} est égale à 10^{-30} . Dans notre cas, afin d'éviter les ambiguïtés entre objets de croyances très proches, nous prenons une valeur de d_{min} égale à 10^{-12} . Dans ce qui suit, nous donnons quelques détails d'implémentation et nous proposons de nouvelles notations :

- $\mathcal{S}_O$ est l'ensemble d'objets potentiels avec $N = |\mathcal{S}_O|$.
- $\{m^{\{o_j, \bar{o}_j\}}\}_{j \in \{1 \dots N\}}$ est l'ensemble des N bbas définies dans les espaces $\{o_j, \bar{o}_j\}$. En effet, les différentes étapes de l'approche proposée manipulent les bbas $m^{\Omega \downarrow \{o_j, \bar{o}_j\}}$ ce qui facilite l'implémentation numérique.
- à un instant t , $\forall j \in \{1 \dots N\}$, $m^{\{o_j, \bar{o}_j\}} = m_{1 \dots t}^{\Omega_{1 \dots t} \downarrow \{o_j, \bar{o}_j\}}$. m^{Ω} est le résultat de la combinaison conjonctive des extensions vides des bbas $m^{\{o_j, \bar{o}_j\}}$ dans Ω . Cependant, la bba m^{Ω} définie dans le produit cartésien Ω n'est utile que pour le calcul de $m(\emptyset/H)$.
- θ^{M_o} est l'ensemble des hypothèses de Ω ayant M_o éléments o_i .
- $\mathcal{S}_O^{(t)}$ est l'ensemble des détections à l'instant t .

- $\left\{m_t^{\{o_j, \bar{o}_j\}}\right\}_{j \in \{1 \dots |\mathcal{S}_O^{(t)}|\}}$ sont les bbas associées aux objets de $\mathcal{S}_O^{(t)}$ c'est-à-dire considérant uniquement l'image à l'instant t .

L'algorithme global de l'estimation du cadre de discernement est donné par l'Algorithme 1.

Les deux versions de l'algorithme proposé (fondé sur le conflit ou le *BetP*) sont efficaces de point de vue temps de calcul. A titre d'exemple, pour une séquence vidéo d'environ 330 images et jusqu'à six objets simultanément présents dans la scène, le temps moyen et maximal (en seconde par image) est respectivement égal à 8.10^{-6} et 3.10^{-4} en utilisant le conflit comme critère de décision et respectivement 2.10^{-3} et 7.10^{-3} en utilisant le *BetP*.

Algorithm 1 Estimation séquentielle de Θ_t

- **En entrée** : séquence d'images de longueur t_{max} , η , τ_a , τ_{it} , M_{max} , ϵ , d_{min} .
- **En sortie** : $\{\Theta_t, t \in \{1 \dots t_{max}\}\}$.
- Initialisations : $\mathcal{S}_O = \emptyset$, $N \leftarrow 0$, α issu de l'équation (5.15).
- **Pour** t de 1 à t_{max} **faire**
 - $\mathcal{S}_O^{(t)} \leftarrow$ détections à l'instant t ;
 - $n_t \leftarrow |\mathcal{S}_O^{(t)}|$;
 - Calculer les n_t bbas $m_t^{\{o_j, \bar{o}_j\}}$, $\forall j \in \{1, \dots, n_t\}$ (cf. Equation (5.13)).
 - **Si** ($t > 1$) **Alors**
 - Calculer les correspondances entre les objets de $\mathcal{S}_O$ et ceux de $\mathcal{S}_O^{(t)}$ (cf. Chapitre 4);
 - $\mathcal{S}_O^{new} \leftarrow$ ensemble des objets de $\mathcal{S}_O^{(t)}$ n'ayant pas des correspondants dans $\mathcal{S}_O$.
 - **Pour** $(o_j, o_k) \in \mathcal{S}_O^2 \mid o_j \equiv o_k, j \neq k$ **faire**
 - Combiner les bbas associées : $m^{\{o_j, \bar{o}_j\}} \leftarrow m^{\{o_j, \bar{o}_j\}} \oslash m^{\{o_k, \bar{o}_k\}}$ connaissant $o_j = o_k$;
 - **Fin Pour**
 - **Pour** $o_j \in \mathcal{S}_O^{(t)}$ **faire**
 - **Si** $\exists o_k \in \mathcal{S}_O \mid o_j \equiv o_k$ **Alors**
 - Combiner les bbas associées $m^{\{o_k, \bar{o}_k\}} \leftarrow m^{\{o_k, \bar{o}_k\}} \oslash m_t^{\{o_j, \bar{o}_j\}}$ connaissant $o_j = o_k$;
 - **Fin Si**
 - **Fin Pour**
 - **Sinon**
 - $\mathcal{S}_O^{new} \leftarrow \mathcal{S}_O^{(t)}$;
 - **Fin Si**
 - $\mathcal{S}_O \leftarrow \mathcal{S}_O \cup \mathcal{S}_O^{new}$, $N \leftarrow N + |\mathcal{S}_O^{new}|$;
 - Actualiser $\{m^{\{o_j, \bar{o}_j\}}\}_{j \in \{1 \dots N\}}$ en ajoutant les bbas $m_t^{\{o_k, \bar{o}_k\}}$ des éléments de $\mathcal{S}_O^{new}$;
 - Actualiser Ω en utilisant l'équation (5.6);
 - Estimer Θ_t selon Algorithme 2 ou Algorithme 3;
 - Affaiblir les bbas $m^{\{o_j, \bar{o}_j\}}$, $\forall j \in \{1, \dots, N\}$ en utilisant l'équation (3.15).
 - **Fin Pour**

Algorithm 2 Estimation de Θ_t en utilisant le critère de décision fondé sur le *BetP*

-
- **En entrée** : $N, \left\{ m_t^{\{o_j, \bar{o}_j\}} \right\}_{j \in \{1 \dots N\}}, M_{max}, d_{min}$.
 - **En sortie** : Θ à l'instant t .
 - Calculer $BetP_t^{\{o_j, \bar{o}_j\}}, \forall j \in \{1, \dots, n_t\}$ en utilisant l'équation (3.49);
 - Dédire $BetP_t$ sur les hypothèses simples de Ω_t en se basant sur l'équation (5.10);
 - $M_o \leftarrow \min(N, M_{max})$;
 - **Si** $N = M_o$ **Alors**
 - $\Theta_t = \{o_1, \dots, o_N\}$
 - **Sinon**
 - **Répéter**
 - Calculer $\theta^{(M_o)}, H^{*(1)} \leftarrow \operatorname{argmax}_{H \in \theta^{(M_o)}} BetP(H)$;
 - $\theta^{(M_o)} \leftarrow \theta^{(M_o)} \setminus \{H^{*(1)}\}, H^{*(2)} \leftarrow \operatorname{argmax}_{H \in \theta^{(M_o)}} BetP(H)$;
 - $d_{opt} \leftarrow BetP(H^{*(1)}) - BetP(H^{*(2)})$;
 - $M_o \leftarrow M_o - 1$;
 - **jusqu'à ce que**
 - $d_{opt} > d_{min}$ et $M_o \geq 1$
 - Dédire Θ_t à partir de $H^{*(1)}$;
 - **Fin Si**
-

Algorithm 3 Estimation de Θ_t en utilisant le critère de décision fondé sur le conflit induit par un *abridgment*.

-
- **En entrée** : $N, \left\{ m_t^{\{o_j, \bar{o}_j\}} \right\}_{j \in \{1 \dots N\}}, M_{max}, d_{min}$.
 - **En sortie** : Θ à l'instant t .
 - Calculer les 2^N éléments focaux de $m_t^{\Omega_t}$ en utilisant l'équation (3.46);
 - Calculer $m_t(\emptyset/H), \forall H \in \Omega_t$ en utilisant l'équation (5.11);
 - $M_o \leftarrow \min(N, M_{max})$;
 - **Si** $N = M_o$ **Alors**
 - $\Theta_t = \{o_1, \dots, o_N\}$
 - **Sinon**
 - **Répéter**
 - Calculer $\theta^{(M_o)}, H^{*(1)} \leftarrow \operatorname{argmin}_{H \in \theta^{(M_o)}} m(\emptyset/H)$;
 - $\theta^{(M_o)} \leftarrow \theta^{(M_o)} \setminus \{H^{*(1)}\}, H^{*(2)} \leftarrow \operatorname{argmin}_{H \in \theta^{(M_o)}} m(\emptyset/H)$;
 - $d_{opt} \leftarrow m(\emptyset/H^{*(2)}) - m(\emptyset/H^{*(1)})$;
 - $M_o \leftarrow M_o - 1$;
 - **jusqu'à ce que**
 - Dédire Θ_t à partir de $H^{*(1)}$;
 - **Fin Si**
-

5.5.5 Récapitulatif des paramètres de la méthode

Les paramètres de la méthode proposée sont les 'entrées' de l'Algorithme 1 : $t_{max}, \eta, \tau_a, \tau_{it}, M_{max}, \epsilon, d_{min}$. Ceux en 'entrée' de l'Algorithme 2 et l'Algorithme 3 dérivent de l'Algorithme 1. Par ailleurs, même si deux autres paramètres à savoir β et α sont utilisés, ils sont directement déduits des paramètres de l'Algorithme 1 : (i) η et τ_a pour calculer

β selon l'équation (5.14); (ii) η , ϵ et τ_{it} pour calculer α selon l'équation (5.15).

Parmi les paramètres de l'Algorithme 1 :

- t_{max} est une donnée en entrée désignant la longueur de la séquence vidéo.
- d_{min} , η et ϵ sont des paramètres internes à la méthode et indépendants des données traitées.
- d_{min} est un seuil pour vérifier la présence d'optimums et donc de solutions ambiguës. Il est positif et très proche de 0. Sa valeur par défaut est égale à 10^{-12} .
- η est la valeur de la masse sur un objet 'presque certain' et prend une valeur par défaut très proche de 1 à savoir 0.98.
- ϵ est la valeur de la masse sur une 'fausse alarme' ou un 'petit objet' qui vient d'apparaître dans la scène et prend une valeur par défaut proche de 0 à savoir 0.1.

τ_a , τ_{it} et M_{max} sont les paramètres 'réels' de la méthode dans le sens où ils dépendent de l'application et des données traitées :

- τ_a est un seuil, en nombre de pixels, de la taille d'un objet considéré 'presque certain' c'est-à-dire de masse égale à η . C'est un paramètre d'échelle qui définit les objets d'intérêt (pour une première apparition) en fonction de la résolution spatiale de l'image ainsi que la taille 'standard' d'un objet et la distance par rapport à la caméra.
- τ_{it} est le nombre d'instantanés consécutifs durant lesquels un objet qui n'est plus détecté est considéré perdu c'est-à-dire de masse égale à ϵ . Ainsi, τ_{it} est un paramètre de temps qui définit, selon la fréquence vidéo, le processus d'oubli pour un objet considéré précédemment 'd'intérêt'.
- M_{max} est le nombre maximal d'objets à suivre. Il est fortement lié à l'application. Dans notre cas, un système de vidéosurveillance fournit des alarmes à un opérateur humain qui se focalise sur un nombre limité, M_{max} , d'alarmes.

Notons que les valeurs typiques pour des séquences vidéo usuelles c'est-à-dire 25 images par seconde et 320×240 pixels sont $\tau_a = 2500$ (pixels i.e. 1 mètre $\times$ 1 mètre pour une distance environ 15 mètres et une focal environ 800 pixels), $\tau_{it} = 25$ (images i.e. 1 seconde). Quelques exemples de la variation de la masse sur un objet par rapport au temps, en utilisant ces valeurs, sont présentés sur la figure 5.4.

5.6 Résultats

Nous illustrons à présent des résultats sur des données simulées ainsi que des données réelles afin d'évaluer quantitativement et qualitativement la méthode proposée.

5.6.1 Résultats sur des données simulées

Les données sont des séquences vidéo simulées présentant un nombre variable d'objets. Une séquence simulée est telle que des objets persistent durant toute la séquence et d'autres uniquement sur une partie de la séquence. La séquence des vérités terrain (une vérité terrain par unité de temps) présente trois périodes successives en terme d'objets, chacune de durée égale à $|\mathcal{T}|$ unités de temps.

Durant la première période $t \in \{1, \dots, |\mathcal{T}|\}$, 4 objets d'intérêt sont présents dans la scène simulée. Ensuite, à l'itération $t = |\mathcal{T}| + 1$, un objet précédemment présent disparaît de la scène et un nouvel objet apparaît de sorte que le nombre total d'objets reste le même. Ces quatre objets persistent jusqu'à $t = 2|\mathcal{T}|$. Puis, à $t = 2|\mathcal{T}| + 1$, deux objets parmi les quatre disparaissent et deux nouveaux objets apparaissent.

Les images d'une séquence sont simulées étant données les probabilités de non détection et de fausses alarmes notées respectivement, p_{nd} et p_{fa} . Pour simuler la fragmentation d'objets, chaque objet est considéré comme étant un ensemble de fragments. Plus précisément, un objet est constitué de 1 à 6 fragments, chacun de taille 10×10 pixels. Ainsi, la présence ou non d'un fragment d'un objet o_i dans l'image S_t (image à l'instant t) dépend de la valeur d'une variable aléatoire z tirée uniformément dans $[0,1]$ et de p_{nd} : si $z \geq p_{nd}$ alors le fragment est présent dans S_t .

Les fausses alarmes sont corrélées temporellement traduisant ainsi, des cas réels tels que les mouvements des arbres à cause du vent, ombre, oiseaux qui passent etc. Pour cela, nous supposons qu'il existe 8 positions possibles pour les fausses alarmes. Pour chaque position, une variable aléatoire z est tirée uniformément dans $[0,1]$. Elle permet de décider de l'existence ou non d'une fausse alarme à cette position de l'image : S_t contient la fausse alarme à cette position si $z < p_{fa}$.

L'Algorithme 1 est appliqué sur des séquences simulées avec $t_{max} = 3 \times |\mathcal{T}|$, $\eta = 0.98$, $\tau_a = 2500$, $M = 4$, $\epsilon = 0.1$, $\tau_{it} \in \{5, 10, 15, 20\}$ et en variant les paramètres p_{nd} , p_{fa} et $|\mathcal{T}|$. Pour un triplet donné, la séquence des cadres de discernement $\{\Theta_t\}_{t \in \{1, \dots, 3|\mathcal{T}|\}}$ est le résultat de l'approche proposée sur une séquence simulée $(S_t, t \in \{1, \dots, 3|\mathcal{T}|\})$.

Etant donnée une séquence de vérités terrain qui correspond à un triplet donné $(p_{nd}, p_{fa}, |\mathcal{T}|)$, nous calculons la moyenne des valeurs de S_e et PPV à t (cf. Section 2.6.2) estimées sur différentes séquences simulées. La figure 5.5 illustre les variations de S_e et PPV en fonction du temps t pour différents triplets. Pour un triplet donné $(p_{nd}, p_{fa}, |\mathcal{T}|)$, les différentes courbes représentent les performances en termes de S_e et PPV , pour différentes valeurs de τ_{it} et les points sur une courbe donnée sont les performances obtenues aux différents instants t . Notons également que les résultats issus des deux critères de décision à savoir la maximisation de la probabilité pignistique (cf. Algorithme 2) et la minimisation du conflit (cf. Algorithme 3) sont les mêmes. Finalement, à partir de la figure 5.5 nous tirons les conclusions suivantes :

- plus le nombre d’images entre deux vérités terrain est important c’est-à-dire plus $|\mathcal{T}|$ est important, meilleurs sont les résultats.
- plus les valeurs p_{nd} et p_{fa} sont faibles, meilleurs sont les résultats de la fusion.
- plus τ_{it} est faible, plus l’algorithme est réactif par rapport au changement de la vérité terrain.

En utilisant les mêmes données simulées pour $(p_{nd}, p_{fa}, |\mathcal{T}|) = (0.5, 0.25, 6)$ et en appliquant l’Algorithme 1 avec les mêmes valeurs des paramètres, nous illustrons à travers la figure 5.6, la variation des résultats par rapport aux paramètres η et ϵ . Les courbes sont les indicateurs de performance S_e et PPV en fonction du temps t et pour différentes valeurs de $\eta \in \{0.9, 0.95, 0.98\}$ et $\epsilon \in \{0.05, 0.1, 0.2\}$, et les points sur une courbe donnée sont les performances obtenues aux différents instants t . Evidemment, les résultats varient car chaque fois des objets disparaissent et d’autres apparaissent. En effet, les paramètres η et ϵ influent sur la réponse de l’algorithme. Notons également que même si les meilleurs résultats sont obtenus pour $\epsilon = 0.05$, nous utilisons plutôt $\epsilon = 0.1$ comme valeur par défaut dans l’Algorithme 1 pour être plus robuste aux occultations temporaires des objets.

5.6.2 Résultats sur des données réelles

Afin d’évaluer qualitativement la méthode proposée, nous l’avons testée sur la séquence de données réelles décrite dans la section 2.5.2. Les détections fragmentaires sont fournies par l’algorithme [15].

La figure 5.7 montre l’évolution des masses $m^{\Omega \downarrow \{o_j, \bar{o}_j\}}(o_j)$ sur les objets o_j en fonction du temps t . Elle nous donne une idée claire sur la pertinence de chaque objet présent dans

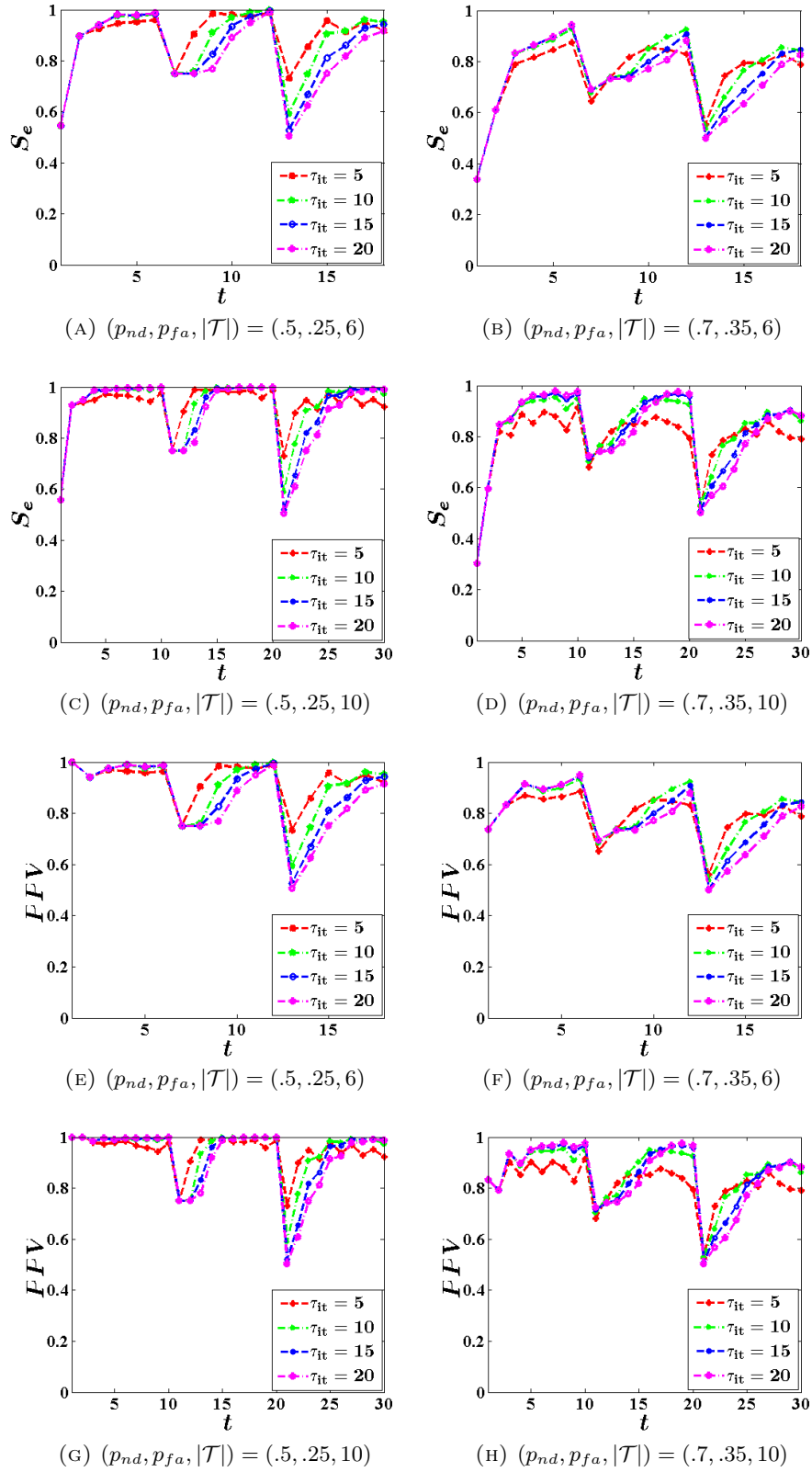


FIGURE 5.5: Indicateurs de performance (S_e et PPV) estimées sur des données simulées en utilisant le *BetP* comme critère de décision; lignes : (p_{nd}, p_{fa}) variable, $|\mathcal{T}|$ fixe; colonnes : (p_{nd}, p_{fa}) fixe, $|\mathcal{T}|$ variable.

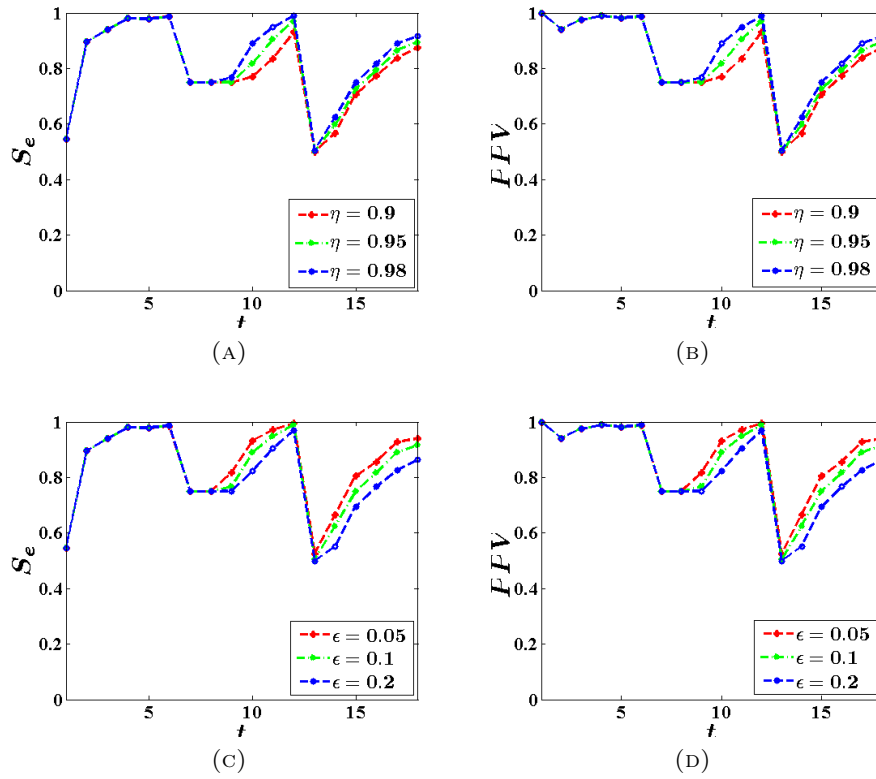


FIGURE 5.6: Indicateurs de performance (S_e et PPV) estimés sur des données simulées en utilisant le *BetP* comme critère de décision ; première lignes : η variable, $\epsilon = 0.1$; deuxième ligne : ϵ variable, $\eta = 0.98$, $(p_{nd}, p_{fa}) = (0.5, 0.25)$, $|\mathcal{T}| = 6$.

la scène même si la décision concernant les objets d'intérêt n'est pas prise directement à partir des masses mais plutôt en calculant le *BetP* ou le conflit induit par un *abridgment*. Dans cette figure, les cercles indiquent la fusion d'hypothèses qui sont reconnues comme étant des fragments d'un même objet physique. C'est le cas par exemple des hypothèses de couleurs 'rouge' et 'saumon', reconnues à l'instant $t=2$, en tant que sous parties de l'homme D et sont alors fusionnées à ce même instant. Un troisième fragment détecté à $t=8$ est reconnu fragment de l'homme D à $t = 10$ et fusionne alors avec les autres fragments. Notons également que la masse sur cet objet (homme D) décroît progressivement suite à son occultation partielle (son déplacement derrière une voiture) puis son occultation totale (sa disparition totale).

Les hypothèses associées aux trois personnes A , B et C sont représentées respectivement en 'cyan', 'vert' et 'rose'. L'hypothèse 'vert' a subi deux fusions (i) la partie inférieure de l'objet B vue initialement en tant qu'objet indépendant est fusionnée avec le reste de l'objet B à $t = 7$; (ii) la tête détectée à partir de l'instant $t=9$ est fusionnée à $t=39$.

La figure 5.8 illustre des exemples d'images parmi l'extrait de la séquence vidéo traitée.

La première colonne présente les données images en composition *RVB* des trois derniers instants. Les dernières détections sont présentées en traits gras et les détections antérieures en traits fins. La deuxième colonne illustre les résultats de l'estimation du cadre de discernement. De façon similaire à la première colonne, les traits gras indiquent les détections à t et les traits fins montrent l'accumulation des détections entre les instants t_0 et $t - 1$ de sorte que les traînées des anciennes détections illustrent les déplacements des objets. Pour ce cas traité, en utilisant les deux critères de décision ($BetP$ ou $m(\emptyset/H)$) les résultats sont les mêmes, ceux-ci sont les cadres de discernement Θ estimés à chaque instant de la séquence pour un nombre maximal d'hypothèses $M_{max} = 5$. Les objets qui correspondent à des hypothèses de Θ (à t) sont codés en couleur (rouge, vert, cyan, magenta) et les autres sont remplis en blanc. A titre d'exemple, dans la figure 5.8d, les carrés blancs correspondent à des éléments non inclus dans Θ et donc des fausses alarmes. Si nous interprétons les résultats en terme d'objets, nous notons les points suivants : (i) les deux sous-parties de l'objet vert fragmenté dans la figure 5.8b sont fusionnées dans la figure 5.8d, notons aussi dans cette même figure la présence d'une fausse alarme qui correspond à une lumière sur le châssis d'une voiture garée ; (ii) dans la figure 5.8f, le camion (objet E) sur la route au fond de l'image est détecté et prend la couleur orange ; (iii) dans la figure 5.8h, la personne D occultée par la voiture garée n'est plus détectée. Par ailleurs, les têtes des personnes A et B codées respectivement en cyan et vert ont été fusionnées à leurs corps.

Nous interprétons à présent les résultats en terme d'opérations dans le cadre de discernement Θ . Les remarques suivantes peuvent être faites : (i) la première opération est la fusion d'hypothèses, e.g. les deux sous parties de l'objet vert fragmenté dans la figure 5.8b sont reconnues appartenant au même objet et codées ainsi avec la même couleur ; (ii) la deuxième opération est l'ajout d'hypothèses dans Θ , e.g. la figure 5.8f illustre l'ajout de l'hypothèse qui correspond à l'objet codé en orange ; (iii) la troisième opération est l'élimination d'hypothèses ; e.g. cas de l'objet rouge qui n'étant plus détecté dès l'instant $t=31$ suite à son occultation par une voiture garée voit sa masse décroître progressivement (suite au processus d'oubli) pour atteindre une valeur si faible qu'il soit éliminé de Θ .

Finalement, notons que Ω garde toute l'information et permet de réintroduire des hypothèses écartées de Θ telles que la voiture sur la route, codée en blanc dans la figure 5.8f en tant que fausse alarme à ce niveau de fusion et en jaune dans la figure 5.8h en tant qu'objet d'intérêt.

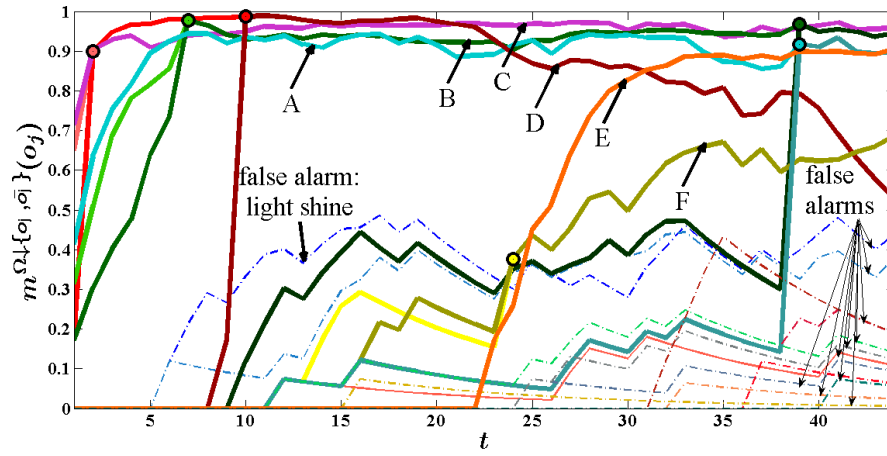


FIGURE 5.7: Données de *SafeAround* : évolution des masses sur les différents objets en fonction du temps illustrant des exemples de renforcement des croyances sur les objets persistants (cas des hypothèses *A*, *B* et *C*), affaiblissement des croyances sur les objets disparus (cas de l'hypothèse *D* et des *fausses alarmes*) et fusion des fragments des objets (cas des hypothèses *B* et *D*).

5.7 Conclusion

Ce chapitre a fait l'objet de deux contributions : une applicative et l'autre théorique. La première contribution concerne un système de surveillance de site pour une application de vidéosurveillance. Les images sont acquises à chaque instant et un algorithme de TI détecte les événements tels que les intrus appelés de manière plus générique objets, relativement à une image de fond. L'approche crédibiliste présentée dans le chapitre précédent nous fournit des objets construits à partir de détections fragmentaires, et le but de l'approche proposée dans ce chapitre est de filtrer ces objets construits (éliminer les fausses alarmes) et présenter à l'opérateur les menaces les plus pertinentes.

La deuxième contribution est l'estimation dynamique de l'ensemble des objets d'intérêt appelé cadre de discernement dans un formalisme évidentiel. Le cadre de discernement est supposé partiellement erroné en plus d'être incomplet. Ainsi, ses modifications possibles sont l'élimination de certaines hypothèses, l'ajout de nouvelles hypothèses et la fusion d'autres. Afin de robustifier la décision sur les modifications, un autre cadre de discernement sous forme d'un produit cartésien a été introduit pour se poser des questions sur la véracité de chacune des hypothèses potentielles.

Dans le chapitre suivant, nous illustrons la possibilité d'étendre notre approche de détection et d'énumération d'objets vers le suivi multi-objets.

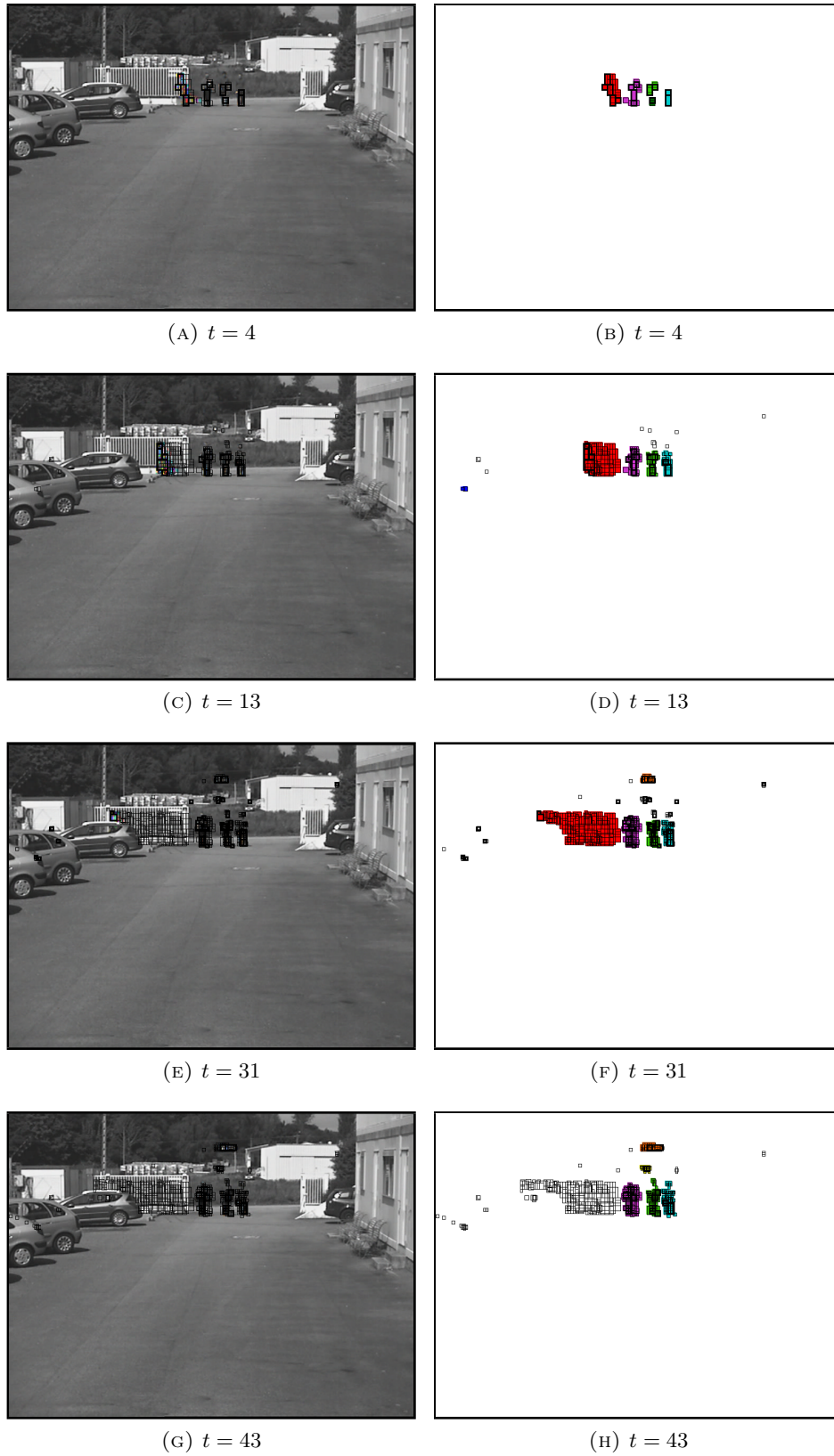


FIGURE 5.8: Données de *SafeAround* : visualisation des hypothèses de Θ projetées sur l'image à l'instant donné; 1^{ère} colonne : composition en *RVB* des détections des 3 derniers instants; 2^{ème} colonne : illustration des hypothèses de Θ (à t) en couleurs et des fausses alarmes en blanc.

Chapitre 6

Vers le suivi d'objets

6.1 Introduction

La méthode proposée dans cette thèse permet d'analyser une scène imagée en terme d'objets ou intrus. Elle est fortement liée au problème de suivi d'objets dans la mesure où elle effectue une partie de l'association de données (détections fragmentaires) mais en utilisant comme unique caractéristique la localisation des objets dans l'image. Elle est alors suffisante pour suivre des objets durant leurs déplacements à condition qu'ils n'interagissent pas. Cependant, elle diffère du problème de suivi dans le sens où elle ne gère pas la continuité des labels suite aux grandes occultations, fusion ou séparation d'objets. En effet, ni des connaissances a priori ni des contraintes sur l'initialisation d'un nouvel objet n'ont été prises en compte dans notre méthode : selon la notion de destruction-construction d'objets (cf. Chapitre 4), un objet se détruit suite à une occultation et se reconstruit en cas de réapparition. Toutefois, même s'il est construit de nouveau, l'objet perd son label.

Pour suivre des objets à partir de détections fragmentaires, sorties d'algorithmes de TI relativement bas niveau, nous proposons une approche à deux niveaux :

- un premier niveau développé dans cette thèse qui consiste à construire les objets d'intérêt présents dans une séquence vidéo. Ce problème s'est ramené (cf. Chapitre 4) à un problème d'association dès lors que les fragments associés entre eux, soit spatialement soit temporellement, forment les objets tandis que ceux non associés sont considérés comme des fausses alarmes, en supposant que la fiabilité d'un objet croît avec sa durée et sa taille (cf. Chapitre 5).

- un deuxième niveau qui assure un lien temporel entre les labels des objets en s'appuyant sur des caractéristiques plus riches (e.g. couleur [25], texture [143], gradient [144], flot optique [145, 146], points d'intérêt [147]) extraites sur les objets d'intérêt construits lors du premier niveau. Pour ce problème, plusieurs approches de suivi classiques existent dans la littérature. Elles sont soit déterministes (mise en correspondance entre les descripteurs des objets, e.g. histogramme [148], filtre de Kalman [149]), soit stochastiques (e.g. filtres particuliers [150]) ou encore des approches évidentielles (e.g. filtre de Kalman évidentiel [151], filtres évidentiels [152], classification évidentielle [153]). Cependant, dans tous les cas, la construction des nouveaux objets qui apparaissent dans la scène ou réapparaissent suite à une séparation d'un autre objet (e.g. un groupe) est une étape clé pour estimer de manière robuste des caractéristiques d'objets, durant leur suivi.

Même si dans ce chapitre nous présentons un résultat de suivi issu du couplage de notre méthode de construction d'objets d'intérêt avec un algorithme de suivi, le deuxième niveau de l'approche hiérarchique demeure un point clé à mieux investiguer et reste une perspective à court terme de nos travaux de thèse.

La suite de ce chapitre est organisée comme suit. Nous introduisons d'abord l'approche du suivi d'objets en utilisant un schéma global qui montre les liens entre les différentes étapes de l'approche. Nous présentons par la suite, l'algorithme de TI qui fournit les détections fragmentaires. Dans la troisième section, nous illustrons l'approche en présentant des résultats sur des données simulées et des données réelles.

6.2 Approche de suivi

La figure 6.1 illustre le schéma global du suivi d'objets qui contient quatre briques. À partir des détections fragmentaires (première brique), les objets sont construits (deuxième brique). Ces derniers sont l'entrée de la troisième brique qui estime l'ensemble des objets d'intérêt Θ . Finalement, la brique *tracking*, à un instant t , prend comme entrée les ensembles des objets d'intérêt estimés à $t - 1$ et à t notés respectivement Θ_{t-1} et Θ_t , et fournit la labellisation des objets d'intérêt.

La première brique fera l'objet de la section suivante concernant l'algorithme de TI qui fournit les détections fragmentaires utilisées dans la section 6.3. La deuxième brique et la

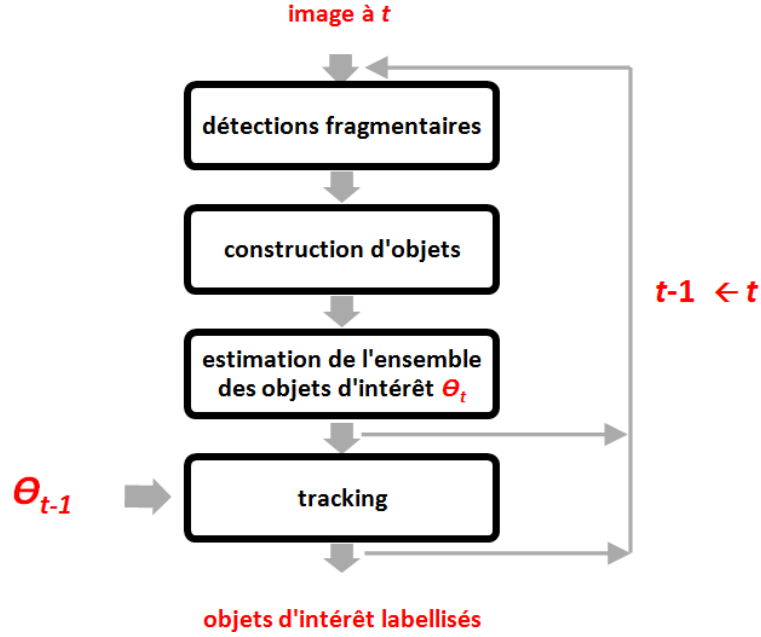


FIGURE 6.1: Schéma global du suivi d'objets.

troisième ont fait l'objet des deux chapitres 4 et 5 alors que le *tracking* est une perspective de notre travail.

6.2.1 Détections fragmentaires

Les approches pour la détection de changement sont nombreuses (cf. [154]). Le choix d'une méthode dépend des données traitées et de l'application. Cependant, dans la perspective de construire des objets en utilisant plusieurs images, les faux négatifs (non détections) sont privilégiés par rapport aux faux positifs (sur détections). Ainsi, dans notre travail, nous nous sommes concentrés sur des approches permettant de contrôler le nombre des fausses alarmes NFA (*number of false alarms*) c'est-à-dire utilisant un critère fondé sur le NFA. Plus particulièrement, les approches a-contrario ont été adoptées dans ce travail. Ces dernières ont été proposées par Desolneux [155] en TI et appliquées par suite dans plusieurs domaines tels que la détection de structure [156], la détection de changement dans des images de télédétection [157] ou encore dans des scènes vidéo [158]. Ces approches détectent des réalisations improbables d'un modèle non structuré, appelé 'modèle naïf'. En modélisant l'absence d'objets c'est-à-dire le bruit représenté par le modèle naïf et elles s'affranchissent de la modélisation des objets et ainsi de l'ajustement de paramètres. En outre, elles traitent le problème de détection (d'où des seuils à ajuster) en tant que problème d'optimisation [15].

Dans [15], le principe a-contrario permet de détecter des objets ne figurant pas dans une scène modèle. Celle-ci peut être une scène a priori (e.g. une scène vide ne contenant que la route) ou classiquement, une image de fond. Deux critères NFA sont utilisés successivement :

- Le premier permet de détecter le sous-ensemble de pixels (c'est-à-dire le sous-domaine de l'image) dont les valeurs des niveaux de gris sont les plus consistantes par rapport au fond. Le modèle naïf utilisé est un bruit gaussien. Ce critère fournit une image binaire B des pixels susceptibles d'appartenir aux objets recherchés.
- Le deuxième critère NFA effectue une mesure de signification sur un petit ensemble de pixels connectés (niveau 'fenêtre'). Il permet de détecter le sous-ensemble de fenêtres présentant le nombre le plus significatif des vrais pixels (de valeur égale à 1) relativement à B . Dans ce cas, le modèle naïf est une distribution binomiale traduisant l'absence de groupes de pixels dans B .

Dans ce travail, contrairement à [15] où l'image de fond est statique (e.g. première image de la séquence), nous avons utilisé plusieurs fonds. Durant la minimisation du premier critère NFA, en chaque pixel, le fond est choisi de sorte à minimiser l'erreur quadratique. La sélection et la mise à jour de l'image de fond sont effectuées en couplant le filtre sigma-delta [159] (cf. Section A.3.1.1) avec le modèle codebook [137] (cf. Section A.3.1.2). Ainsi, en minimisant le NFA successivement selon chaque modèle naïf, l'algorithme de détection fournit un ensemble non dense de détections (des fragments qui ne se chevauchent pas).

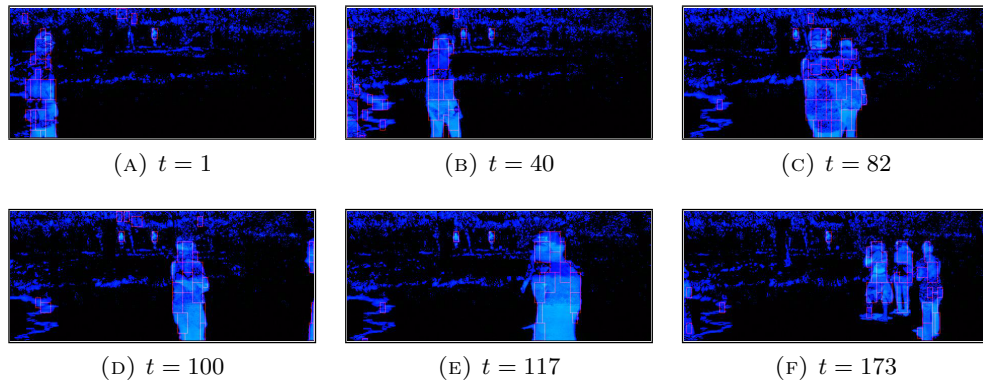


FIGURE 6.2: Résultat de détection issu de l'algorithme [15] et en actualisant l'image de fond, appliqué sur la séquence *Digiteo ComFuCet* (cf. Section 2.5.1) : le résultat de la détection au niveau pixel issu du premier critère NFA est représenté en bleu. Les détections fragmentaires (taille maximale égale à 10×20) issues du deuxième critère NFA sont en rouge.

6.3 Exemples de résultats

Nous illustrons à présent à travers des exemples de résultats (cf. Figures 6.4, 6.3) la possibilité d'étendre la méthode de construction d'objets d'intérêt (cf. Chapitre 4 et Chapitre 5) pour suivre des objets multiples c'est-à-dire la possibilité de maintenir les labels des objets après fusion et scission. L'idée principale est d'extraire des caractéristiques à partir des objets d'intérêt (cf. Section 6.1).

A chaque instant t , nous disposons d'un premier ensemble Θ_{t-1} , d'objets d'intérêt, estimé à $t-1$ et d'un nouvel ensemble Θ_t estimé à t . Il s'agit d'un problème d'association type 1-1 entre les ensembles Θ_{t-1} et Θ_t . Le coût d'association d'un objet de Θ_{t-1} avec un objet de Θ_t se déduit d'une mesure de similarité entre leurs descripteurs. Spécifiquement, nous considérons comme descripteur d'un objet, son histogramme normalisé des niveaux de gris. Dans ce cas, le coût d'associer un objet o_i de Θ_{t-1} d'histogramme h_i avec un objet o_j de Θ_t d'histogramme h_j est la distance Bhattacharyya [160] :

$$c_{ij} = 1 - \left(1 - \sum_{k=1}^b \sqrt{h_{ki} \times h_{kj}} \right)^{\frac{1}{2}}, \quad (6.1)$$

où b dénote le nombre de bins, h_{ki} et h_{kj} sont respectivement le $k^{\text{ème}}$ bin de h_i et le $k^{\text{ème}}$ bin de h_j . Par ailleurs, nous spécifions un coût de non association. Ainsi, un objet de Θ_t non associé initie un nouvel objet ayant un nouveau label. Il peut correspondre soit à un groupe d'objets (qui n'a pas de correspondant dans Θ_{t-1}) soit à un nouvel objet qui vient d'apparaître dans la scène. Finalement, le problème d'association est résolu de manière globale, en utilisant l'algorithme Hongrois (cf. Section A.1). L'approche de suivi est appliquée sur la séquence *Digiteo ComFuCet* (cf. Section 2.5.1) où les détections sont données par l'algorithme [15] et en actualisant l'image de fond (cf. Section 6.2.1). Notre approche d'estimation du cadre de discernement (cf. Chapitre 5) qui nous fournit à chaque instant l'ensemble Θ des objets d'intérêt est telle que le nombre maximal d'objets $M_{max} = 4$. Dans la figure 6.3, les fausses alarmes ne sont pas affichées. Elles sont représentées en noir dans la figure 6.4. Les objets d'intérêt (ceux appartenant à Θ) prennent différentes couleurs.

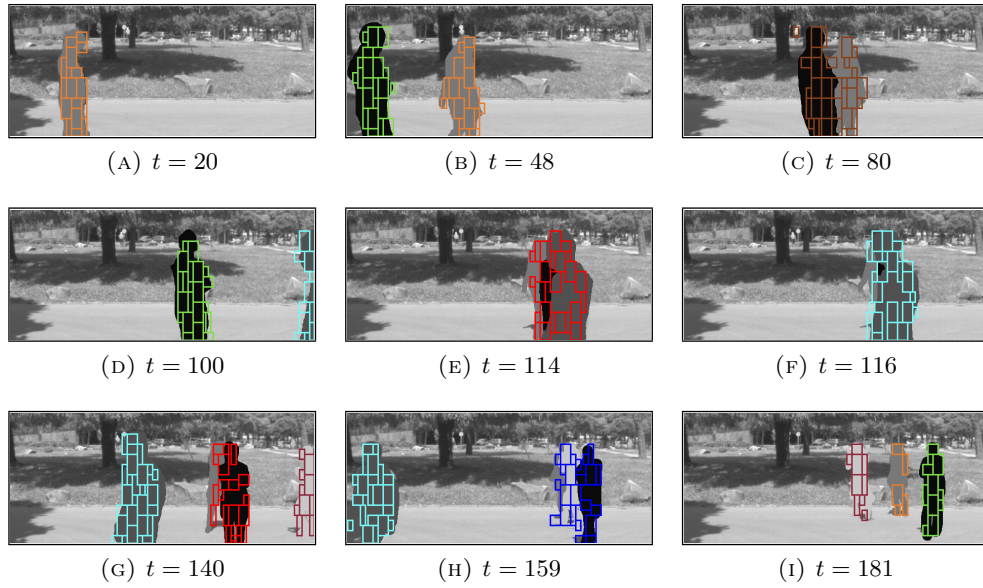


FIGURE 6.3: Exemple de résultat de suivi d'objets sur des données simulées. Les détections sont celles issues de l'algorithme [15] et en actualisant l'image de fond. Les différents objets d'intérêt prennent différents labels.

6.3.1 Résultats sur des données simulées

La figure 6.3 montre un exemple de résultat de suivi sur des données simulées. Ces dernières sont telles que chaque objet se voit attribuer une couleur unique de sorte que son histogramme d'intensité ne varie pas au cours du temps.

Cette figure montre qu'il est possible de garder les labels des objets après fusion et séparation. Nous prenons comme exemple la personne labélisée en cyan (cf. Figures 6.3f, 6.3g et 6.3h) qui a gardé son label après séparation du groupe d'objets labellisé en rouge dans la figure 6.3e, ainsi que les trois personnes labellisées en vert, orange et violet, fusionnées à $t = 159$ (groupe d'objets labellisé en bleu dans la figure 6.3h) et séparées à $t = 181$ (cf. Figure 6.3i).

6.3.2 Résultats sur des données réelles

La figure 6.4 présente un résultat de suivi sur des données réelles (les histogrammes des objets n'ont pas été modifiés) et illustre les points suivants :

- L'approche de construction (cf. Chapitre 4) fait face au problème de croisement d'objets. Ceci est illustré à travers les figures 6.4d, 6.4e, 6.4g. En effet, même si deux

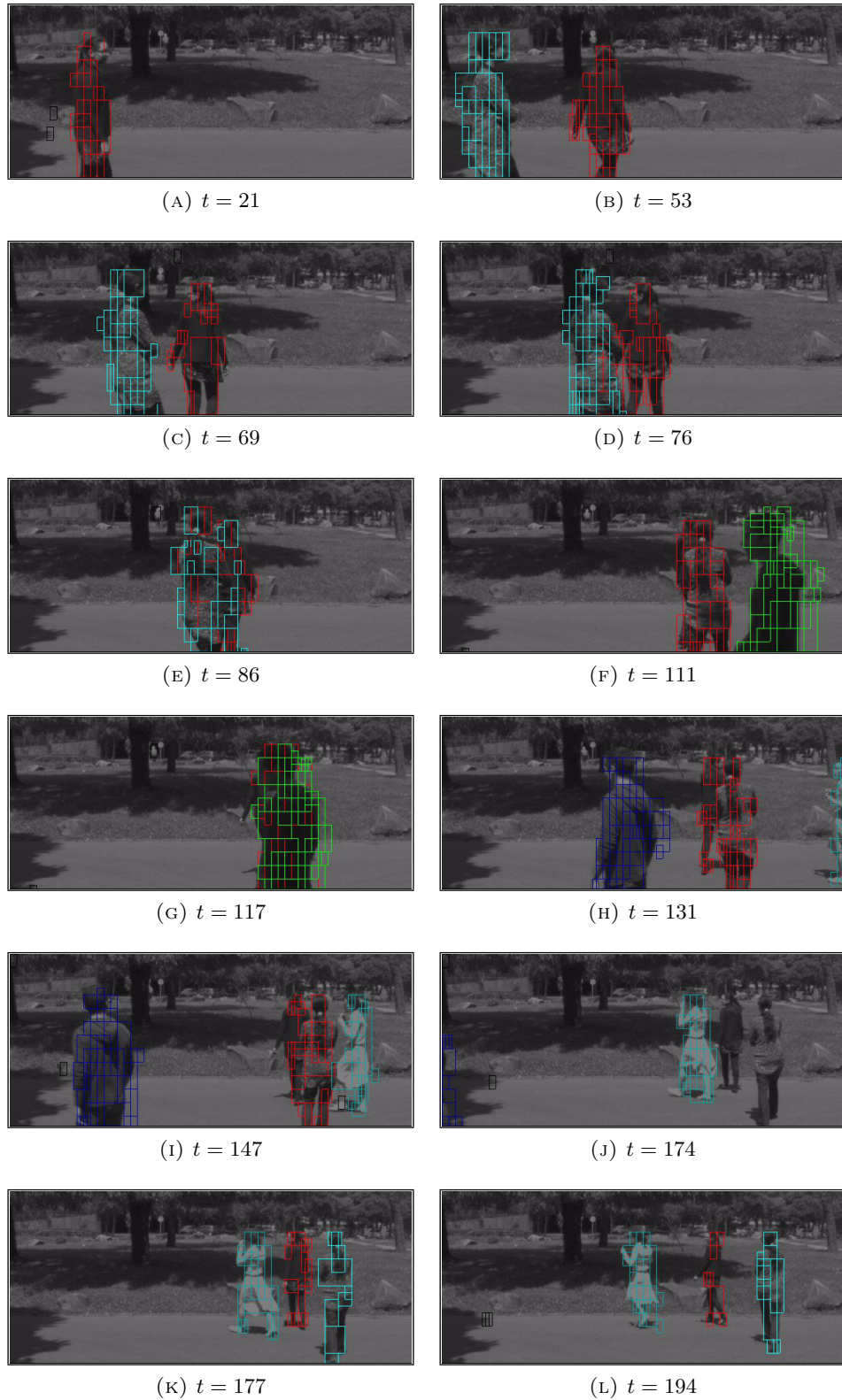


FIGURE 6.4: Exemple de résultat de suivi d'objets sur la séquence *Digitéo ComFu-Cet* (cf. Section 2.5.1). Les détections sont issues de l'algorithme [15] et en actualisant l'image de fond. Les fausses alarmes sont présentées en noir. Les différents objets d'intérêt prennent différents labels.

objets se croisent, leurs labels persistent : le rouge et le cyan foncé au niveau des figures 6.4d, 6.4e, le rouge et le vert au niveau de la figure 6.4g.

- La personne qui entre dans la scène avec le label vert (cf. Figure 6.4f) sort de la scène avec un label différent à savoir le bleu (cf. Figure 6.4h,6.4i). Par conséquent, l'algorithme de reconnaissance d'objets utilisant l'histogramme d'intensité n'arrive pas à reconnaître cet objet (vert) et lui attribue alors un nouveau label (bleu). Ceci s'explique par le changement de la luminosité au niveau de la séquence : cet objet passe de l'ombre (donc des niveaux de gris foncés) vers une région ensoleillée (présentant des niveaux de gris plus clairs).
- Les figures 6.4j et 6.4k illustrent la destruction-construction des objets. En effet, les trois personnes à droite de la scène dans la figure 6.4i ont été fusionnées à des instants ultérieurs. Dans la figure 6.4j, deux objets parmi les trois (les deux personnes à droite de la scène) se détruisent suite au conditionnement spatial (cf. Section 4.3.3) qui ne garde que la composante principale (ici l'objet de label cyan foncé). Les deux objets ainsi reconstruits dans la figure 6.4k ont bien gardé leurs labels (cyan foncé, rouge) grâce à l'algorithme de reconnaissance.
- L'aspect temporel (cf. Section 5.5.3) permet d'oublier un objet qui n'est plus détecté. C'est le cas de l'objet de label bleu à gauche de la scène dans la figure 6.4i et qui ne figure plus dans les figures 6.4k, 6.4l.

A la suite de ces remarques, nous pouvons conclure que, dans ce cas traité, l'histogramme des intensités n'est probablement pas la caractéristique la plus adéquate pour discerner les objets entre eux. En effet, la séquence est acquise à l'extérieur et présente des difficultés notamment en terme de changement brusque de luminosité. Dans un tel cas, le HOG peut être une alternative.

6.4 Conclusion

Ce chapitre montre l'intérêt de la construction d'objets et l'estimation de l'ensemble des objets d'intérêt, dans la résolution du problème de suivi multi-objets. D'une part, le suivi ne peut pas se produire au niveau fenêtre (c'est-à-dire directement à partir des détections fragmentaires) mais plutôt au niveau objet, d'où l'utilité de la construction d'objets à partir des détections. D'autre part, l'estimation de l'ensemble des objets à suivre en éliminant les fausses alarmes peut limiter le temps de calcul. Cependant, la

dernière étape qui assure le suivi des objets en cas d'occultations est une perspective de notre travail. Toutefois, nous avons présenté dans ce chapitre, des exemples de résultats de suivi sur des données simulées ainsi que sur des données réelles, où l'algorithme de *tracking* est fondé sur la mise en correspondance des histogrammes des intensités.

En utilisant les données simulées, l'algorithme de *tracking* réussit à maintenir les labels des objets après occultations. Cependant, les résultats sur les données réelles sont moins satisfaisants. Ceci s'explique par les variations brusques des intensités (séquence acquise à l'extérieur). Ainsi, nous proposons dans nos prochains travaux d'utiliser d'autres descripteurs tels que le HOG et de tester l'algorithme sur des séquences acquises à l'intérieur, qui sont probablement moins complexes que celle étudiée.

Chapitre 7

Conclusion générale

Pour faire le bilan de ces travaux de thèse et avant d'en énoncer les perspectives, rappelons le contexte et les choix méthodologiques effectués.

7.1 Bilan des travaux

Notre travail s'est situé dans le cadre d'une application réelle de vidéosurveillance. Typiquement, il pourrait s'agir d'un système de protection sur un site présentant des zones interdites pour différents types d'intrus typiquement les véhicules, les individus et les groupes.

Dans notre travail, nous avons fait l'hypothèse qu'une menace est d'autant plus importante que l'intrus associé est de taille importante (dans l'image donc soit effectivement plus grand, soit, à taille égale, plus proche) et temporellement persistant. Nous avons utilisé la redondance temporelle pour, d'une part, distinguer les fausses alarmes, décorréées dans le temps, et évaluer la persistance temporelle d'un intrus (typiquement, un intrus peut 'par erreur' pénétrer sur le site).

La seconde hypothèse de travail prise est la découpe fonctionnelle de notre système de vidéosurveillance. En se restreignant à la partie traitement du flux vidéo, nous nous sommes concentrés sur le cas où la détection et le suivi d'objets se font disjointement. En effet, du point de vue des systèmes complexes, le fait d'avoir un système modulaire permet de définir un certain nombre de briques répondant à des sous-problèmes pour

lesquels des solutions pourront être apportées séparément. Nous avons postulé qu'une de ces briques concerne la détection de fragments correspondant potentiellement à des objets.

Notre troisième hypothèse de travail est le choix du type d'algorithme de détection. Précisément, nous avons choisi des algorithmes de TI (Traitement d'Images) de détection se situant entre le niveau pixel et le niveau objet, compromis entre les algorithmes de détection niveau pixel (bas niveau et/ou aveugle) pouvant générer de nombreuses fausses alarmes et les algorithmes de segmentation (tels que le Mean-Shift, Graph-Cuts, contour actif) ainsi que ceux de la classification supervisée nécessitant des connaissances a priori sur les objets (telles que la forme, texture, couleur, etc.). Les algorithmes utilisés n'étant pas pour autant exempts de défauts, ils présentent généralement un aspect fragmentaire pour les objets que nous traitons en terme de construction de l'objet, aspect qui peut exister avec les autres approches également (e.g. en cas d'occultation partielle des objets) et que nous aurions dû traiter de toutes façons.

Notre quatrième hypothèse est que la seule information disponible à partir de ces détections fragmentaires (fenêtres de changement) est leur localisation spatiale dans l'image. Elle est donc l'unique information que nous avons exploitée pour les sous-problèmes de représentation, association et filtrage d'objets. Cependant, une telle information est insuffisante pour résoudre des problèmes d'occultations (lorsque des objets ont la même localisation spatiale dans l'image 2D). Ainsi, pour étendre ce travail vers le suivi multi-objets c'est-à-dire résoudre le problème supplémentaire de reconnaissance d'objets occultés, nous avons cherché à utiliser des informations plus haut niveau (telles que la couleur, texture, trajectoire, etc.) extraites à partir des objets construits (mais ce travail reste principalement une perspective).

Etant données ces contraintes pour notre problème de surveillance, nos contributions ont été les suivantes :

- Représentation des objets en termes de localisation imprécise et incertaine. Du point de vue de la fusion et l'intelligence artificielle, la représentation pertinente des informations est un point essentiel pour les problèmes d'estimation ou la prise de décision. Une bonne représentation se reconnaît au fait qu'en découlent des critères simples et

- performants pour résoudre des sous-problèmes. Ainsi dans notre cas, la représentation proposée a été valorisée par le fait qu'un critère d'association (entre nouvelles détections et objets en construction) a pu être défini d'une façon simple et rigoureuse.
- Filtrage des objets en termes d'estimation dynamique du cadre de discernement. De ce point de vue, les mises à jour du cadre de discernement sont liées à la détection de deux principaux types d'erreurs : la duplication de certains objets en raison de leur caractère fragmentaire, les fausses alarmes dues au bruit inhérent aux données.

7.2 Perspectives

Nos travaux de thèse ouvrent trois perspectives dont l'une est à court terme et concerne le *tracking* d'objets. La deuxième découle de l'a priori sur l'indépendance entre les hypothèses du cadre de discernement c'est-à-dire les objets. La troisième porte sur la fiabilité des associations à considérer lors de l'estimation de l'ensemble des objets d'intérêt.

- Le but du *tracking* est d'estimer la trajectoire de chaque objet présent dans la scène c'est-à-dire maintenir son label au cours du temps. Cependant, ceci est généralement difficile en présence d'occultations. Ainsi, certains algorithmes de *tracking* se fondent sur des caractéristiques riches extraites à partir des objets telles que la couleur, la texture, la forme, etc. D'autres utilisent des filtres de prédiction tels que le filtre de Kalman et le filtre particulaire. Ceux-ci, étant donnés un modèle dynamique (linéaire ou non linéaire) décrivant l'objet à suivre, un vecteur d'état estimé antérieurement et de nouvelles observations, prédisent l'état de l'objet à l'instant actuel. L'état de l'objet peut concerner sa localisation dans l'image.

L'approche d'analyse de scène proposée dans cette thèse assure le suivi d'objets dans le cas où les objets n'interagissent pas c'est-à-dire en absence d'occultations. Dans le cas contraire, un objet occulté se détruit (perd alors son label) et se reconstruit en ayant un nouveau label, en cas de réapparition. Ainsi, afin de gérer le problème d'occultation, notre perspective pour ce travail est de le coupler avec un algorithme de *tracking* pour assurer le suivi des objets construits à partir des détections fragmentaires, en utilisant soit un modèle de prédiction soit une mise en correspondance entre les caractéristiques des objets (couleur, texture, etc.).

- Lors de l'estimation de l'ensemble des objets d'intérêt, nous avons supposé l'indépendance totale entre les objets (hypothèses du cadre de discernement). En effet, dans

notre application de détection d'objets, la présence d'un objet (e.g. homme, voiture, chien, etc.) n'informe pas sur la présence ou non d'un autre objet. Cette indépendance est une condition forte pour l'utilisation du produit cartésien dont le but est de robustifier la décision sur les modifications du cadre de discernement.

Toutefois, dans d'autres applications telles que la classification multi-label de scènes, il pourrait y avoir une corrélation entre certaines hypothèses : (i) une corrélation positive lorsque certaines hypothèses sont susceptibles de se produire simultanément (e.g. les labels 'voiture' et 'route') ; (ii) une corrélation négative lorsque la présence d'une hypothèse réfute la présence d'une autre (e.g. les labels 'urbain' et 'tigre'). Par conséquent, dans les prochains travaux nous visons à considérer la corrélation éventuelle entre les hypothèses lors de l'estimation du cadre de discernement.

- Les associations entre les nouvelles détections et les objets définissent les relations entre les hypothèses (objets) du cadre de discernement. Elles sont supposées suffisamment fiables. Or, une association peut être erronée ce qui induit l'augmentation erronée de la croyance sur une hypothèse au détriment d'une autre. Ainsi, une perspective pour ce travail est d'estimer la fiabilité de chaque association et la considérer lors de l'estimation du cadre de discernement.

Annexe A

Annexe au chapitre 3

A.1 Algorithme Hongrois

L'algorithme Hongrois [49], [121] repose sur le théorème suivant :

Théorème 1. Soit M_c de taille $n \times n$ une matrice de coûts. Le classement des permutations est inchangé si un même nombre est ajouté ou supprimé de tous les éléments d'une ligne ou d'une colonne de M_c . De telles modifications n'altèrent pas l'association optimale, mais uniquement le coût optimal. En effet, elles permettent de naviguer entre des matrices correspondant à des problèmes équivalents.

Il itère en transformant la matrice de coûts en une suite de matrices équivalentes, jusqu'à l'obtention d'une solution optimale. La matrice finale est telle que toutes ses valeurs sont soit positives soit nulles. Par ailleurs, une association faisant appel seulement aux valeurs nulles est possible. Cette association de coût nul est alors nécessairement optimale. Dans ce qui suit, nous décrivons les étapes essentielles de l'algorithme Hongrois et nous les illustrons au travers des deux exemples qui suivent.

Trois étapes sont à itérer jusqu'à l'obtention de n zéros indépendants c'est-à-dire n zéros tels qu'un 0 par ligne et par colonne :

1. éliminer le plus petit élément de chacune des lignes (respectivement colonnes) de M_c afin de transformer le problème d'association vers un problème équivalent décrit par une matrice de coûts ayant un 0 par ligne et par colonne.

2. sélectionner le nombre maximal de 0 indépendants sachant que le problème d'association se résout en obtenant n zéros indépendants. Pour ce faire, deux sous étapes sont nécessaires :
 - considérer la ligne ayant un nombre minimal de 0. Marquer l'un des 0 de cette ligne et barrer les 0 qui se trouvent sur la même ligne ou la même colonne que le 0 marqué. Répéter tant qu'il existe des 0 non traités (marqués ou barrés).
 - cocher '*' : (i) les lignes qui ne contiennent aucun 0 marqué ; (ii) les colonnes qui ont un ou plusieurs 0 barrés dans une ligne * ; (iii) les lignes qui ont un 0 marqué dans une colonne *.
3. considérer de nouveau un problème équivalent :
 - barrer les lignes non * ainsi que les colonnes *.
 - soit la valeur minimale dans la matrice réduite M'_c . Retirer cette valeur des éléments de M'_c et l'ajouter aux éléments 'doublement' éliminés de $\mathcal{K}$ (la valeur est retirée de tous les éléments de M_c et ajoutée aux lignes de $M_c \setminus M'_c$ puis aux colonnes de $M_c \setminus M'_c$).

Exemple 1. (cf. Figure A.1) Dans cet exemple illustratif, nous considérons un problème d'affectation tâches-employés où chaque employé ne peut accomplir qu'une seule tâche et chaque tâche ne peut être exécutée que par un seul employé. L'affectation parfaite des tâches aux employés est celle qui vérifie les contraintes (association bijective) et minimise la somme des coûts locaux. La matrice de coûts est donnée par la figure A.1a. La matrice équivalente présentée dans la figure A.1i montre une affectation de coût nul. Cette affectation est alors optimale : l'employé 1 exécute la tâche 2, l'employé 2 exécute la tâche 5, l'employé 3 exécute la tâche 1, l'employé 4 exécute la tâche 3 et l'employé 5 exécute la tâche 4. Son coût, faisant appel aux coûts locaux originaux, est égal à $3+7+5+4+4=23$.

Exemple 2. (cf. Figure A.2) Dans cet exemple, nous considérons de nouveau un problème d'affectation tâches-employés et nous supposons avoir trois tâches et seulement deux employés. Ainsi, la matrice de coûts est rectangulaire. Du fait que l'algorithme Hongrois ne s'applique que sur des matrices carrées, nous créons un troisième employé fictif qui s'affecte automatiquement à la tâche non accomplie par les deux véritables employés. Nous lui attribuons un coût fixe $c = 0$. Par ailleurs, nous supposons que l'employé 1 ne peut pas exécuter la tâche 1 ce qui se traduit par un coût $c(1,1) = k$ assez élevé. Dans les figures A.2a et A.2b, nous présentons respectivement, la matrice de coûts initiale

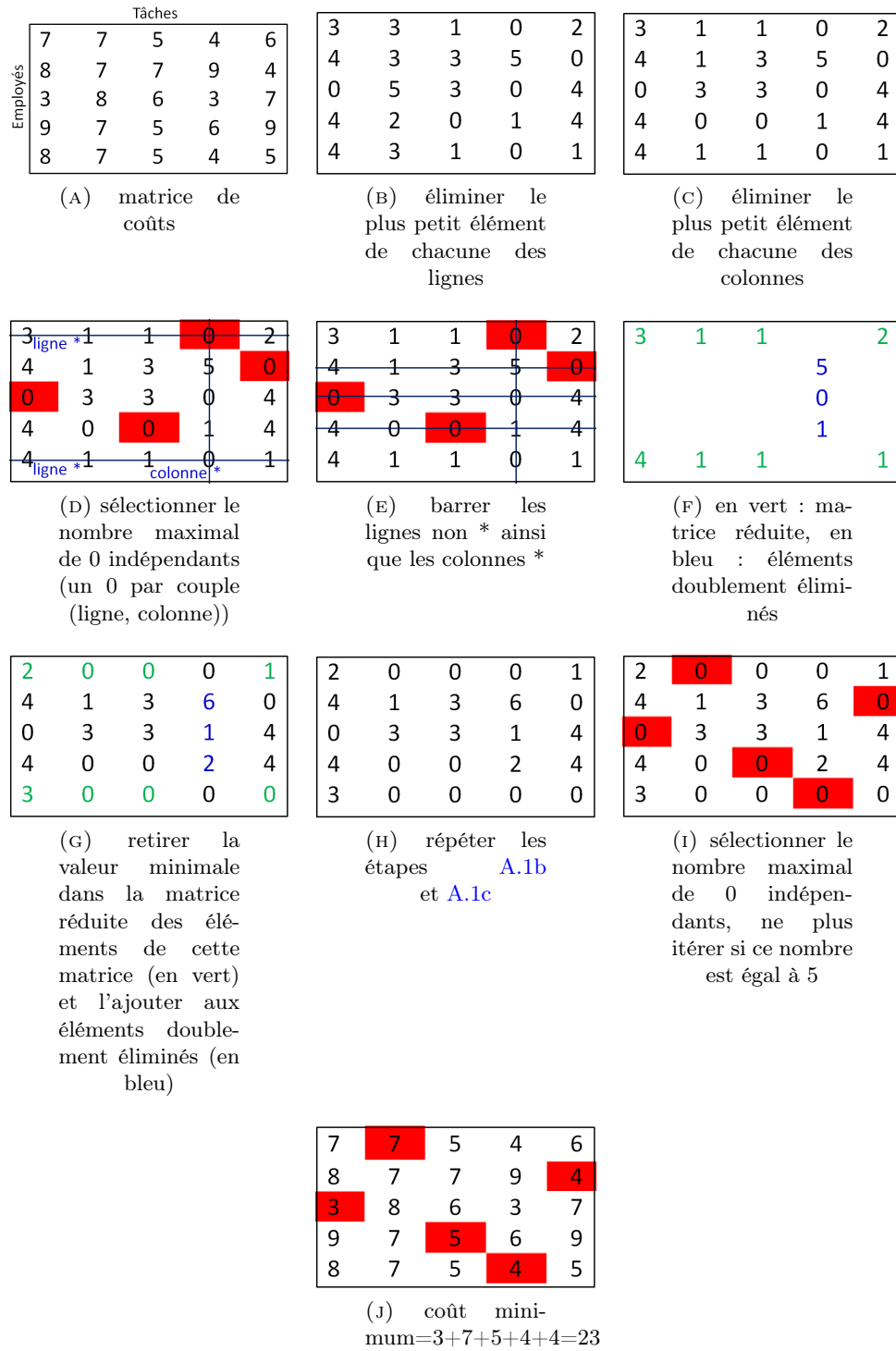


FIGURE A.1: Premier exemple illustrant l'algorithme Hongrois.

et la matrice de coûts étendue. Selon l'affectation optimale (cf. Figure A.2i) l'employé 1 exécute la tâche 2 et l'employé 2 exécute la tâche 3 alors que la tâche 1 reste non affectée. Le coût de cette affectation est égal à $8+4+0=12$ (cf. Figure A.2j).

Dans les deux exemples, l'affectation optimale est obtenue au bout d'une itération. Cependant, nous mentionnons qu'il est parfois nécessaire d'itérer plusieurs fois pour faire converger l'algorithme vers la solution optimale.

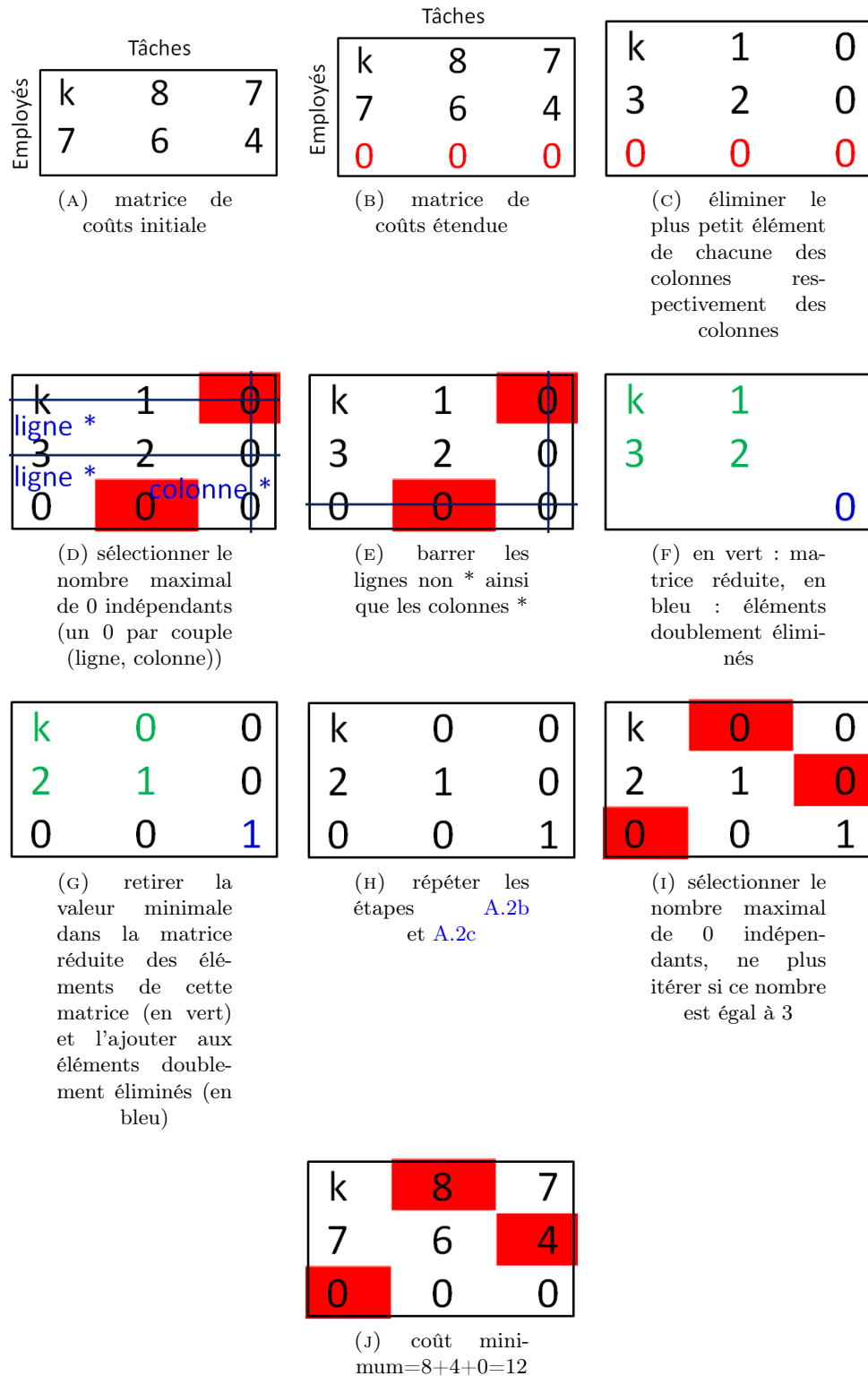


FIGURE A.2: Deuxième exemple illustrant l'algorithme Hongrois.

A.2 Modification de l'algorithme Hongrois

La version modifiée de l'algorithme Hongrois considère une matrice de coûts M_c symétrique. Elle utilise le même principe que l'algorithme Hongrois sauf qu'elle préserve la symétrie de la matrice de coûts au cours de ses transformations en matrices équivalentes : les opérations effectuées sur les lignes ou les colonnes sont effectuées de manière symétrique sur les colonnes ou les lignes. L'algorithme est le suivant :

- étape 1 : éliminer le plus petit élément de chaque ligne, et de la colonne de même indice.
- étape 2 : sélectionner les zéros indépendants. Si la dimension de la matrice est impaire, nous imposons de choisir comme 'zéro' la valeur minimale sur la diagonale. Si le nombre de zéros indépendants est égal à la dimension de la matrice alors ces zéros définissent la solution de coût minimum, sinon nous passons à l'étape 3 :
- considérer la ligne ayant un nombre minimal de 0. Marquer l'un des 0 de cette ligne et barrer les 0 qui se trouvent sur la même ligne ou la même colonne que le 0 marqué. Répéter tant qu'il existe des 0 non traités (marqués ou barrés).
- cocher '*' : (i) les lignes qui ne contiennent aucun 0 marqué ; (ii) les colonnes qui ont un ou plusieurs 0 barrés dans une ligne * ; (iii) les lignes qui ont un 0 marqué dans une colonne *.
- barrer les lignes non * ainsi que les colonnes *.
- créer la matrice M présentant des 0 sur les lignes barrées et les colonnes non barrées ou sur les colonnes barrées et lignes non barrées, des 1 sur les lignes et les colonnes barrées, des -1 sur le reste de la matrice.
- symétriser M en prenant le maximum entre un élément et son symétrique.
- $M_c \leftarrow M_c + M$.
- répéter jusqu'à l'obtention de la solution de coût minimum.

La figure [A.3](#) illustre l'algorithme Hongrois modifié sur un cas simple.

A.3 Détection de changement

La détection de changement est la détection d'objets présents dans une scène imagée par rapport à un modèle à t qui est soit une image de fond (en anglais *background* et signifie l'image en l'absence d'objet d'intérêt) soit une image précédente dans la séquence.

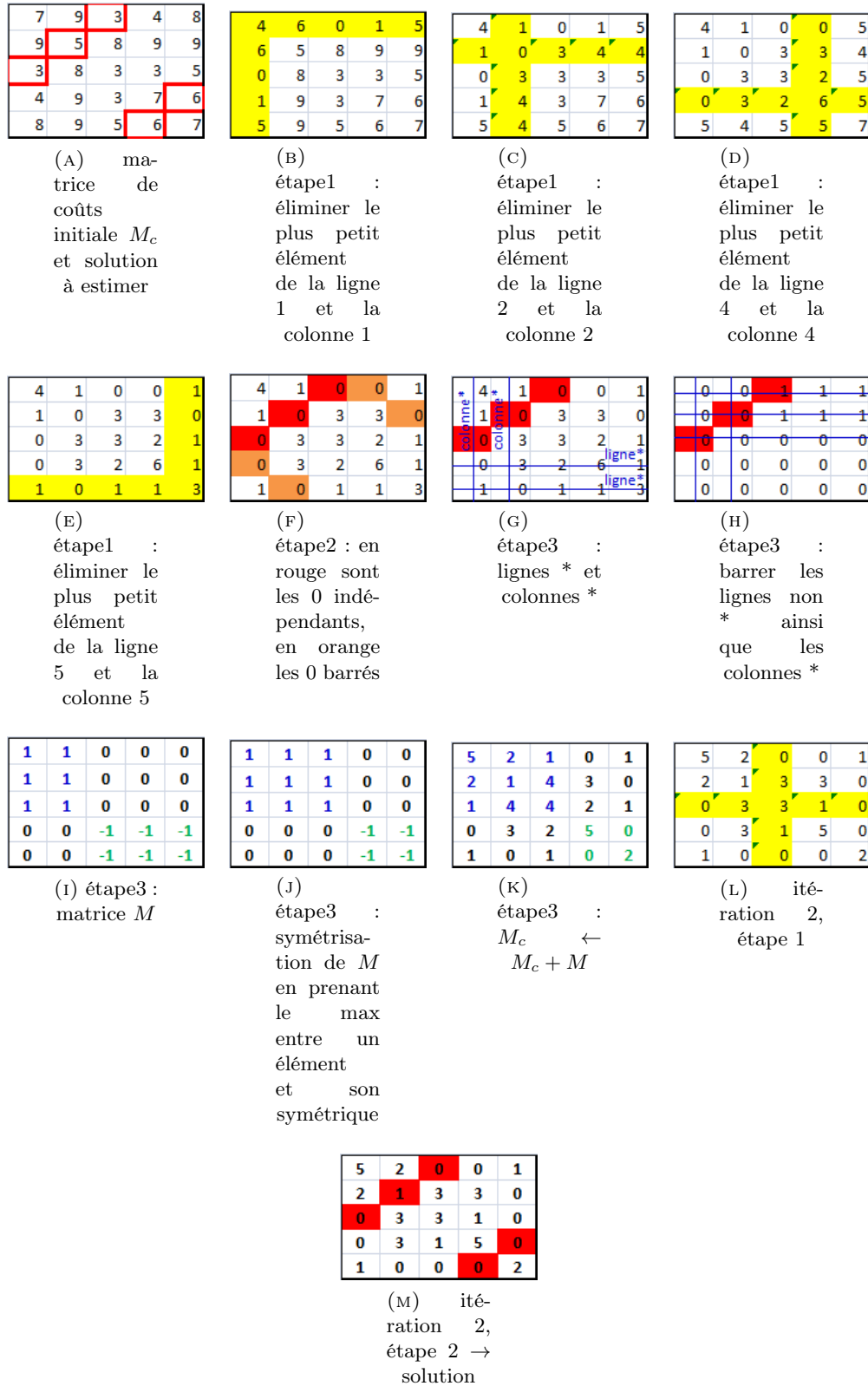


FIGURE A.3: Exemple illustrant l'algorithme Hongrois modifié.

Précisons aussi que les méthodes de détection de changement diffèrent de celles de reconnaissance qui visent à identifier des objets connus dans une scène inconnue.

En détection de changement, nous distinguons principalement deux types d'approches : (i) des approches qui visent à suivre un objet. Ainsi, pour détecter un objet dans l'image courante, des informations sur l'objet telles que sa position, sa couleur, sa forme etc. sont nécessaires ; (ii) des approches qui détectent des changements entre deux images. Ce type d'approches auquel nous nous sommes intéressés ne demande que l'information issue de l'image de fond. Donc, une première question basique est comment estimer l'image de fond, une deuxième est comment détecter les différences entre l'image de fond et l'image courante ?

A.3.1 Estimation de l'image de fond

L'image de fond peut simplement être choisie parmi les images de la séquence qui ne contiennent aucun objet à détecter. Cependant, dans la pratique, ceci est généralement insuffisant. En effet, les scènes réelles présentent des structures et des configurations de paysages pouvant varier suite à un changement de positions des objets statiques ou bien aux conditions d'acquisition (luminosité, nuage, etc.). Par conséquent, plusieurs travaux se sont focalisés sur le développement de techniques robustes d'actualisation de l'image de fond qui consistent à assimiler progressivement les changements dans cette image. Parmi ces techniques, nous présentons les plus connues à savoir le filtre sigma-delta et le modèle codebook [137]. Ces techniques ont été utilisées dans les algorithmes de TI cités dans la section 4.8.

A.3.1.1 Filtre sigma-delta

Ce filtre [159] est à la fois simple et efficace. Il consiste à modéliser chaque pixel i de l'image de fond par une moyenne m_i et un bruit s_i qui les fait évoluer en fonction des nouvelles observations p_i^t selon la démarche suivante :

- actualisation de m_i :
 - si $m_i > p_i^t + \delta$ alors $m_i = m_i - \delta$,
 - si $m_i < p_i^t - \delta$ alors $m_i = m_i + \delta$,
 - sinon $m_i = p_i^t$.
- actualisation de s_i :
 - si $s_i > |p_i^t - m_i|$ alors $s_i = s_i - 1$,
 - si $s_i < |p_i^t - m_i|$ alors $s_i = s_i + 1$.

avec δ un paramètre traduisant la réactivité du filtre face au changement. Ce modèle malgré sa simplicité est très robuste et ne dépend que du paramètre δ qui prend généralement la valeur 1. Cependant, son inconvénient majeur est son assimilation rapide des changements. Prenons l'exemple d'un objet en mouvement qui s'arrête provisoirement dans la scène alors ce dernier sera rapidement assimilé dans l'image de fond et ne sera plus détecté en tant que changement.

A.3.1.2 Modèle codebook

Ce modèle [137] a été conçu pour traiter des images comportant des zones ne pouvant être modélisées par une seule valeur. C'est le cas par exemple de la végétation qui peut brusquement changer de couleur à cause du vent ou de la luminosité. Ainsi, un pixel i de l'image de fond n'est pas modélisé par une seule valeur (cas du filtre sigma-delta) mais plutôt par plusieurs triplets (m_i^j, s_i^j, c_i^j) ainsi qu'un triplet (m_i, s_i, c_i) pour l'apprentissage. Ces derniers sont actualisés selon les observations p_i^t acquises à chaque instant t comme suit :

- chercher le triplet j dont la moyenne m_i^j minimise $|m_i^j - p_i^t|$:
 - si $|m_i^j - p_i^t| < \gamma s_i^j$ alors actualiser m_i^j et s_i^j selon le filtre sigma-delta et $c_i^j = t$,
 - si $|m_i - p_i^t| < \gamma s_i^j$ alors mettre à jour le triplet d'apprentissage selon le filtre sigma-delta et $c_i = c_i + 1$,
 - sinon réinitialiser le triplet d'apprentissage : $m_i = p_i^t$, $s_i = s_{init}$ et $c_i = 1$.
- parcourir les triplets (m_i^j, s_i^j, c_i^j) , si $c_i^j < t - dt$ alors supprimer le triplet j .
- si $c_i > c_{min}$ alors ajouter le triplet d'apprentissage au modèle.

Ce modèle fait intervenir trois paramètres à savoir :

- γ : seuil sur la différence entre une observation p_i^t et un triplet du modèle.
- dt : délai minimum pour supprimer un triplet non observé.
- c_{min} : nombre minimum d'observations pour ajouter un triplet au cours d'apprentissage au modèle.

En revanche, l'inconvénient majeur de ce modèle est sa lourdeur en terme de temps de calcul.

A.3.2 Détection des différences par rapport à l'image de fond

L'algorithme de TI cité dans la section 4.8 en tant que variante de [137] a été développé dans le cadre des travaux de la thèse [123]. Nous présentons dans ce qui suit l'idée générale et nous référons à [123] pour plus de détails.

Pour détecter les différences entre l'image courante et l'image de fond, [123] distingue deux niveaux de changement à savoir le niveau pixel et le niveau fenêtre. Le niveau pixel est intuitif. En effet, pour chaque pixel, un modèle du fond et une valeur courante sont définis. Ainsi, évaluer si un pixel i de l'image courante est un pixel de changement repose sur un seuillage sur la valeur $|m_i - p_i^t|$. Si cette valeur dépasse un certain seuil alors i est un pixel de changement.

Selon [123], le niveau fenêtre est plus robuste que le niveau pixel. En effet, il s'agit de comparer l'image courante et l'image de fond non pas à l'échelle d'un pixel mais plutôt à l'échelle d'une fenêtre de taille supérieure au pixel. Par conséquent, des zones de changements de tailles inférieures à celle d'une fenêtre peuvent être détectées. L'avantage de ce niveau est qu'un changement radiométrique local mais significatif à l'intérieur d'une fenêtre est susceptible de changer la statistique de la fenêtre de manière à la détecter en tant que zone de changement.

Annexe B

Annexe au chapitre 4

B.1 Démonstration de l'équation 5.7

Dans cette section, nous montrons l'égalité donnée par l'équation 5.7. Comme la démonstration est quasiment la même pour toute dimension N , nous nous limitons au cas $N = 2$. Pour $N = 2$, $\Omega = \{o_1, \bar{o}_1\} \times \{o_2, \bar{o}_2\}$, $2^{\Omega_1} = \{\emptyset, o_1, \bar{o}_1, \{o_1, \bar{o}_1\}\}$ et $2^{\Omega_2} = \{\emptyset, o_2, \bar{o}_2, \{o_2, \bar{o}_2\}\}$. Soient $m_{11} = (m_1^{\downarrow\{o_1, \bar{o}_1\}})^{\uparrow\Omega}$ et $m_{21} = (m_2^{\downarrow\{o_1, \bar{o}_1\}})^{\uparrow\Omega}$. Puisque les éléments focaux de $m_1^{\downarrow\{o_1, \bar{o}_1\}}$ et $m_2^{\downarrow\{o_1, \bar{o}_1\}} \in 2^{\Omega_1}$, alors les éléments focaux de m_{11} et $m_{21} \in \{(\emptyset, \{o_2, \bar{o}_2\}), (o_1, \{o_2, \bar{o}_2\}), (\bar{o}_1, \{o_2, \bar{o}_2\}), (\{o_1, \bar{o}_1\}, \{o_2, \bar{o}_2\})\}$. Calculons le résultat de $m_{11} \oplus m_{21}$ dont les éléments focaux sont de la forme $(\omega_1, \{o_2, \bar{o}_2\})$, avec $\omega_1 \in 2^{\Omega_1}$:

$$\begin{aligned}
 \forall \omega_1 \in 2^{\Omega_1}, \quad & (m_{11} \oplus m_{21})((\omega_1, \{o_2, \bar{o}_2\})) \\
 &= \sum_{\substack{(A,B) \in 2^{\Omega_1} \times 2^{\Omega_1} \\ A \cap B = \omega_1}} m_{11}(A, \{o_2, \bar{o}_2\}) m_{21}(B, \{o_2, \bar{o}_2\}) \\
 &= \sum_{\substack{(A,B) \in 2^{\Omega_1} \times 2^{\Omega_1} \\ A \cap B = \omega_1}} m_1^{\downarrow\{o_1, \bar{o}_1\}}(A) m_2^{\downarrow\{o_1, \bar{o}_1\}}(B) \\
 &= \left(m_1^{\downarrow\{o_1, \bar{o}_1\}} \oplus m_2^{\downarrow\{o_1, \bar{o}_1\}} \right) (\omega_1)
 \end{aligned} \tag{B.1}$$

Exemple : $m_{11} \oplus m_{21}((\emptyset, \{o_2, \bar{o}_2\})) = m_1^{\downarrow\{o_1, \bar{o}_1\}}(\emptyset) + m_2^{\downarrow\{o_1, \bar{o}_1\}}(\emptyset) - m_1^{\downarrow\{o_1, \bar{o}_1\}}(\emptyset) m_2^{\downarrow\{o_1, \bar{o}_1\}}(\emptyset) + m_1^{\downarrow\{o_1, \bar{o}_1\}}(o_1) m_2^{\downarrow\{o_1, \bar{o}_1\}}(\bar{o}_1) + m_1^{\downarrow\{o_1, \bar{o}_1\}}(\bar{o}_1) \times m_2^{\downarrow\{o_1, \bar{o}_1\}}(o_1)$.

Par ailleurs, $\forall \omega_1 \in 2^{\Omega_1}$,

$$\left(m_1^{\downarrow\{o_1, \bar{o}_1\}} \odot m_2^{\downarrow\{o_1, \bar{o}_1\}}\right)(\omega_1) = \left(m_1^{\downarrow\{o_1, \bar{o}_1\}} \odot m_2^{\downarrow\{o_1, \bar{o}_1\}}\right)^{\uparrow\Omega}((\omega_1, \{o_2, \bar{o}_2\})). \quad (\text{B.2})$$

Par conséquent :

$$(m_1^{\downarrow\{o_1, \bar{o}_1\}})^{\uparrow\Omega} \odot (m_2^{\downarrow\{o_1, \bar{o}_1\}})^{\uparrow\Omega} = (m_1^{\downarrow\{o_1, \bar{o}_1\}} \odot m_2^{\downarrow\{o_1, \bar{o}_1\}})^{\uparrow\Omega}. \quad (\text{B.3})$$

Finalement, supposons que les bbas m_i^Ω , $i \in \{1, 2\}$ sont elles mêmes issues d'une combinaison conjonctive de bbas 'partielles' $m_i^{\downarrow\{o_j, \bar{o}_j\}}$, $j \in \{1, 2\}$: $m_i^\Omega = \left(m_i^{\downarrow\{o_1, \bar{o}_1\}}\right)^{\uparrow\Omega} \odot \left(m_i^{\downarrow\{o_2, \bar{o}_2\}}\right)^{\uparrow\Omega}$. Ainsi, en se basant sur l'associativité de la règle conjonctive :

$$\begin{aligned} m_1^\Omega \odot m_2^\Omega &= \\ &\left(\left(m_1^{\downarrow\{o_1, \bar{o}_1\}}\right)^{\uparrow\Omega} \odot \left(m_1^{\downarrow\{o_2, \bar{o}_2\}}\right)^{\uparrow\Omega}\right) \odot \left(\left(m_2^{\downarrow\{o_1, \bar{o}_1\}}\right)^{\uparrow\Omega} \odot \left(m_2^{\downarrow\{o_2, \bar{o}_2\}}\right)^{\uparrow\Omega}\right) = \\ &\left(\left(m_1^{\downarrow\{o_1, \bar{o}_1\}}\right)^{\uparrow\Omega} \odot \left(m_2^{\downarrow\{o_1, \bar{o}_1\}}\right)^{\uparrow\Omega}\right) \odot \left(\left(m_1^{\downarrow\{o_2, \bar{o}_2\}}\right)^{\uparrow\Omega} \odot \left(m_2^{\downarrow\{o_2, \bar{o}_2\}}\right)^{\uparrow\Omega}\right) = \\ &\quad (m_1^{\downarrow\{o_1, \bar{o}_1\}} \odot m_2^{\downarrow\{o_1, \bar{o}_1\}})^{\uparrow\Omega} \odot (m_1^{\downarrow\{o_2, \bar{o}_2\}} \odot m_2^{\downarrow\{o_2, \bar{o}_2\}})^{\uparrow\Omega}. \end{aligned}$$

Bibliographie

- [1] W. Rekik, S. Le Hégarat-Masclé, C. André, R. Reynaud, A. Kallel, and A. Ben Hamida. Data association for object enumeration using belief function theory. In *Proceedings of Belief Functions : Theory and Applications, Oxford*, pages 383–392. 2014.
- [2] W. Rekik, S. Le Hégarat-Masclé, C. André, R. Reynaud, A. Kallel, and A. Ben Hamida. Object reconstruction in an image based on belief function representation. In *21st IEEE International Conference on Image Processing (ICIP), Paris*, pages 1633–1637, 2014.
- [3] B. Mourllion, D. Gruyer, C. Royère, and S. Théroutte. Multi-hypotheses tracking algorithm based on the belief theory. In *Proceedings of the 8th International Conference on Information Fusion*, volume 2, pages 8–pp, 2005.
- [4] D. Gruyer, C. Royère, and V. Cherfaoui. Heterogeneous multi-criteria combination with partial or full information. In *Proceedings of the 6th International Conference on Information Fusion*, volume 2, pages 1186–1193. IEEE, 2003.
- [5] D. Gruyer. *Etude du traitement de données imparfaites pour le suivi multi-objets : Application aux situations routières*. PhD thesis, 1999.
- [6] D. Mercier, E. Lefèvre, and D. Jolly. Object association with belief functions, an application with vehicles. *Information Sciences*, 181(24) :5485–5500, 2011.
- [7] N. El Zoghby, V. Cherfaoui, and T. Denoeux. Optimal object association from pairwise evidential mass functions. In *16th International Conference on Information Fusion (FUSION)*, pages 774–780, July 2013.

- [8] T. Denœux, N. El Zoghby, V. Cherfaoui, and A. Jouglet. Optimal object association in the Dempster–Shafer framework. *IEEE Transactions on Cybernetics*, 44 :2521–2531, 2014.
- [9] W. Rekik, S. Le Hégarat-Masclé, R. Reynaud, A. Kallel, and A. Ben Hamida. Dynamic estimation of the discernment frame in belief function theory : Application to object detection. *Information Sciences, Elsevier*, 306 :132–149, 2015.
- [10] W. Rekik, S. Le Hégarat-Masclé, R. Reynaud, A. Kallel, and A. Ben Hamida. Dynamic estimation of the discernment frame in belief function theory. In *16th International Conference on Information Fusion (FUSION), Istanbul*, pages 1135–1142, July 2013.
- [11] W. Rekik, S. Le Hégarat-Masclé, R. Reynaud, A. Kallel, and A. Ben Hamida. Sequential object detection using belief function theory. In *1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse*, pages 19–24. IEEE, 2014.
- [12] A. Yilmaz, O. Javed, and M. Shah. Object tracking : A survey. *Acm Computing Surveys*, 38, 2006.
- [13] URL <http://dirac.epucfe.eu/projets/wakka.php?wiki=P08A08index>.
- [14] K. C. Smith. *Bayesian methods for visual multi-object tracking with applications to human activity recognition*. PhD thesis, 2007.
- [15] M. Ammar, S. Le Hégarat-Masclé, M. Vasiliu, and R. Reynaud. An a-contrario approach for object detection in video sequence. *International Journal of Pure and Applied Mathematics*, 89 :173–201, 2013.
- [16] D. Comaniciu and P. Meer. Mean shift : A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5) : 603–619, 2002.
- [17] S. Jianbo and M. Jitendra. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8) :888–905, 2000.
- [18] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International journal of computer vision*, 22(1) :61–79, 1997.

- [19] H. A. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1) :23–38, 1998.
- [20] C. P. Papageorgiou, M. Oren, and T. Poggio. A general framework for object detection. In *6th international conference on Computer vision, 1998*, pages 555–562. IEEE, 1998.
- [21] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2) :91–110, 2004.
- [22] D. Comaniciu, V. Ramesh, and P. Meer. Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5) :564–577, 2003.
- [23] A. Yilmaz, X. Li, and M. Shah. Contour-based object tracking with occlusion handling in video acquired using mobile cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(11) :1531–1536, 2004.
- [24] A. Ali and J. K. Aggarwal. Segmentation and recognition of continuous human activity. In *IEEE Workshop on Detection and recognition of events in video*, pages 28–35, 2001.
- [25] G. Paschos. Perceptually uniform color spaces for color texture analysis : an empirical evaluation. *IEEE Transactions on Image Processing*, 10(6) :932–937, 2001.
- [26] R. M. Haralick, K. Shanmugam, and H. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, (6) :610–621, 1973.
- [27] G. J. Edwards, C. J. Taylor, and T. F. Cootes. Interpreting face images using active appearance models. In *3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pages 300–305, 1998.
- [28] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7) : 696–710, 1997.
- [29] M. J. Black and A. D. Jepson. Eigentracking : Robust matching and tracking of articulated objects using a view-based representation. *International Journal of Computer Vision*, 26(1) :63–84, 1998.

- [30] R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1) :35–45, 1960.
- [31] M. Athans, R. P. Wishner, and A. Bertolini. Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements. *IEEE Transactions on Automatic Control*, 13(5) :504–514, 1968.
- [32] S. J. Julier and J. K. Uhlmann. A new extension of the kalman filter to nonlinear systems. In *Int. symp. aerospace/defense sensing, simul. and controls*, volume 3, pages 3–2. Orlando, FL, 1997.
- [33] T. J. Broida and R. Chellappa. Estimation of object motion parameters from noisy images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1) : 90–99, 1986.
- [34] R. Rosales and S. Sclaroff. A framework for heading-guided recognition of human activity. *Computer Vision and Image Understanding*, 91(3) :335–367, 2003.
- [35] P. Del Moral, G. Rigal, and G. Salut. Estimation et commande optimale non linéaire : un cadre unifié pour la résolution particulière. *Rapport technique LAAS/CNRS, Toulouse*, 1992.
- [36] N. J. Gordon, D. J. Salmond, and A. FM. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEEE Proceedings on Radar and Signal Processing*, volume 140, pages 107–113, 1993.
- [37] G. Kitagawa. Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of computational and graphical statistics*, 5(1) :1–25, 1996.
- [38] M. Isard and A. Blake. Contour tracking by stochastic propagation of conditional density. In *Computer Vision ECCV'96*, pages 343–356. 1996.
- [39] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P. J. Nordlund. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on Signal Processing*, 50(2) :425–437, 2002.
- [40] J. Vermaak, C. Andrieu, A. Doucet, and S. J. Godsill. Particle methods for bayesian modeling and enhancement of speech signals. *IEEE Transactions on Speech and Audio Processing*, 10(3) :173–185, 2002.

- [41] O. Frank, J. Nieto, J. Guivant, and S. Scheduling. Multiple target tracking using sequential monte carlo methods and statistical data association. In *IEEE International Conference on Intelligent Robots and Systems, 2003*, volume 3, pages 2718–2723, 2003.
- [42] E. Erdem, S. Dubuisson, and I. Bloch. Fragments based tracking with adaptive cue integration. *Computer vision and image understanding*, 116(7) :827–841, 2012.
- [43] V. Berge. *Contribution à la gestion des incertitudes en fusion multicapteurs Application à la persection du contexte de conduite*. Habilitation à diriger des recherches, Université de Technologie de Compiègne, 2009.
- [44] T Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1) :21–27, 1967.
- [45] Y. Bar-Shalom and E. Tse. Tracking in a cluttered environment with probabilistic data association. *Automatica*, 11(5) :451–460, 1975.
- [46] K. Fukunaga and T. E. Flick. An optimal global nearest neighbor metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (3) :314–318, 1984.
- [47] Y. Bar-Shalom, F. Daum, and J. Huang. The probabilistic data association filter. *IEEE Control Systems*, 29(6) :82–100, 2009.
- [48] R. L. Streit and T. E. Luginbuhl. Probabilistic multi-hypothesis tracking. Technical report, DTIC Document, 1995.
- [49] J. Munkres. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial & Applied Mathematics*, 5(1) :32–38, 1957.
- [50] Y. Bar-Shalom. *Tracking and data association*. Academic Press Professional, 1987.
- [51] Y. L. Chang and J. K. Aggarwal. 3d structure reconstruction from an ego motion sequence using statistical estimation and detection theory. In *Proceedings of the IEEE Workshop on Visual Motion*, pages 268–273. IEEE, 1991.
- [52] C. Rasmussen and G. D. Hager. Probabilistic data association methods for tracking complex visual objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6) :560–576, 2001.

- [53] D. B. Reid. An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, 24(6) :843–854, 1979.
- [54] K. G. Murty. Letter to the editor an algorithm for ranking all the assignments in order of increasing cost. *Operations Research*, 16(3) :682–687, 1968.
- [55] S. Le Hégarat-Masclé, R. Reynaud, and C. André. Association plots-pistes dans un système de protection périmétrique. In *XXIIe colloque GRETSI*. GRETSI, Groupe d’Étude du Traitement du Signal et des Images, 2009.
- [56] M. A. Zulkifley and B. Moran. Robust hierarchical multiple hypothesis tracker for multiple-object tracking. *Expert Systems with Applications*, 39 :12319–12331, November 2012.
- [57] B. Ristic and D. Salmond. A study of a nonlinear filtering problem for tracking an extended target. In *Proc. 7th Intl. Conf. on Information Fusion*, pages 503–509. Citeseer, 2004.
- [58] K. Gilholm and D. Salmond. Spatial distribution model for tracking extended objects. In *IEEE Proceedings on Radar, Sonar and Navigation*, volume 152, pages 364–371. IET, 2005.
- [59] A. P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *The annals of mathematical statistics*, pages 325–339, 1967.
- [60] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton, 1976.
- [61] R. Fagin and J. Y. Halpern. Uncertainty, belief, and probability. *Computational Intelligence*, 7(3) :160–173, 1991.
- [62] P. R. Halmos and H. Mathematician. *Measure theory*, volume 2. van Nostrand New York, 1950.
- [63] J. Y. Halpern and R. Fagin. Two views of belief : belief as generalized probability and belief as evidence. *Artificial Intelligence*, 54(3) :275–317, 1992.
- [64] J. Pearl. Reasoning with belief functions : An analysis of compatibility. *International Journal of Approximate Reasoning*, 4(5) :363–389, 1990.

- [65] L. A. Zadeh. Review of a mathematical theory of evidence. *Artificial Intelligence magazine*, 5(3) :81, 1984.
- [66] P. Smets. The combination of evidence in the transferable belief model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5) :447–458, 1990.
- [67] I. Bloch. Some aspects of Dempster-Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account. *Pattern Recognition Letters*, 17(8) :905–919, 1996.
- [68] S. Le Hégarat-Masclé and R. Seltz. Automatic change detection by evidential fusion of change indices. *Remote Sensing of Environment*, 91(3) :390–404, 2004.
- [69] P. Kumar, A. Mittal, and P. Kumar. Addressing uncertainty in multi-modal fusion for improved object detection in dynamic environment. *Information Fusion*, 11(4) :311–324, 2010.
- [70] F. Janez and A. Appriou. Theory of evidence and non-exhaustive frames of discernment : Plausibilities correction methods. *International Journal of Approximate Reasoning*, 18(1) :1–19, 1998.
- [71] D. Dubois, H. Farreny, and H. Prade. Sur divers problèmes inhérents à l’automatisation des raisonnements de sens commun. In *Proc. 5th Cong. AFCET Reconnaissance des Formes et IA, Grenoble*, pages 321–328, 1985.
- [72] P. Smets. Analyzing the combination of conflicting belief functions. *Information Fusion*, 8(4) :387–412, 2007.
- [73] P. Smets. Belief functions and the transferable belief model. *The Imprecise Probabilities Project*.
- [74] F. Janez. *Fusion de sources d’information définies sur des référentiels non exhaustifs différents. Solutions proposées sous le formalisme de la théorie de l’évidence*. PhD thesis, 1996.
- [75] I. Bloch. *Fusion d’informations en traitement du signal et des images*. Hermes Science Publications, 2003.
- [76] B. Fassinut-Mombot and J. B. Choquel. Gestion entropique du flot d’informations et minimisation des ressources matérielles en fusion d’informations. *Traitement du Signal*, 20(4) :389, 2003.

- [77] P. Smets. The concept of distinct evidence. In *IPMU*, pages 789–794, 1992.
- [78] P. Smets. Belief functions : the disjunctive rule of combination and the generalized bayesian theorem. *International Journal of Approximate Reasoning*, 9(1) :1–35, 1993.
- [79] S. Le Hégarat-Mascle, R. Reynaud, C. André, and A. Kallel. Utilisation contextuelle d’une règle prudente adaptative pour la combinaison d’évidences séparables. *REE. Revue de l’électricité et de l’électronique*, (8), 2009.
- [80] A. Appriou. Discrimination multisignal par la théorie de l’évidence. *Décision et Reconnaissance des formes en signal*, Hermes Science Publication, pages 219–258, 2002.
- [81] G. L. Rogova and V. Nimier. Reliability in information fusion : literature survey. In *Proceedings of the 7th International Conference on Information Fusion*, pages 1158–1165, 2004.
- [82] Z. Elouedi, K. Mellouli, and P. Smets. Assessing sensor reliability for multisensor data fusion within the transferable belief model. *IEEE Transactions on Systems, Man, and Cybernetics, Part B : Cybernetics*, 34(1) :782–787, 2004.
- [83] D. Mercier. *Fusion d’informations pour la reconnaissance automatique d’adresses postales dans le cadre de la théorie des fonctions de croyance*. PhD thesis, Université de technologie de compiègne, 2006.
- [84] D. Mercier, B. Quost, and T. Denœux. Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, 9 (2) :246–258, 2008.
- [85] C. Zeng and P. Wu. A reliability discounting strategy based on plausibility function of evidence. In *10th International Conference on Information Fusion*, pages 1–6. IEEE, 2007.
- [86] R. R. Yager. Entropy and specificity in a mathematical theory of evidence. *International Journal of General System*, 9(4) :249–260.
- [87] J. Abellán and S. Moral. Upper entropy of credal sets, applications to credal classification. *International Journal of Approximate Reasoning*, 39(2) :235–255.

- [88] D. Dubois and H. Prade. The principle of minimum specificity as a basis for evidential reasoning. In *Uncertainty in knowledge-based systems*, pages 75–84. Springer, 1987.
- [89] R. R. Yager. On the Dempster-Shafer framework and new combination rules. *Information Sciences*, 41(2) :93–137, 1987.
- [90] P. Smets. Quantified epistemic possibility theory seen as an hyper cautious transferable belief model. In *Rencontres Francophones sur la Logique Floue et ses Applications*. Citeseer, 2000.
- [91] R. R. Yager. The entailment principle for Dempster-Shafer granules. *International Journal of Intelligent Systems*, 1(4) :247–262, 1986.
- [92] F. Klawonn and P. Smets. The dynamic of belief in the transferable belief model and specialization-generalization matrices. In *Proceedings of the 8th international conference on Uncertainty in artificial intelligence*, pages 130–137. Morgan Kaufmann Publishers Inc., 1992.
- [93] T. Denœux. Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. *Artificial Intelligence*, 172(2–3) :234–264, 2008.
- [94] D. Dubois and H. Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4(3) :244–264, 1988.
- [95] P. Smets and R. Kennes. The transferable belief model. *Artificial intelligence*, 66(2) :191–234, 1994.
- [96] D. Dubois and H. Prade. On the unicity of Dempster rule of combination. *International Journal of Intelligent Systems*, 1(2) :133–142, 1986.
- [97] D. Dubois, H. Prade, and P. Smets. New semantics for quantitative possibility theory. In *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, pages 410–421. Springer, 2001.
- [98] T. Denœux. The cautious rule of combination for belief functions and some extensions. In *9th International Conference on Information Fusion*, pages 1–8. IEEE, 2006.

- [99] D. Dubois and H. Prade. A set-theoretic view of belief functions logical operations and approximations by fuzzy sets. *International Journal Of General System*, 12 (3) :193–226, 1986.
- [100] P. Smets. The canonical decomposition of a weighted belief. In *IJCAI*, volume 95, pages 1896–1901, 1995.
- [101] P. Smets. The application of the matrix calculus to belief functions. *International Journal of Approximate Reasoning*, 31(1) :1–30, 2002.
- [102] R. Kennes. Evidential reasoning in a categorial perspective : conjunction and disjunction of belief functions. In *Proceedings of the 7th conference on Uncertainty in Artificial Intelligence*, pages 174–181, 1991.
- [103] T. Denœux and A. Ben Yaghlane. Approximating the combination of belief functions using the fast moebius transform in a coarsened frame. *International Journal of Approximate Reasoning*, 31(1) :77–101, 2002.
- [104] M. H. DeGroot. *Optimal statistical decisions*, volume 82. John Wiley & Sons, 2005.
- [105] F. P. Ramsey. Truth and probability. *The foundations of mathematics and other logical essays*, pages 156–198, 1931.
- [106] L. J. Savage. *The foundations of statistics*. Courier Dover Publications, 2012.
- [107] T. Denœux. Analysis of evidence-theoretic decision rules for pattern classification. *Pattern recognition*, 30(7) :1095–1107, 1997.
- [108] D. Dubois and H. Prade. On several representations of an uncertain body of evidence. *Fuzzy information and decision processes*, pages 167–181, 1982.
- [109] P. Smets. Belief functions on real numbers. *International journal of approximate reasoning*, 40(3) :181–223, 2005.
- [110] A. Appriou. Approche générique de la gestion de l’incertain dans les processus de fusion multisenseur. *Traitement du signal*, 22 :307 – 319, 2005.
- [111] S. Ben Hariz, Z. Elouedi, and K. Mellouli. Clustering approach using belief function theory. In *Artificial Intelligence : Methodology, Systems, and Applications*, pages 162–171. Springer, 2006.

- [112] Q. Ye, X. Wu, and Z. Chen. An approach for evidence clustering using generalized distance. *Journal of electronics (china)*, 26(1) :18–23, 2009.
- [113] M. J. Beynon. The role of the DS/AHP in identifying inter-group alliances and majority rule within group decision making. *Group decision and negotiation*, 15(1) :21–42, 2006.
- [114] J. Schubert. Clustering decomposed belief functions using generalized weights of conflict. *International journal of approximate reasoning*, 48(2) :466–480, 2008.
- [115] A. Martin, A. L. Jousselme, and C. Osswald. Conflict measure for the discounting operation on belief functions. In *11th International Conference on Information Fusion*, pages 1–8. IEEE, 2008.
- [116] M. C. Florea, A. L. Jousselme, and E. Bossé. Dynamic estimation of evidence discounting rates based on information credibility. *RAIRO-Operations Research*, 44(04) :285–306, 2010.
- [117] C. Liang-zhou, S. Wen-kang, D. Yong, and Z. Zhen-fu. A new fusion approach based on distance of evidences. *Journal of Zhejiang University Science A*, 6(5) : 476–482, 2005.
- [118] J. Klein, O. Colot, et al. Automatic discounting rate computation using a dissent criterion. In *Workshop on the theory of belief functions*, pages 1–6, 2010.
- [119] B. Ristic and P. Smets. The TBM global distance measure for the association of uncertain combat id declarations. *Information fusion*, 7(3) :276–284, 2006.
- [120] A. Roquel. *Exploitation du conflit entre capteurs pour la gestion d’un système complexe multi-capteurs*. PhD thesis, Université Paris Sud-Paris XI, 2012.
- [121] R. Jonker and A. Volgenant. A shortest augmenting path algorithm for dense and sparse linear assignment problems. *Computing*, 38(4) :325–340, 1987.
- [122] V. Kolmogorov and R. Zabini. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2) : 147–159, 2004.
- [123] C. André. *Approche crédibiliste pour la fusion multi capteurs décentralisée*. PhD thesis, Université Paris Sud, 2013.

- [124] C. Royère, D. Gruyer, and V. Cherfaoui. Data association with believe theory. In *Proceedings of the 3rd International Conference on Information Fusion*, volume 1, pages TUD2–3. IEEE, 2000.
- [125] A. Ayoun and P. Smets. Data association in multi-target detection using the transferable belief model. *International Journal of Intelligent Systems*, 16(10) : 1167–1182, 2001.
- [126] B. Ristic and P. Smets. Global cost of assignment in the tbm framework for association of uncertain id reports. *Aerospace science and technology*, 11(4) :303–309, 2007.
- [127] G. Gennari, A. Chiuso, F. Cuzzolin, and R. Frezza. Integrating shape and dynamic probabilistic models for data association and tracking. In *Proceedings of the 41st IEEE Conference on Decision and Control*, volume 3, pages 2409–2414. IEEE, 2002.
- [128] G. Gennari, A. Chiuso, F. Cuzzolin, and R. Frezza. Integration of shape constraints in data association filters. In *43rd IEEE Conference on Decision and Control*, volume 3, pages 2668–2673. IEEE, 2004.
- [129] F. Cuzzolin. *Visions of a Generalized Probability Theory*. PhD thesis, Department of Information Engineering, University of Padova, Italy, 2001.
- [130] M. Rombaut. Decision in multi-obstacle matching process using theory of belief. *AVCS*, 98 :1–3, 1998.
- [131] L. W. Perry and H. E. Stephanou. Belief function divergence as a classifier. In *IEEE International Symposium In Intelligent Control*, 1991.
- [132] D. Harmanec. Measures of uncertainty and information. *Imprecise Probability Project*, 1999.
- [133] B. J. Tessem et al. Approximations for efficient computation in the theory of evidence. *Artificial Intelligence*, 61(2) :315–329, 1993.
- [134] A. L. Jousselme, D. Grenier, and E. Bossé. A new distance between two bodies of evidence. *Information fusion*, 2(2) :91–101, 2001.

- [135] D. Yong, S. WenKang, Z. ZhenFu, and L. Qi. Combining belief functions based on distance of evidence. *Decision support systems*, 38(3) :489–493, 2004.
- [136] T. Denœux. Extending stochastic ordering to belief functions on the real line. *Information Sciences*, 179(9) :1362–1376, 2009.
- [137] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis. Real-time foreground–background segmentation using codebook model. *Real-time imaging*, 11(3) :172–185, 2005.
- [138] S. Le Hégarat-Masclé, I. Bloch, and D. Vidal-Madjar. Application of Dempster-Shafer evidence theory to unsupervised classification in multisource remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 35(4) :1018–1031, jul 1997.
- [139] J. Schubert. Constructing and evaluating alternative frames of discernment. *International Journal of Approximate Reasoning*, 53 :176–189, 2012.
- [140] C. E. Shannon. A mathematical theory of communication. *Mobile Computing and Communications Review*, 5(1) :3–55, 2001.
- [141] R. VL. Hartley. Transmission of information. *Bell System technical journal*, 7(3) : 535–563, 1928.
- [142] R. Haenni. Shedding new light on Zadeh’s criticism of Dempster’s rule of combination. In *Proceedings of the 8th International Conference on Information Fusion*, volume 2, pages 1–6. IEEE, 2005.
- [143] M. Mirmehdi, X. Xie, and J. Suri. *Handbook of texture analysis*. Imperial College Press, 2009.
- [144] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (6) :679–698, 1986.
- [145] B. K. Horn and B. G. Schunck. Determining optical flow. In *Technical Symposium East*, pages 319–331. International Society for Optics and Photonics, 1981.
- [146] K. Fragkiadaki, W. Zhang, G. Zhang, and J. Shi. Two-granularity tracking : Mediating trajectory and detection graphs for tracking under occlusions. In *Computer Vision*, pages 552–565. 2012.

- [147] A. L. Blum and P. Langley. Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97(1) :245–271, 1997.
- [148] H. Schweitzer, J. W. Bell, and F. Wu. Very fast template matching. In *Computer Vision*, pages 358–372. 2002.
- [149] A. D. Jepson, D. J. Fleet, and T. F. El-Maraghi. Robust online appearance models for visual tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10) :1296–1311, 2003.
- [150] N. Bouaynaya, W. Qu, and D. Schonfeld. An online motion-based particle filter for head tracking applications. In *Proceedings of ICASSP (2)*, pages 225–228, 2005.
- [151] P. Smets and B. Ristic. Kalman filter and joint tracking and classification based on belief functions in the TBM framework. *Information Fusion*, 8(1) :16–27, 2007.
- [152] R. Munoz-Salinas, R. Medina-Carnicer, F. J. Madrid-Cuevas, and A. Carmona-Poyato. Multi-camera people tracking using evidential filters. *International Journal of Approximate Reasoning*, 50(5) :732–749, 2009.
- [153] S. Hachour, F. Delmotte, D. Mercier, and E. Lefèvre. Object tracking and credal classification with kinematic data in a multi-target context. *Information Fusion*, 20 :174–188, 2014.
- [154] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms : a systematic survey. *IEEE Transactions on Image Processing*, 14(3) : 294–307, 2005.
- [155] A. Desolneux, M. Lionel, and J. M. Morel. Meaningful alignments. *International Journal of Computer Vision*, 40(1) :7–23, 2000.
- [156] A. Desolneux, L. Moisan, and J. M. Morel. A grouping principle and four applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 508–513, April 2003.
- [157] A. Robin, L. Moisan, and S. Le Hégarat-Masclé. An a-contrario approach for sub-pixel change detection in satellite imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32 :1977–1993, November 2010.

- [158] F. Dibos, S. Pelletier, and G. Koepfler. Real-time segmentation of moving objects in a video sequence by a contrario detection. In *Proceedings of the IEEE International Conference on Image Processing*, volume 1, pages 1065–1068, 2005.
- [159] A. Manzanera and J. C. Richefeu. A new motion detection algorithm based on sigma delta background estimation. *Pattern Recognition Letters*, 28(3) :320–328, 2007.
- [160] A. Bhattacharyya. On a measure of divergence between two statistical populations defined by their probability distributions. *Bull. Calcutta Math. Soc.*, 35 :99–109, 1943.

Résumé

Les systèmes intelligents sont de plus en plus présents dans notre société à l'instar des systèmes de surveillance et de protection de sites civils ou militaires. Leur but est de détecter les intrus et remonter une alarme ou une menace à un opérateur distant. Dans nos travaux, nous nous intéressons à de tels systèmes avec comme objectif de gérer au mieux la qualité de l'information présentée à l'opérateur en termes de fiabilité et précision. Nous nous concentrons sur la modalité image en vue de gérer des détections à la fois incertaines et imprécises de façon à présenter des objets fiables à l'opérateur.

Pour préciser notre problème nous posons les contraintes suivantes. La première est que le système soit modulaire, l'une des briques (ou sous-fonctions) du système étant la détection de fragments correspondant potentiellement à des objets. Notre deuxième contrainte est alors de n'utiliser que des informations issues de la géométrie des détections fragmentaires : localisation spatiale dans l'image et taille des détections. Une menace est alors supposée d'autant plus importante que les détections sont de tailles importantes et temporellement persistantes.

Le cadre formel choisi est la théorie des fonctions de croyance qui permet de modéliser des données à la fois imprécises et incertaines. Les contributions de cette thèse concernent la représentation des objets en termes de localisation imprécise et incertaine et le filtrage des objets.

La représentation pertinente des informations est un point clé pour les problèmes d'estimation ou la prise de décision. Une bonne représentation se reconnaît au fait qu'en découlent des critères simples et performants pour résoudre des sous-problèmes. La représentation proposée dans cette thèse a été valorisée par le fait qu'un critère d'association entre nouvelles détections (fragments) et objets en construction, a pu être défini d'une façon simple et rigoureuse. Rappelons que cette association est une étape clé pour de nombreux problèmes impliquant des données non étiquetées, ce qui étend notre contribution au-delà de l'application considérée.

Le filtrage des données est utilisé dans de nombreuses méthodes ou algorithmes pour robustifier les résultats en s'appuyant sur la redondance attendue des données s'opposant à l'inconsistance du bruit. Nous avons alors formulé ce problème en termes d'estimation dynamique d'un cadre de discernement contenant les 'vraies hypothèses'. Ce cadre est estimé dynamiquement avec la prise en compte de nouvelles données (ou observations) permettant de détecter deux principaux types d'erreurs : la duplication de certaines hypothèses (objets dans notre application), la présence de fausses alarmes (dus au bruit ou aux fausses détections dans notre cas).

Pour finir nous montrons la possibilité de coupler nos briques de construction des objets et de filtrage de ces derniers avec une brique de suivi utilisant des informations plus haut niveau, telle que les algorithmes de tracking classiques de traitement d'image. Mots-clés : théorie des fonctions de croyance, association de données, filtrage.

Abstract

Intelligent systems are more and more present in our society, like the systems of surveillance and civilian or military sites protection. Their purpose is to detect intruders and present the alarms or threats to a distant operator. In our work, we are interested in such systems with the aim to better handle the quality of information presented to the operator in terms of reliability and precision. We focus on the image modality and we have to handle detections that are both uncertain and imprecise in order to present reliable objects to the operator.

To specify our problem, we consider the following constraints. The first one is that the system is modular ; one subpart of the system is the detection of fragments corresponding potentially to objects. Our second constraint is then to use only information derived from the geometry of these fragmentary detections : spatial location in the image and size of the detections. Then, a threat is supposed all the more important as the detections have an important size and are temporally persistent.

The chosen formal framework is the belief functions theory that allows modeling imprecise and uncertain data. The contributions of this thesis deal with the objects representation in terms of imprecise and uncertain location of the objects and object filtering.

The pertinent representation of information is a key point for estimation problems and decision making. A representation is good when simple and efficient criteria for the resolution of sub problems can be derived. The representation proposed has allowed us to derive, in a simple and rigorous way, an association criterion between new detections (fragments) and objects under construction. We remind that this association is a key step for several problems with unlabelled data that extends our contribution beyond of the considered application.

Data filtering is used in many methods and algorithms to robustify the results using the expected data redundancy versus the noise inconsistency. Then, we formulated our problem in terms of dynamic estimation of a discernment frame including the ‘true hypotheses’. This frame is dynamically estimated taking into account the new data (or observations) that allow us to detect two main types of errors, namely the duplication of some hypotheses (objects in our application) and the presence of false alarms (due to noise or false detections in our case).

Finally, we show the possibility of coupling our sub-functions dealing with object construction and their filtering with a tracking process using higher level information such as classical tracking algorithm in image processing.

Mots-clés : belief functions theory, data association, filtering.

Liste des publications

1. W. Rekik, S. Le Hégarat-Masclé, R. Reynaud, A. Kallel, and A. Ben Hamida. Dynamic estimation of the discernment frame in belief function theory : Application to object detection. *Information Sciences, Elsevier*, 306 :132-149, 2015.
2. W. Rekik, S. Le Hégarat-Masclé, R. Reynaud, A. Kallel, and A. Ben Hamida. Dynamic estimation of the discernment frame in belief function theory. In *16th International Conference on Information Fusion (FUSION), Istanbul*, pages 1135–1142, July 2013.
3. W. Rekik, S. Le Hégarat-Masclé, R. Reynaud, A. Kallel, and A. Ben Hamida. Sequential object detection using belief function theory. In *1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse*, pages 19–24. IEEE, 2014.
4. W. Rekik, S. Le Hégarat-Masclé, C. André, R. Reynaud, A. Kallel, and A. Ben Hamida. Data association for object enumeration using belief function theory. In *Proceedings of Belief Functions : Theory and Applications, Oxford*, pages 383–392, 2014.
5. W. Rekik, S. Le Hégarat-Masclé, C. André, R. Reynaud, A. Kallel, and A. Ben Hamida. Object reconstruction in an image based on belief function representation. In *21st IEEE International Conference on Image Processing (ICIP), Paris*, pages 1633–1637, 2014.

Représentation, Filtrage Dynamique et Association dans un Cadre Evidentiel : Vers le Suivi d'Objets

Wafa REKIK

Résumé : Les systèmes intelligents sont de plus en plus présents dans notre société à l'instar des systèmes de surveillance et de protection de sites civils ou militaires. Leur but est de détecter les intrus et remonter une alarme ou une menace à un opérateur distant. Dans nos travaux, nous nous intéressons à de tels systèmes avec comme objectif de gérer au mieux la qualité de l'information présentée à l'opérateur en termes de fiabilité et précision. Nous nous concentrons sur la modalité image en vue de gérer des détections à la fois incertaines et imprécises de façon à présenter des objets fiables à l'opérateur. Pour préciser notre problème nous posons les contraintes suivantes. La première est que le système soit modulaire, l'une des briques (ou sous-fonctions) du système étant la détection de fragments correspondant potentiellement à des objets. Notre deuxième contrainte est alors de n'utiliser que des informations issues de la géométrie des détections fragmentaires : localisation spatiale dans l'image et taille des détections. Une menace est alors supposée d'autant plus importante que les détections sont de tailles importantes et temporellement persistantes. Le cadre formel choisi est la théorie des fonctions de croyance qui permet de modéliser des données à la fois imprécises et incertaines. Les contributions de cette thèse concernent la représentation des objets en termes de localisation imprécise et incertaine et le filtrage des objets. La représentation pertinente des informations est un point clé pour les problèmes d'estimation ou la prise de décision. Une bonne représentation se reconnaît au fait qu'en découlent des critères simples et performants pour résoudre des sous-problèmes. La représentation proposée dans cette thèse a été valorisée par le fait qu'un critère d'association entre nouvelles détections (fragments) et objets en construction, a pu être défini d'une façon simple et rigoureuse. Rappelons que cette association est une étape clé pour de nombreux problèmes impliquant des données non étiquetées, ce qui étend notre contribution au-delà de l'application considérée. Le filtrage des données est utilisé dans de nombreuses méthodes ou algorithmes pour robustifier les résultats en s'appuyant sur la redondance attendue des données s'opposant à l'inconsistance du bruit. Nous avons alors formulé ce problème en termes d'estimation dynamique d'un cadre de discernement contenant les 'vraies hypothèses'. Ce cadre est estimé dynamiquement avec la prise en compte de nouvelles données (ou observations) permettant de détecter deux principaux types d'erreurs : la duplication de certaines hypothèses (objets dans notre application), la présence de fausses alarmes (dus au bruit ou aux fausses détections dans notre cas). Pour finir nous montrons la possibilité de coupler nos briques de construction des objets et de filtrage de ces derniers avec une brique de suivi utilisant des informations plus haut niveau, telle que les algorithmes de *tracking* classiques de traitement d'image.

Abstract: Intelligent systems are more and more present in our society, like the systems of surveillance and civilian or military sites protection. Their purpose is to detect intruders and present the alarms or threats to a distant operator. In our work, we are interested in such systems with the aim to better handle the quality of information presented to the operator in terms of reliability and precision. We focus on the image modality and we have to handle detections that are both uncertain and imprecise in order to present reliable objects to the operator. To specify our problem, we consider the following constraints. The first one is that the system is modular ; one subpart of the system is the detection of fragments corresponding potentially to objects. Our second constraint is then to use only information derived from the geometry of these fragmentary detections : spatial location in the image and size of the detections. Then, a threat is supposed all the more important as the detections have an important size and are temporally persistent. The chosen formal framework is the belief functions theory that allows modeling imprecise and uncertain data. The contributions of this thesis deal with the objects representation in terms of imprecise and uncertain location of the objects and object filtering. The pertinent representation of information is a key point for estimation problems and decision making. A representation is good when simple and efficient criteria for the resolution of sub problems can be derived. The representation proposed has allowed us to derive, in a simple and rigorous way, an association criterion between new detections (fragments) and objects under construction. We remind that this association is a key step for several problems with unlabelled data that extends our contribution beyond of the considered application. Data filtering is used in many methods and algorithms to robustify the results using the expected data redundancy versus the noise inconsistency. Then, we formulated our problem in terms of dynamic estimation of a discernment frame including the 'true hypotheses'. This frame is dynamically estimated taking into account the new data (or observations) that allow us to detect two main types of errors, namely the duplication of some hypotheses (objects in our application) and the presence of false alarms (due to noise or false detections in our case). Finally, we show the possibility of coupling our sub-functions dealing with object construction and their filtering with a tracking process using higher level information such as classical tracking algorithm in image processing.

Mots clés: théorie des fonctions de croyance, association de données, filtrage.

Key-words: belief functions theory, data association, filtering.