

Numéro d'identification :

Laboratoire d'Informatique, de Traitement de l'Information et des Systèmes

# THESE DE DOCTORAT

pour l'obtention du Grade de  
DOCTEUR DE L'UNIVERSITE DE ROUEN

Discipline : **Physique**

Spécialité : **Sciences de l'information**

soutenue par

**John KLEIN**

---

## Suivi robuste d'objets dans des séquences d'images par fusion de sources, application au suivi de véhicules dans des scènes routières

---

Directeur de Thèse :

M. Pierre MICHÉ

Encadrant de Thèse :

Mme Christèle LECOMTE

Soutenue le 4 décembre 2008 devant le jury composé de :

Claude Laugeau, Professeur, Ecole des Mines de Paris

Olivier Colot, Professeur, Université des Sciences et Technologies de Lille

Thierry Dencœur, Professeur, Université de Technologie de Compiègne

Laurent Trassoudaine, Professeur, Université Blaise Pascal de Clermont Ferrand

Pierre Miché, Professeur, Université de Rouen

Christèle Lecomte, Maître de Conférences, Université de Rouen

Président du jury

Rapporteur

Rapporteur

Examineur

Directeur de thèse

Co-Encadrante de thèse

remerciement labo



---

# Remerciements

*remerciements*

---

# Résumé

Ces travaux de thèse sont destinés à une application de suivi visuel de véhicules. Une caméra est placée à l'intérieur d'un véhicule ; il s'agit alors de déterminer, dans chaque image fournie par la caméra, la position du véhicule précédent dont seule la face arrière est *a priori* visible tout au long de la séquence d'images. Cette application s'inscrit dans le domaine des systèmes de transport intelligents et fait appel à des algorithmes de vision par ordinateur.

La vision par ordinateur a connu un intérêt croissant depuis plusieurs décennies et de nombreuses approches de suivi d'objets ont déjà été proposées. A l'issue d'une étude bibliographique, nous avons constaté que les algorithmes de suivi n'ayant recours qu'à une seule source d'information fournissent un suivi de qualité nettement inférieure à ceux ayant recours au multi-sources. Nous avons donc choisi de concentrer nos efforts sur l'amélioration d'un point particulier d'un algorithme de suivi : la fusion de données.

Nous utilisons dans cette thèse un algorithme de suivi par filtrage particulaire, qui a fait ses preuves en termes de qualité de suivi. La fusion de données s'opère habituellement dans un filtre particulaire à l'aide de techniques probabilistes ou à l'aide de stratégies permettant de faire appel séparément à plusieurs sources lors de différentes étapes du filtre. Ces méthodes ne permettent pas une prise en compte explicite du conflit entre sources ou de la quantité d'information pertinente délivrée par les sources. A cet égard, nous proposons une méthode pour réaliser une fusion de données plus précise et plus efficace dans ce filtre grâce au cadre des fonctions de croyance qui permet de combler les défauts évoqués. Nous contribuons également à fournir un opérateur de combinaison pour les fonctions de croyance approprié aux besoins d'un suivi de véhicules, à savoir la fusion de sources non fiables et indéterminées. Cet opérateur s'adapte aux masses de croyance à fusionner et permet ainsi de cumuler certaines propriétés que des règles de combinaison ne peuvent posséder conjointement. L'approche proposée est validée à l'aide de plusieurs tests sur des séquences routières de natures variées : différents types de voie de circulation, situations climatiques dégradées, changements brutaux d'illumination, ombres portées, dénivelé ou virages. Une seule image du véhicule est supposée connue, l'approche ne requiert donc qu'une faible quantité d'*a priori*.

**Mots-clés :** suivi visuel d'objets, extraction de caractéristiques visuelles, filtrage particulaire, fusion de données, fonctions de croyance, théorie de Dempster-Shafer, véhicules intelligents.

---

# Abstract

This PhD thesis is meant to be applied to preceding vehicle rear view tracking problems. Using a camera installed inside a vehicle, our goal is to determine the position of the preceding vehicle in each image delivered by the camera. The vehicle's rear view is *a priori* the only visible part of the vehicle throughout an on-road sequence of images. This application is part of the field named intelligent transportation systems and exploits computer vision algorithms.

The interest for computer vision witnessed a fast growth in the past decades and various object tracking algorithms have already been proposed. After examining the state-of-the-art, it appears that tracking approaches which only use one source of information produce far worse results than multi-source approaches. Therefore, it was decided to focus our efforts on improving a particular part of tracking approaches : data fusion.

Particle filters were retained as the main tracking procedure for our work, because of their well-known advantages in an object tracking context. Data fusion inside a particle filter is usually operated using probabilistic techniques or strategies that enable a usage of several sources at different steps of the filter. These methods cannot explicitly take into account such things as the conflict between sources or the amount of relevant information brought by the sources. Thus, we propose a more efficient and accurate data fusion method inside particle filters using the framework of belief functions, which can deal with these issues. We also contribute to obtaining a new contextual combination pattern for belief functions which is more appropriate to specific needs of object tracking, *i.e.* unreliable and imprecise source combination. This pattern adapts itself automatically to belief masses that are to be aggregated; consequently it can cumulate some properties that a single combination rule cannot conjugate. The approach proposed is tested through several on-road sequences of various kinds : different road types, damaged weather conditions, sudden illumination changes, shadows, slopes and turns. Only one image of the tracked vehicle is known beforehand, therefore the system is based on little *a priori* information.

**Key words :** visual object tracking, visual characteristic extraction methods, particle filtering, data fusion, belief functions, Dempster-Shafer theory, intelligent vehicles.

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Modélisation d'objets pour le suivi</b>	<b>5</b>
1.1 Introduction et positionnement du problème de suivi . . . . .	5
1.2 Représentation des objets . . . . .	7
1.2.1 Boîte englobante . . . . .	8
1.2.2 Contour . . . . .	9
1.2.3 Squelette . . . . .	10
1.2.4 Maillage . . . . .	10
1.2.5 Ensemble de régions . . . . .	10
1.2.6 Ensemble de points d'intérêt et graphes . . . . .	11
1.2.7 Représentation choisie . . . . .	12
1.2.8 De la représentation vers le modèle d'objet . . . . .	12
1.3 Extraction de caractéristiques couleur et texture . . . . .	12
1.3.1 Méthodes colorimétriques . . . . .	13
1.3.2 Méthodes texture . . . . .	15
1.3.3 Méthodes texture-couleur . . . . .	19
1.3.4 Bilan sur l'information texture-couleur . . . . .	20
1.4 Extraction de caractéristiques de forme . . . . .	21
1.4.1 Détection de contours . . . . .	22
1.4.2 Squelettisation . . . . .	23
1.4.3 Méthodes structurelles . . . . .	24
1.4.4 Points d'intérêt et graphes . . . . .	25
1.4.5 Attributs statistiques . . . . .	27
1.4.6 Cartes de symétrie . . . . .	29
1.4.7 Autres méthodes de caractérisation de forme . . . . .	29
1.4.8 Bilan sur l'information de forme . . . . .	30
1.5 Extraction de caractéristiques de mouvement . . . . .	30
1.5.1 Méthodes basées sur la différence inter-images . . . . .	31
1.5.2 Méthodes d'analyse du fond de la scène . . . . .	36
1.5.3 Extraction en couches de mouvement . . . . .	37
1.5.4 Méthodes à base de mouvement cohérent . . . . .	38
1.5.5 Bilan sur l'information de mouvement . . . . .	38
1.6 Conclusion . . . . .	39
<b>2 Procédures de suivi d'objets</b>	<b>41</b>
2.1 Introduction . . . . .	41
2.2 Détection et suivi . . . . .	42
2.3 Méthodes d'optimisation d'une grandeur scalaire . . . . .	43
2.3.1 Maximisation de l'inter-corrélation . . . . .	43
2.3.2 Minimisation de l'erreur . . . . .	43
2.3.3 Minimisation d'une fonctionnelle d'énergie . . . . .	45

2.3.4	Minimisation d'un coût . . . . .	45
2.3.5	Minimisation d'une distance ou d'une mesure de dissimilarité . . . . .	46
2.3.6	Remarques sur les méthodes d'optimisation . . . . .	48
2.4	Méthodes probabilistes . . . . .	49
2.4.1	Filtrage Optimal . . . . .	49
2.4.2	Algorithme EM . . . . .	54
2.5	Méthodes d'appariements de détection . . . . .	55
2.5.1	Segmentation séquentielle . . . . .	55
2.5.2	Appariements probabilistes et crédibilistes . . . . .	57
2.6	Remarques supplémentaires . . . . .	58
2.6.1	Mise à jour des modèles . . . . .	58
2.6.2	Suivi d'objets multiples . . . . .	59
2.6.3	Évaluation des performances . . . . .	60
2.7	Les principales difficultés en suivi d'objets . . . . .	61
2.7.1	Les occultations . . . . .	61
2.7.2	Le fouillis . . . . .	62
2.7.3	Les mouvements particuliers . . . . .	63
2.7.4	Les changements d'illumination . . . . .	64
2.8	Conclusion et approche retenue . . . . .	65
<b>3</b>	<b>Processus de fusion de données applicable au filtrage particulaire</b>	<b>67</b>
3.1	Rappel sur les filtres particuliers . . . . .	68
3.1.1	Principe de Monte Carlo . . . . .	68
3.1.2	Échantillonnage d'importance . . . . .	69
3.1.3	Estimation récursive des poids . . . . .	70
3.1.4	Problème de dégénérescence . . . . .	71
3.1.5	Autres éléments sur les filtres particuliers . . . . .	75
3.2	Filtre particulaire adapté au problème du suivi . . . . .	76
3.3	Processus et stratégies de fusion dans les filtres particuliers . . . . .	78
3.3.1	Approches existantes . . . . .	79
3.3.2	Méthodes à apprentissage . . . . .	82
3.3.3	Théorie des possibilités . . . . .	83
3.3.4	Théories des fonctions de croyance . . . . .	87
3.3.5	Justification du choix des fonctions de croyance comme méthode de fusion	88
3.4	Formalisme des fonctions de croyance . . . . .	90
3.4.1	Masse de croyance élémentaire . . . . .	90
3.4.2	Le principe du minimum d'information . . . . .	92
3.4.3	Règles de combinaison . . . . .	92
3.4.4	Autres représentations des croyances . . . . .	94
3.4.5	Notion de distinction . . . . .	95
3.4.6	Opération d'affaiblissement . . . . .	95
3.4.7	Transformation pignistique . . . . .	96
3.4.8	Autres processus crédibilistes . . . . .	97
3.5	Approche proposée . . . . .	97
3.5.1	Définition du cadre de discernement $\Omega$ . . . . .	97
3.5.2	Calcul des masses de croyance pour chaque source . . . . .	100

3.5.3	Fusion des FDM et intégration de la masse fusionnée au filtre particulaire	104
3.5.4	Tests de suivi génériques	105
3.6	Conclusion	119
<b>4</b>	<b>Contribution à l'obtention d'une stratégie de fusion contextuel</b>	<b>121</b>
4.1	Introduction	121
4.2	Les principales règles de combinaisons	122
4.2.1	Règles particulières	122
4.2.2	Familles de règles	126
4.2.3	Représentation matricielle	128
4.2.4	Remarques supplémentaires sur les règles	129
4.3	Mécanismes et propriétés des règles de combinaisons	130
4.3.1	Propriétés algébriques	130
4.3.2	Nature des combinaisons	132
4.3.3	Influence de l'élément d'évidence	133
4.3.4	Propriétés liées aux sous-ensembles	134
4.3.5	Autres propriétés	138
4.3.6	Comparatif des règles	139
4.4	Stratégie de combinaison contextuel	139
4.4.1	Combinaison de sources anormales	140
4.4.2	Utilisation du contexte dans une stratégie de combinaison de sources non fiables	145
4.5	Expérimentations	149
4.5.1	Cas de sources défaillantes	150
4.5.2	Cas de sources leurrées	151
4.5.3	Comparatif global des règles et de la stratégie de combinaison	156
4.6	Conclusion	160
<b>5</b>	<b>Application au suivi du véhicule précédent</b>	<b>161</b>
5.1	Introduction	161
5.2	Remarques préliminaires	162
5.2.1	Remarques sur les extracteurs	163
5.2.2	Fusion et véhicules intelligents	164
5.2.3	Modifications algorithmique mineures	165
5.2.4	Remarques sur la complexité	165
5.2.5	Résumé du raisonnement scientifique formulé	166
5.2.6	Protocole expérimental	166
5.3	Robustesse du suivi dans le temps	169
5.4	Environnement d'acquisition	171
5.4.1	Conditions climatiques	171
5.4.2	Conditions d'éclairage	176
5.4.3	Milieu routier	180
5.4.4	Géométrie de la route	183
5.5	Influence du nombre de sources	186
5.6	Remarques et bilan des expérimentations	187
5.7	Conclusion	190

---

<b>Conclusion et Perspectives</b>	<b>191</b>
<b>Annexes</b>	<b>197</b>
<b>A Matrices de cooccurrences couleur pondérées</b>	<b>197</b>
A.1 Ajout de la pondération . . . . .	197
A.2 Test comparatif sur une séquence routière . . . . .	199
A.3 Test de résistance aux changements d'illumination . . . . .	201
A.4 Test d'invariance aux rotations et changements d'échelle . . . . .	202
<b>B Filtres de Kalman</b>	<b>205</b>
<b>C Preuves</b>	<b>209</b>
C.1 Preuve de la proposition 1 . . . . .	209
C.2 Preuve de la proposition 2 . . . . .	210
C.3 Preuve de la proposition 3 . . . . .	211
C.4 Preuve de la proposition 4 . . . . .	212
C.5 Contraintes liées à l'accumulation de propriétés . . . . .	212
C.5.1 Associativité, commutativité, consistance faible et règle directe . . . . .	212
C.5.2 Idempotence, commutativité et règle directe . . . . .	213
C.6 Unicité et liens entre sous-ensembles absorbants et de rejet . . . . .	213
<b>Bibliographie</b>	<b>215</b>

# Liste des figures

0.1	Exemple de faces arrières de véhicules. . . . .	2
1.1	Procédure de suivi sous une forme généraliste. . . . .	6
1.2	Exemple de boîte englobante pour un véhicule. . . . .	8
1.3	Exemple de boîte englobante orientée pour un visage extraite des travaux de Nguyen [Nguyen 04]. . . . .	8
1.4	Exemple de suivi d'une peluche. Les pixels blancs correspondent à l'estimation du contour obtenu par la méthode de suivi de Meribout [Meribout 04]. . . . .	9
1.5	Exemple de suivi d'une personne. La B-spline en cyan est estimée par l'algorithme CONDENSATION. . . . .	9
1.6	A gauche, image d'un piéton, suivi d'un traitement par squelettisation, d'une image binaire du squelette et d'une représentation graphique du squelette. . . . .	10
1.7	Exemple d'un "Minitel". Les pixels blancs correspondent à l'estimation de la projection du maillage dans l'image par la méthode de suivi [Gerard 00]. . . . .	10
1.8	Objet et son ensemble de région relatif pour une image donnée. Figure inspirée de [Cavallaro 02]. . . . .	11
1.9	L'objet suivi est une main, segmentée en plusieurs régions. Chaque région est représentée par un nœud du graphe, constituant une représentation de la main. . . . .	11
1.10	Texture très simple avec deux valeurs pixellaires : blanc et noir. . . . .	13
1.11	Texture rudimentaire construite avec deux textons. . . . .	16
1.12	Contour d'objet résumé par un ensemble de primitives de types courbes. . . . .	24
1.13	Détection de points d'intérêt de Harris multi-échelles : points détectés en jaune. . . . .	26
1.14	Image d'un véhicule de sa carte de symétrie correspondante. La présence de symétrie se traduit par des raies noires au niveau de la carte. . . . .	29
2.1	Chaîne reliant les états et les observations. . . . .	50
2.2	Segmentation de deux véhicules, puis labélisation et estimation des régions globales correspondant à chaque véhicule. . . . .	55
2.3	Champ de Markov aléatoire spatio-temporel. . . . .	56
2.4	Exemple d'occultation. . . . .	61
2.5	Exemple de situation de fouillis, où trois gardes nationaux grecques présentent des propriétés visuelles similaires. . . . .	63
2.6	Exemple d'apparences d'un véhicule à l'intérieur d'une séquence routière. . . . .	64
3.1	Exemples de deux chaînes d'état pour l'approche de Wu et Huang, figure tirée de [Wu 04]. L'exposant $s$ fait référence au filtre utilisant l'information de forme. L'exposant $c$ fait référence au filtre utilisant l'information de couleur. . . . .	82
3.2	Processus décisionnel dans le TBM. . . . .	96

3.3	Exemple d'imagettes correspondant à différentes valeurs de $X_t^{(i)}$ . Dans le cadre du filtrage particulaire, les imagettes sont tirées au hasard selon la densité d'importance et peuvent correspondre à ces différents cas. Pour le cadre vert, la situation correspond à celle décrite par l'hypothèse $\omega_1$ , cette hypothèse doit donc se voir attribuer un crédit très important dans un tel cas de figure. Le raisonnement est analogue pour le cadre rouge vis-à-vis de $\omega_2$ et pour le cadre bleu vis-à-vis de $\omega_3$ . . . . .	98
3.4	Chaîne de traitement pour l'obtention d'une FDMS pour une source et pour un élément focal. . . . .	102
3.5	Chaîne de traitement pour l'obtention d'une FDM pour une source. . . . .	103
3.6	En haut : suivi par filtre particulaire texture-couleur, en bas : suivi par FPFCC (3 sources). . . . .	107
3.7	Évaluation des performances dans le cas d'une source défaillante. . . . .	107
3.8	suivi par FPFCC (3 sources). . . . .	109
3.9	Évaluation des performances en présence de défaillances pour chaque source. . .	110
3.10	Variabilité visuelle de l'objet suivi. . . . .	110
3.11	suivi avec FPFCC (3 sources). . . . .	111
3.12	Évaluation des performances en présence de défaillances pour chaque source. . .	112
3.13	Comparaison entre l'approche probabiliste FPFCC et l'approche crédibiliste FPFCC.	113
3.14	Comparaison entre l'approche probabiliste FPFCC modifiée et l'approche crédibiliste FPFCC. . . . .	114
4.1	Schéma décrivant la stratégie de combinaison contextuelle (SCC). . . . .	148
4.2	Résultats de l'étape de classification des FDM de la SCC. <b>1</b> : classe normale, <b>2</b> : classe défaillante, <b>3</b> : classe leurrée. . . . .	151
4.3	Évaluation des performances en présence de défaillances pour différentes méthodes de fusion. . . . .	151
4.4	Résultats de l'étape de classification des FDM de la SCC. <b>1</b> : classe normale, <b>2</b> : classe défaillante, <b>3</b> : classe leurrée. . . . .	152
4.5	suivi par FPFCC (3 sources). . . . .	153
4.6	Évaluation des performances en présence d'informations erronées pour différentes méthodes de fusion. . . . .	153
4.7	Résultats de l'étape de classification des FDM de la SCC. <b>1</b> : classe normale, <b>2</b> : classe défaillante, <b>3</b> : classe leurrée. . . . .	154
4.8	suivi avec FPFCC (3 sources). . . . .	155
4.9	Évaluation des performances en présence d'informations erronées pour différentes méthodes de fusion. . . . .	156
5.1	Système d'acquisition embarqué. . . . .	162
5.2	Architecture proposée pour le suivi de la face arrière du véhicule précédent. . .	167
5.3	suivi par FPFCC. 5 sources : $S_1$ à $S_5$ (cf. 5.2.4), séquence n°1. . . . .	170
5.4	Performance moyenne pour la séquence test n°1 selon quatre approches de fusion différentes. $t_{def} = 10^{-6}$ pour le FPFCC modifié. A gauche, histogrammes pour chaque approche, la couleur correspond à la catégorie de qualité de suivi, de $r_1$ en bleu foncé à $r_5$ en vert. A droite, mêmes données présentées sous forme d'histogrammes cumulés rassemblés sur un même graphe. . . . .	171

5.5	suivi par FPFCC. 5 sources : $S_1$ à $S_5$ (cf. 5.2.4), séquence n°2. . . . .	172
5.6	Performance moyenne pour la séquence test n°2 selon quatre approches de fusion différentes. $t_{def} = 10^{-6}$ pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés. . . . .	173
5.7	Variations d'apparence du véhicule suivi dans la séquence test n°1. . . . .	173
5.8	suivi par FPFCC, 3 sources : $S_1$ , $S_2$ et $S_4$ (cf. 5.2.4), séquence n°3. . . . .	174
5.9	Performance moyenne pour la séquence test n°3 selon quatre approches de fusion différentes. $t_{def} = 10^{-6}$ pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés. . . . .	174
5.10	suivi par FPFCC. 4 sources : $S_1$ à $S_4$ (cf. 5.2.4), séquence n°4. . . . .	175
5.11	Performance moyenne pour la séquence test n°4 selon quatre approches de fusion différentes. $t_{def} = 10^{-6}$ pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés. . . . .	176
5.12	suivi par FPFCC. 3 sources : $S_1$ , $S_2$ et $S_4$ (cf. 5.2.4), séquence n°5. . . . .	177
5.13	Performance moyenne pour la séquence test n°5 selon quatre approches de fusion différentes. $t_{def} = 10^{-6}$ pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés. . . . .	178
5.14	suivi par FPFCC. 5 sources : $S_1$ à $S_5$ (cf. 5.2.4). . . . .	179
5.15	suivi par FPFCC. 4 sources : $S_1$ à $S_4$ (cf. 5.2.4), séquence n°6. . . . .	179
5.16	Performance moyenne pour la séquence test n°6 selon quatre approches de fusion différentes. $t_{def} = 10^{-3}$ pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés. . . . .	180
5.17	suivi par FPFCC. 2 sources : $S_2$ et $S_3$ (cf. 5.2.4), séquence n°7. . . . .	181
5.18	Performance moyenne pour la séquence test n°7 selon quatre approches de fusion différentes. $t_{def} = 10^{-6}$ pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés. . . . .	182
5.19	suivi par FPFCC. 2 sources : $S_2$ et $S_4$ (cf. 5.2.4), séquence n°8. . . . .	183
5.20	Performance moyenne pour la séquence test n°8 selon quatre approches de fusion différentes. $t_{def} = 10^{-6}$ pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés. . . . .	184
5.21	suivi par FPFCC. 3 sources : $S_1$ , $S_3$ et $S_4$ (cf. 5.2.4), séquence n°9. . . . .	185
5.22	Performance moyenne pour la séquence test n°9 selon quatre approches de fusion différentes. $t_{def} = 10^{-6}$ pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés. . . . .	186
5.23	suivi par FPFCC. 2 sources : $S_2$ et $S_3$ (cf. 5.2.4), extrait d'une séquence avec dénivelé. . . . .	186
5.24	suivi par FPFCC. 2 sources : $S_2$ et $S_3$ (cf. 5.2.4), extrait d'une séquence avec virage. . . . .	187
5.25	Performance moyenne pour la séquence test n°1 avec le FPFCC pour différentes fusions de sources. M varie de 1 à 5. . . . .	188
5.26	Véhicule autonome TerraMax. Figure extraite de [Braid 06]. . . . .	194
A.1	Matrice de cooccurrence pour un vecteur $\vec{d}$ donné.. . . .	198
A.2	Matrice de cooccurrence pour un vecteur $\vec{d}$ donné.. . . .	198
A.3	Différents noyaux utilisés pour les MCCP.. . . .	199

---

A.4	(a) : suivi avec cooccurrences classiques, (b) : suivi avec MCCP Epanechnikov, (c) : suivi avec MCCP Gaussien, (d) : suivi avec MCCP rationnel, (e) suivi avec Gabor. . . . .	200
A.5	Évolution du taux de suivi au cours d'une séquence pour les différentes approches texture-couleur examinées. . . . .	201
A.6	Changement artificiel de luminosité, première image de la séquence $\gamma = 1$ , dernière image $\gamma = 0,5$ . . . . .	201
A.7	Évolution du taux de suivi en fonction du temps et de la correction $\gamma$ . . . . .	202
A.8	Suivi avec MCCP gaussien pour un objet en rotation. . . . .	202
A.9	Exemple d'un changement d'échelle d'un carré vert sur fond jaune. Surfaces et périmètres ne sont pas modifiés par le même rapport. . . . .	203
A.10	Suivi avec MCCP gaussien pour un objet changeant d'échelle. . . . .	203

# Liste des tableaux

2.1	Comparaison des extracteurs d'information. . . . .	65
3.1	Comparatif des différentes approches de fusion dans un filtre particulaire par taux de suivi moyen ( $\mu$ ) et l'écart-type du taux ( $\sigma$ ). Les résultats les plus significatifs sont en gras. . . . .	118
3.2	Comparatif des différentes approches de fusion dans un filtre particulaire par taux de suivi moyen ( $\mu$ ) et l'écart-type du taux ( $\sigma$ ). . . . .	119
4.1	Différentes règles de combinaison et leurs propriétés - partie n°1. $\times$ : la règle possède la propriété correspondante. . . . .	139
4.2	Différentes règles de combinaison et leurs propriétés - partie n°2. $\times$ : la règle possède la propriété correspondante. $\star$ : la règle possède la propriété correspondante sous réserve qu'un ou plusieurs paramètres soient correctement choisis. . . . .	140
4.3	Moyenne et écart-type de $r$ pour différentes techniques de combinaison. Les résultats les plus significatifs sont indiqués en gras. $*$ : méthodes utilisant de l'information contextuelle. . . . .	157
4.4	Moyenne et écart-type de $r$ pour différentes techniques de combinaison et avec un affaiblissement contextuel. Les résultats les plus significatifs sont indiqués en gras. $\star$ : les sources défaillantes sont évincées du processus de fusion. . . . .	158
5.1	Récapitulatifs des résultats et des caractéristiques des séquences test. $*$ : mesures n'ayant pas fait l'objet d'un relevé précis ; elles sont présentées à titre indicatif. En outre, nous ne pouvons certifier que le logiciel d'acquisition respecte rigoureusement la cadence d'acquisition indiquée. . . . .	189



# Liste des notations

## Notations générales

$t$	: Variable temporelle discrète
$I_t$	: Image obtenue à l'instant $t$ à partir d'une séquence
$\vec{p} = (i, j)$	: Pixel de coordonnées $(i, j)$
$I_t(i, j)$	: Valeur de l'image pour le pixel de coordonnées $(i, j)$ dans l'image $I_t$ . Cette valeur est un triplet pour une image en couleur
$\mathcal{P}$	: Grille des pixels sur laquelle une image $I_t$ est définie
$X_t$	: Ensemble des paramètres permettant de déterminer la position de l'objet dans l'image $I_t$ . Dans le cadre des filtre particulière cet ensemble est vu comme un vecteur stochastique
$N_{entity}$	: Entier naturel représentant le cardinal d'une entité <i>entity</i>
$\#(D)$	: Cardinal d'un ensemble $D$
$r_t$	: Taux de suivi instantané

## Notations mathématiques

$ \cdot $	: Valeur absolue d'une variable réelle
$\ \cdot\ $	: Distance euclidienne
$\delta(x)$	: Distribution de Dirac
$\delta_X(x) = \delta(x - X)$	: Distribution de Dirac centrée en $X$
$*$	: Produit de convolution
$\mathbf{i}$	: Nombre complexe tel que $\mathbf{i}^2 = -1$
$\mathbb{R}$	: Ensemble des nombres réels
$\mathbb{C}$	: Ensemble des nombres complexes
$\hat{x}$	: Estimation de la variable $x$
$\xrightarrow{p.s.}$	: Convergence presque sûre
$\mathcal{N}(\mu, \sigma)$	: Loi normale de moyenne $\mu$ et d'écart-type $\sigma$
$\lfloor x \rfloor$	: Partie entière de $x$
$\lceil x \rceil$	: Partie fractionnaire de $x$
$\propto$	: Proportionnel à
$X \sim p(\cdot)$	: La variable aléatoire $X$ est échantillonnée selon la loi $p(\cdot)$

## Traitement d'images et caractérisation

$c = (r, g, b)$	: Triplet définissant le couleur d'un pixel
$q_c$	: Densité de couleur évaluée en la couleur $c$
$M_{\vec{d}}(i, j)$	: Matrice de cooccurrence pour le couple de couleurs $i$ et $j$ et le voisinage $\vec{d}$
$\lambda_{mn}$	: Moments de Legendre
$A_{mn}$	: Moments de Zernike
$\delta I_t$	: Différence d'image entre l'image actuelle et l'image précédente
$\nabla I_t$	: Gradient de l'image actuelle

## Filtrage particulière

$N$	: Nombre de particules du filtre
$X_t$	: Vecteur d'état stochastique
$Y_t$	: Vecteur aléatoire correspondant aux observations à $t$
$Y_{0:t}$	: Ensemble des valeurs de la variable aléatoire $Y$ du temps 0 au temps $t$
$p(\cdot, \cdot)$	: Densité de probabilité jointe de l'état et des observations
$q(\cdot)$	: Densité de probabilité d'échantillonnage d'importance
$w_t^{(i)}$	: Poids de la particule d'indice $i$ à l'instant $t$
$N_{eff}$	: Taille efficace du $N$ échantillon
$\lambda_t^{(i)}$	: Poids auxiliaire d'une particule d'indice $i$ à l'instant $t$
$Z_t^{(i)}$	: Particule intermédiaire
$a_n$	: Paramètre de la densité de transition

## Théorie des fonctions de croyance

$\Omega$	: Cadre de discernement
$K$	: Nombre d'hypothèses
$\omega_k$	: $k^{\text{ème}}$ hypothèse singleton
$A$	: Proposition, sous-ensemble de $\Omega$
$2^\Omega$	: Ensemble des sous-partie de $\Omega$
$m(\cdot)$	: Fonction de masse
$Ev$	: Élément d'évidence
$\mathcal{F}_m^\Omega$	: Ensemble des éléments focaux de $m$
$bel(\cdot)$	: Fonction de crédibilité
$pl(\cdot)$	: Fonction de plausibilité
$q(\cdot)$	: Fonction de communalité
$S_i$	: $i^{\text{ème}}$ source d'information
$m[S](\cdot)$	: Fonction de masse issue de la source $S$
$M$	: Nombre de sources
$A^\omega(\cdot)$	: Fonction de masse simple
$\eta(\cdot)$	: Mesure de non-spécificité.
$\mathfrak{B}^\Omega$	: Ensemble des fonctions de masse définies sur $\Omega$
$\oplus$	: Règle de combinaison de Dempster.
$\odot$	: Règle de combinaison conjonctive
$\oslash$	: Règle de combinaison disjonctive
$\oslash$	: Règle de combinaison conjonctive prudente
$\oslash$	: Règle de combinaison disjonctive hardie
$\oslash$	: Règle disjonctive exclusive
$\odot$	: Règle de combinaison quelconque
$\mu(\cdot)$	: Fonction de masse simple généralisée
$m_{\text{symbole}}$	: Fonction de masse résultant de la combinaison suivant la règle symbolisée par <i>symbole</i>
$m_{\text{acronyme}}$	: Fonction de masse résultant de la combinaison suivant la règle référencée par l'acronyme <i>acronyme</i>
$\kappa$	: Conflit (masse de croyance conflictuelle) généré lors de la combinaison conjonctive
$\alpha$	: Coefficient d'affaiblissement

$m[\alpha, S]$	: Fonction de masse affaiblie par le coefficient $\alpha$
$BetP$	: Probabilité pignistique
$w(\cdot)$	: Fonctions des poids conjonctifs
$v(\cdot)$	: Fonctions des poids disjonctifs
$\xi_{l,j}$ et $\alpha_{L,j}$	: Paramètres du modèle de fonction de masse
$R_i$	: Coefficient de fiabilité
$\beta(\cdot)$	: Fonction de poids de Lefèvre
$\gamma_1(\cdot)$ et $\gamma_2(\cdot)$	: Fonction de poids de Florea
$\delta_1(\cdot, \cdot)$ et $\delta_2(\cdot, \cdot)$	: Fonction de poids de Martin
$\epsilon$	: Paramètre d'affaiblissement de la règle DPCR.
$\vec{m}$	: Représentation vectoriel d'une FDM $m$
$M_{\odot}^m$	: Représentation matricielle de la combinaison avec la FDM $m$ selon la règle $\odot$
$\psi_A(\cdot, \cdot)$	: Fonction caractéristique d'une règle de combinaison directe
$\mathfrak{R}$	: Ensemble de règles
$\odot_d$	: Règle de combinaison duale de la règle $\odot$
$t_{def}$	: Seuil de détection de la défaillance d'une source
$var_t$	: Variance pondérée des particules
$t_{var}$	: Seuil de détection d'une situation de leurre



# Liste des abréviations

## Traitement d'images et caractérisation

RVB/RGB	: Système trichromatique de représentation de la couleur selon les couleurs primaires (rouge, vert et bleu)
XYZ et YCbCr	: Systèmes de représentation de la couleur selon une composante de luminance et deux composantes chromatiques
HSV	: Système de représentation de la couleur selon la teinte, la saturation et l'éclat
ACP	: Analyse en composantes principales
ACI	: Analyse en composantes indépendantes
LoG	: Laplacien de Gauss
DoG	: Différences de gaussiennes
ECMA	: Équation de contrainte du mouvement apparent

## Fonctions de croyance

TBM	: Transferable Belief Model
FDM	: Fonction de masse
FDMS	: Fonction de masse simple
FDMSG	: Fonction de masse simple généralisée
FDPC	: Fonction de poids conjonctifs
FDPD	: Fonction de poids disjonctifs
TBG	: Théorème de Bayes généralisé
PCR	: Proportional conflict redistribution
DPCR	: Discounted proportional conflict redistribution
MDPCR	: Mixed discounted proportional conflict redistribution
WAO	: Weighted average operator
OCC	: Opérateur de fusion contextuel

## Algorithmes

EM	: Espérance-maximisation
CONSENSATION	: Algorithmes de suivi par filtrage particulaire proposés par Isard et Blake
ICONDENSATION	
MAP	: Maximum a posteriori
MRF	: Champs aléatoires de Markov
ICM	: Recuits simulés itérés
CRF	: Champs aléatoires conditionnels
MHT	: Multiple hypotheses tracking
PDAF	: Probability data association filtering
JPDAF	: Joint probability data association filtering
MC	: Méthodes de Monte Carlo
MCMC	: Méthode de Monte Carlo par chaîne de Markov
FP	: Filtre particulaire

SIR	: Sampling importance resampling
SIS	: Sampling importance sampling
FPA	: Filtre particulaire auxiliaire
SVM	: Séparateurs à vaste marge
k-ppv	: Algorithme des k plus proches voisins
FPFC	: Filtre particulaire avec fusion crédibiliste
FPFB	: Filtre particulaire avec fusion bayésienne
FPFP	: Filtre particulaire avec fusion possibiliste
FPFCC	: Filtre particulaire avec fusion crédibiliste contextuel (via l'OCC)

# Introduction

Les thématiques liées de près ou de loin à l'intelligence artificielle ont toujours suscité un vif intérêt et ce bien au delà de la simple communauté scientifique. Depuis des siècles, les chercheurs de tous horizons tentent de mettre au point une machine ou un dispositif autonome quelconque capable d'imiter voire de parfaire un comportement cognitif humain. La vision par ordinateur s'inscrit en droite lignée de ces démarches. Elle a pour but ultime de doter une machine d'une perception visuelle. Cet objectif reste à l'heure actuelle utopiste car le problème soulevé va bien au delà des frontières déjà existantes en termes de calcul algorithmique et de modélisation mathématique. Les thèmes concrètement abordés par les chercheurs en vision par ordinateur sont plus modestes et réalistes. Ces thématiques sont par exemple : la reconnaissance d'objets, le suivi d'objets, l'indexation d'image, la reconstruction de scène 3D ou encore l'estimation du mouvement d'un objet.

Les travaux présentés dans ce mémoire de thèse ciblent la problématique de suivi d'objets dans les séquences vidéos. Il s'agit d'être capable de positionner, d'isoler un ou plusieurs objets au sein d'une scène acquise à l'aide d'un dispositif de type caméra. Une fois capté par une caméra, le signal optique est converti en signal numérique, qu'il est possible de conserver sous forme matricielle dans la mémoire d'un système informatique. Pour une séquence vidéo, cette matrice possède trois dimensions : une temporelle et deux spatiales. Pour une coordonnée temporelle fixée, nous obtenons une image de la séquence. Cette image correspond à une grille dont les éléments sont appelés pixels. Un triplet de données est associé à chaque pixel, permettant de lui attribuer une couleur. Dans chaque image, un algorithme de suivi d'objets doit déterminer de manière plus ou moins précise la zone comprenant les pixels relatifs à l'objet recherché.

La notion d'objet, relativement vague dans son usage commun, mérite d'être précisée. Un objet est défini dans ce mémoire comme une entité dont les contours délimitent une région de l'espace de dimensions finies et ayant une certaine cohérence visuelle. Concernant les dimensions, celles-ci doivent être suffisamment faibles pour que l'intégralité de l'objet puisse entrer dans le champ d'une caméra. La cohérence visuelle traduit le fait que l'aspect de l'objet possède des propriétés détectables et analysables permettant de le dissocier du reste de la scène.

Il n'existe pas de méthode universelle de suivi d'objets. Les méthodes existantes prennent appui sur des approches et des hypothèses variées. Selon ces méthodes, des thématiques annexes sont souvent abordées. Certains concepts sont empruntés entre autres :

- au traitement et à l'analyse d'images numériques,
- au filtrage numérique,
- à la reconnaissance de forme et à l'apprentissage,
- à la fusion de donnée.

Les méthodes de suivi ont par conséquent des avantages et des inconvénients qui leur sont propres. Il est donc important de préciser le contexte applicatif de ces travaux de thèse, afin de se donner des critères permettant de choisir parmi l'ensemble des méthodes, celles qui présentent le meilleur compromis.

## Contexte

Concernant l'acquisition de données, nous nous plaçons dans le cadre d'un suivi mono-caméra couleur. Nous excluons donc de notre étude les méthodes de stéréo-vision ou multi-caméra. Sont également exclues les approches multi-capteur ; en effet des capteurs radars ou sonores par exemple peuvent être exploités pour cibler un objet. Les méthodes conçues pour des images en niveaux de gris peuvent être utilisées de différentes manières sur des images couleur. Il appartient alors de justifier leur emploi par rapport à des méthodes couleur. Toutefois la dimension colorimétrique devra être exploitée à un étage quelconque de l'architecture proposée. Les informations *a priori* sur l'objet à suivre seront de plus limitées, ainsi de nombreuses méthodes d'apprentissage ne pourront être employées.

Pour ce qui est du domaine applicatif de la présente thèse, celle-ci s'inscrit dans le cadre des systèmes de transport intelligents. L'objet à suivre est la face arrière d'un véhicule. Ce véhicule en mouvement précède le véhicule dans lequel la caméra est disposée : la caméra est donc également en mouvement. En conséquence, des changements dynamiques importants auront lieu dans les séquences acquises. La figure 0.1 présente un cas type d'image extraite d'une vidéo à traiter. Ces

**Figure 0.1** : Exemple de faces arrières de véhicules.

séquences sont qualifiées de routières. Les intérêts applicatifs d'un algorithme capable de suivre un véhicule dans des séquences routières sont multiples :

- assistance à la conduite (dépassement automatisé du véhicule précédent) ;
- amélioration des systèmes de sécurité routière (non respect des distances de sécurité, comportements à risque) ;
- conduite en convois autonomes (les véhicules suivent de façon autonome le véhicule de tête, les moyens humains nécessaires au convois sont réduits).

Le laboratoire LITIS et son équipe "Systèmes de Transport Intelligents" ont pour ambition de lever les verrous scientifiques liés, entre autres, au suivi de véhicules dans les séquences routières. De nombreux travaux ont déjà été réalisés dans le cadre du projet ARCOS ou de collaborations auprès des constructeurs et équipementiers automobiles tels que Valéo et Renault. Plus récemment, le pôle de compétitivité Mov'eo offre des perspectives de recherche supplémentaires pour les systèmes de transport intelligents. Ces travaux de thèse ont vocation à être valorisés dans cet axe.

## Démarche adoptée

La démarche scientifique naturelle conduit dans un premier temps à analyser l'état de l'art des méthodes de suivi. Il est important de faire un tour d'horizon des méthodes existantes, afin de pouvoir justifier de nos choix et de nos contributions. Cette analyse sera proposée dans les deux premiers chapitres du mémoire. Le premier chapitre s'attachera à apporter des réponses à un premier problème : comment modéliser correctement un objet au sein d'une machine pour pouvoir le suivre ? Plus précisément, les différentes représentations informatiques possibles pour un objet seront passées en revue afin de retenir la plus adaptée au suivi de face arrière des véhicules. Ensuite, la caractérisation d'objets, seconde étape essentielle de la modélisation d'un objet, sera examinée sous trois angles principaux :

- la couleur et les textures,
- la forme,
- le mouvement.

Cette classification des méthodes de caractérisation se justifie par la différence de nature entre les informations extraites. Notons que nous contribuerons dans cette partie à fournir un des opérateurs texture-couleur présentés. Les détails sur la conception de cet opérateur sont donnés en annexe A. Dans le second chapitre, les procédures de suivi, c'est-à-dire le cœur algorithmique permettant le traitement itératif des images, seront présentées. Nous pourrions voir que sur nos séquences les méthodes à base de filtres particulières constituent une procédure de suivi particulièrement avantageuse.

Suite aux deux études précédentes, une incompatibilité entre les méthodes existantes et les défis imposés par la nature des séquences à traiter sera mis en avant. En effet, une certaine catégorie d'événements inéluctables dans les scènes naturelles routières ne peuvent être gérés en couplant simplement une méthode de modélisation à une procédure de suivi. La solution soutenue dans cette thèse est alors d'exploiter plusieurs méthodes de caractérisation, aux propriétés complémentaires, au sein d'une méthode de suivi.

Pour parvenir à un tel résultat la fusion de données est indispensable. La fusion de données est un thème de recherche vaste qui nécessite lui aussi une étude approfondie. En troisième partie du mémoire, nous étudierons les possibilités de fusion de données au sein d'une procédure de suivi par filtrage particulière. Les méthodes d'extractions retenues formeront autant de sources d'information à fusionner. L'état de l'art montrera que le cadre dit des fonctions de croyances constitue un outil performant. Une architecture globale et unifiée sera alors introduite. Elle sera validée en termes de suivi d'objets à l'aide d'une gamme de tests où la variation des paramètres permettra une évaluation objective. Cette architecture constitue la contribution méthodologique principale du mémoire. Néanmoins, les résultats montreront que la nouvelle approche ne répond pas convenablement à l'ensemble des événements perturbateurs liés aux séquences routières.

En s'appuyant sur le constat que l'approche proposée est complémentaire vis-à-vis du filtrage particulière classique, le quatrième chapitre s'attachera à une étude plus théorique de la fusion de données dans le cadre des fonctions de croyance. Afin d'unifier les deux filtres complémentaires, un opérateur de combinaison de sources sera proposé en alternative. Cet opérateur constitue la contribution théorique principale de ces travaux de thèse. Son intérêt sera justifié en comparaison à d'autres méthodes de fusion dans le cadre des fonctions de croyance.

Enfin, le dernier chapitre proposera une architecture plus concrète pour le suivi spécifique de la face arrière du véhicule précédant le véhicule instrumenté. La méthode sera poussée dans ses retranchements afin de faire ressortir les perspectives de travail les plus pertinentes.

## Abus de langage et anglicisme

Avant d'aller plus loin, nous tenons à préciser que certains abus de langage seront commis dans le document. Afin de faciliter la lecture de ce dernier, nous précisons dès à présent les principaux. Un pixel fait référence à un point fixe d'une grille image, or nous dirons qu'un pixel appartient à un objet, alors que cet objet peut être en mouvement. On parlera éventuellement alors de *déplacements des pixels*.

La variable temporelle sera notée  $t$ , et sera dans la majorité des cas une variable discrète, du fait de la nature numérique des supports analysés. Nous ne préciserons pas l'aspect discret de cette variable, le contexte dans lequel elle est employée permettant assez nettement de lever l'ambiguïté.

Par rapport au suivi d'objets, nous parlerons souvent de *position d'un objet*, il s'agira en réalité plus exactement de l'ensemble des paramètres permettant de définir la position de cet objet. Le nombre et la nature de ces paramètres dépendent de la représentation choisie de l'objet. La "position de l'objet" est la sortie attendue d'un algorithme de suivi. Concernant la fusion de données, nous confondrons souvent sources d'information et extracteurs d'information appliqués aux images. D'autres abus plus ciblés seront évoqués au fil du mémoire.

Enfin la plupart des termes anglo-saxons utilisés dans les domaines abordés dans le mémoire trouvent une traduction adéquate en langue française. Nous nous efforcerons d'en traduire un maximum. Toutefois, un nombre très limité n'ont pas d'équivalents français pertinents ; nous nous permettrons alors de les conserver. Pour plus de commodité, certains acronymes anglais très fréquemment utilisés dans la communauté seront également conservés.

# Chapitre 1

## Modélisation d'objets pour le suivi

### Sommaire

---

<b>1.1</b>	<b>Introduction et positionnement du problème de suivi . . . . .</b>	<b>5</b>
<b>1.2</b>	<b>Représentation des objets . . . . .</b>	<b>7</b>
1.2.1	Boîte englobante . . . . .	8
1.2.2	Contour . . . . .	9
1.2.3	Squelette . . . . .	10
1.2.4	Maillage . . . . .	10
1.2.5	Ensemble de régions . . . . .	10
1.2.6	Ensemble de points d'intérêt et graphes . . . . .	11
1.2.7	Représentation choisie . . . . .	12
1.2.8	De la représentation vers le modèle d'objet . . . . .	12
<b>1.3</b>	<b>Extraction de caractéristiques couleur et texture . . . . .</b>	<b>12</b>
1.3.1	Méthodes colorimétriques . . . . .	13
1.3.2	Méthodes texture . . . . .	15
1.3.3	Méthodes texture-couleur . . . . .	19
1.3.4	Bilan sur l'information texture-couleur . . . . .	20
<b>1.4</b>	<b>Extraction de caractéristiques de forme . . . . .</b>	<b>21</b>
1.4.1	Détection de contours . . . . .	22
1.4.2	Squelettisation . . . . .	23
1.4.3	Méthodes structurelles . . . . .	24
1.4.4	Points d'intérêt et graphes . . . . .	25
1.4.5	Attributs statistiques . . . . .	27
1.4.6	Cartes de symétrie . . . . .	29
1.4.7	Autres méthodes de caractérisation de forme . . . . .	29
1.4.8	Bilan sur l'information de forme . . . . .	30
<b>1.5</b>	<b>Extraction de caractéristiques de mouvement . . . . .</b>	<b>30</b>
1.5.1	Méthodes basées sur la différence inter-images . . . . .	31
1.5.2	Méthodes d'analyse du fond de la scène . . . . .	36
1.5.3	Extraction en couches de mouvement . . . . .	37
1.5.4	Méthodes à base de mouvement cohérent . . . . .	38
1.5.5	Bilan sur l'information de mouvement . . . . .	38
<b>1.6</b>	<b>Conclusion . . . . .</b>	<b>39</b>

---

### 1.1 Introduction et positionnement du problème de suivi

Comme pour toutes les disciplines voisines, la complexité et l'efficacité des méthodes de suivi ont crû avec le progrès des outils de calculs informatiques. Néanmoins, nous constatons que

depuis les prémices de la recherche en suivi, la dépendance des méthodes aux vidéos traitées est une constante. Certes, certaines méthodes sont plus génériques que d'autres, mais l'universalité n'est toujours pas d'actualité.

Ces aspects sont généralement méconnus des personnes extérieures à la discipline, car pour l'humain, le suivi d'objets est un acte naturel presque inconscient et ne nécessite pas d'activité intellectuelle intense, tout du moins en apparence. Or ce qui est effectué avec une facilité déconcertante par l'homme, se trouve être bien moins évident pour une machine.

En voulant produire un algorithme de suivi, les scientifiques se sont heurtés à un premier problème : comment représenter et reconnaître un objet ? Pour ce qui est de la reconnaissance, cela revient d'un point de vue informatique à être capable de calculer des caractéristiques sur les pixels appartenant à l'objet qui permettent de le distinguer du reste de la scène. On parle alors de caractérisation d'objets. La nature de ces caractéristiques est très large, nous reviendrons sur ce point plus loin dans ce chapitre.

Concernant la représentation structurelle de l'objet, les auteurs cherchent à résumer l'apparence de l'objet sous une forme compressée, par exemple : boîte englobante, maillage, contours, ensembles de région, etc.

La représentation ne peut être totalement dissociée des caractéristiques. **C'est le couple représentation/descripteurs qui forme le modèle de l'objet.** Une représentation de l'objet ne peut être appareillée à une observation qu'en extrayant des caractéristiques de l'observation.

Une fois cette première difficulté surmontée, il faut à présent être capable de suivre ce modèle dans le temps. Il convient donc de mettre au point un algorithme dont les entrées sont les caractéristiques du modèle évoqué à l'instant  $t$  ainsi que la trajectoire passée de l'objet et dont la sortie est la position actuelle de l'objet. Nous parlerons dans ce mémoire de procédure de suivi. La procédure sous une forme simple et générale est rappelée par la figure 1.1. Il existe

**Figure 1.1 :** Procédure de suivi sous une forme généraliste.

naturellement une grande variété de procédures de suivi, il est possible par exemple de prendre

en compte les positions passées de manières plus ou moins explicites et ne prendre en compte que la position précédente, ou bien encore utiliser tout l'historique. Ce qu'il est important de retenir est qu'un algorithme de suivi se décompose en **deux phases principales : la modélisation de l'objet et la procédure de suivi**.

Dans la suite du chapitre, nous commencerons par détailler les différentes grandes catégories de représentation des objets. Afin de pouvoir constituer un modèle d'objet, les méthodes de caractérisation d'objets seront présentées juste après les représentations d'objets. Les liens entre représentation et caractérisation seront évoqués. A la vue de l'ensemble des approches rassemblées dans ce chapitre, une large gamme de modèles pourra être conçue. Les procédures de suivi seront quant à elles évoquées dans le second chapitre. Comme le modèle et la procédure forment un tout, le choix des méthodes de caractérisation sera affiné en conclusion du chapitre 2.

Veillez noter enfin que ces deux premiers chapitres n'ont pas pour vocation de constituer un état de l'art exhaustif en suivi. Il existe de nombreuses manières de constituer un tel état de l'art et la prolifération des algorithmes de suivi dans la dernière décennie rend très ambitieux l'établissement d'un état de l'art exhaustif. Toutefois, nous cherchons à faire un tour d'horizon des principaux algorithmes de suivi existant dans la littérature afin d'en discuter l'intérêt par rapport aux visées applicatives de ces travaux de thèse. Un état de l'art récent et pertinent est donné par Yilmaz *et al.* [Yilmaz 06] et constitue un point de vue complémentaire par rapport au notre.

Avant de présenter plus en avant ces méthodes, nous tenons à donner quelques notations générales appelées à être très fréquemment utilisées :

- On notera  $I_t(i, j)$  la fonction discrète représentant la séquence d'images numériques à traiter.

$$I_t(i, j) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

- $t$  est la dimension temporelle, ainsi à  $t$  fixée  $I_t$  est une image.  $i$  est l'indice des lignes et  $j$  l'indice des colonnes.  $I_t(i, j)$  est un triplet de couleur.
- $\vec{p} = (i, j)$  est le pixel de coordonnées  $(i, j)$ .
- $\mathcal{P}$  est la grille des pixels sur laquelle une image  $I_t$  est définie.

Une liste de notations et d'abréviations est par ailleurs disponible dans la liste de notations située en préambule du mémoire.

## 1.2 Représentation des objets

Cette partie a pour but de lister les manières les plus classiques de résumer la structure d'un objet d'un point de vue informatique. La représentation de l'objet n'est qu'une facette de la modélisation. Dans la majorité des cas, la représentation choisie peut être combinée à n'importe quelle méthode de caractérisation. Seules quelques unes imposent le recours à certains procédés de traitement d'images. Étant donné notre cadre applicatif, nous n'aborderons pas dans cette section les représentations spécifiques aux objets non-rigides.

### 1.2.1 Boîte englobante

La boîte englobante se définit comme le rectangle le plus petit dont la surface contient entièrement celle de l'objet-cible. C'est la manière la plus simple de caractériser la position d'un objet au sein d'une image donnée d'une séquence. La boîte est entièrement définie par quatre paramètres :

- la position du centre de la boîte en abscisse,
- la position du centre de la boîte en ordonnée,
- la largeur de la boîte,
- la hauteur de la boîte.

L'objet est donc représentable par un vecteur de quatre coordonnées. La figure 1.2 donne un exemple de véhicule encadré par sa boîte englobante. Il est possible d'ajouter une cinquième co-

**Figure 1.2 :** Exemple de boîte englobante pour un véhicule.

ordonnée représentant l'orientation de la boîte. Cette cinquième est utile lorsque l'objet subit des rotations selon l'axe orthogonal au plan image. La boîte sélectionne ainsi une zone plus proche du contour de l'objet et donne une information supplémentaire sur son déplacement. La figure 1.3 illustre le rôle d'une coordonnée d'orientation. Notons enfin qu'il est possible de choisir une autre

**Figure 1.3 :** Exemple de boîte englobante orientée pour un visage extraite des travaux de Nguyen [Nguyen 04].

forme que le rectangle, notamment un cercle ou une ellipse, parfois plus adaptée au type d'objet suivi [Comaniciu 00] ou encore des polygones moins réguliers que le rectangle [Georgescu 04]. Les paramètres géométriques codant cette forme remplacent ceux précédemment cités. Le principal reproche adressé à ce type de représentation de l'objet est le caractère approximatif du positionnement de l'objet. Cela dépend en réalité de la précision du suivi qui est attendue en application. Plus la boîte englobante possède de paramètres, plus elle tend vers une représentation de type

contour.

### 1.2.2 Contour

Il se définit comme l'ensemble des pixels de l'image appartenant à la frontière entre la région correspondant à l'objet et le reste de l'image. Le contour permet une localisation précise de l'objet dans chaque image de la scène. De manière idéale, le contour est large d'un pixel. Il peut être stocké à l'aide d'une image binaire, où les pixels blancs sont des pixels de contour. L'approche retenue par Méribout *et al.* [Meribout 04] donne un exemple de représentation de l'objet par contour (*cf.* figure 1.4). Il est également possible de représenter l'objet par une approximation de

**Figure 1.4 :** Exemple de suivi d'une peluche. Les pixels blancs correspondent à l'estimation du contour obtenu par la méthode de suivi de Meribout [Meribout 04].

son contour. Cette approximation peut être réalisée par des courbes paramétriques [Lefèvre 01], des courbes de Bézier ou encore des B-splines [Isard 98a, Isard 98b]. Les B-splines constituent une des représentations les plus utilisées en pratique, car l'approximation du contour est bonne et la représentation du contour se résume en un ensemble de points appelés points de contrôle. Les points de contrôle conditionnent localement la courbure du contour. La manipulation informatique de ces points est plus simple que celle d'une image binaire de contour. Un des algorithmes de suivi les plus célèbres, CONDENSATION de Isard et Blake [Isard 98a], utilise une B-spline (*cf.* figure 1.5).

**Figure 1.5 :** Exemple de suivi d'une personne. La B-spline en cyan est estimée par l'algorithme CONDENSATION.

### 1.2.3 Squelette

Le squelette est d'une certaine manière le pendant du contour en terme de représentation. Il est très principalement utilisé pour le suivi de piétons, le squelette ne se prêtant pas à tout type d'objets en terme de représentation. En général le squelette n'est pas paramétré, il est obtenu en sortie d'un algorithme de squelettisation, coûteux en temps de calcul et sensible au bruit. S'il est paramétré, il l'est en général au niveau de ses extrémités et de ses nœuds. Il correspond alors plutôt à une représentation par graphes (*cf.* 1.2.6). Dans ses travaux de thèse, Suard [Suard 06] utilise ces différents niveaux de représentations squelettiques (*cf.* figure 1.6).

**Figure 1.6 :** A gauche, image d'un piéton, suivi d'un traitement par squelettisation, d'une image binaire du squelette et d'une représentation graphique du squelette.

### 1.2.4 Maillage

Un maillage est une surface tridimensionnelle dont les arrêtes et les sommets résument la forme tridimensionnelle d'un l'objet. Plus l'objet est de forme simple et plus le maillage possède de sommets, meilleure est l'approximation. La figure 1.7 extraite de Gérard et Gagalowicz [Gerard 00] donne un exemple de représentation d'objets par maillage. Le maillage suppose

**Figure 1.7 :** Exemple d'un "Minitel". Les pixels blancs correspondent à l'estimation de la projection du maillage dans l'image par la méthode de suivi [Gerard 00].

une connaissance *a priori* particulièrement précise de l'objet. De plus, il se prête mieux à la représentation d'objets avec des arrêtes marquées, plutôt que des objets avec de grandes courbures. Le maillage est typiquement utilisé dans un cadre 3D, il faut donc en général, via l'algorithme de suivi, trouver la projection du maillage dans le plan image qui correspond le mieux à l'observation. Cette représentation permet alors une localisation précise dans l'espace.

### 1.2.5 Ensemble de régions

Quand un objet possède des régions aux propriétés visuelles nettement distinctes, il peut être intéressant de subdiviser le problème de suivi au niveau de ces régions (*cf.* fig 1.8). Cela

est aussi particulièrement appréciable pour des objets articulés. Une région est individuellement représentée par son contour et son centre, aussi appelé centroïde. Il existe en effet une certaine dualité entre contour et région. La topologie des régions finit de représenter l'objet-cible. Cette représentation est par exemple utilisée par Cavallaro [Cavallaro 02] ou dans Huang *et al.* [Huang 02]. Le fait de décrire les objets plus localement offre plus de précision qu'une mesure

**Figure 1.8 :** Objet et son ensemble de région relatif pour une image donnée. Figure inspirée de [Cavallaro 02].

globale. De plus, cela permet également de s'adapter plus facilement à certains événements tels que les occultations partielles (*cf.* 2.7.1). En revanche l'ensemble de région est plus complexe à mettre en œuvre et requiert une phase de segmentation dont l'efficacité n'est pas toujours garantie.

### 1.2.6 Ensemble de points d'intérêt et graphes

Un ensemble de points d'intérêt bien particuliers peut constituer une façon de représenter un objet. La distribution de ces points dans l'image est une forme de représentation. Elle est toutefois peu parlante pour un utilisateur humain. On lui préfère donc souvent une extension sous forme de graphe, reliant les points qui constituent les sommets du graphe. La création des arrêtes du graphe est un procédé à définir et fait partie de la représentation.

On note par ailleurs qu'une représentation par régions peut être aussi traduite sous forme de graphe. Les centroïdes des régions forment les sommets, la notion de voisinage définit les arrêtes. Un graphe possède également certaines similarités avec un maillage, la différence étant que le maillage est nécessairement en 3D. Dans la figure 1.9 issue de Graciano *et al.* [Graciano 07], l'objet suivi est représenté par un graphe. Le problème des représentations type ensemble de points est

**Figure 1.9 :** L'objet suivi est une main, segmentée en plusieurs régions. Chaque région est représentée par un nœud du graphe, constituant une représentation de la main.

qu'elle impose une étape de détection de points d'intérêts. Cette détection n'est pas toujours stable au cours de la séquence. De plus, certains objets ne présentent pas de point d'intérêt particulier (notamment des objets sphériques et lisses). Les défauts des graphes dépendent de la façon dont ils sont conçus, et du sens que l'on donne à leurs sommets et arrêtes. Dans le cas où un sommet est une région, les graphes souffrent des mêmes défauts que la représentation par ensemble de régions.

### 1.2.7 Représentation choisie

Par rapport à l'application de suivi de faces arrières de véhicules, notre choix de représentation s'est porté sur la boîte englobante. En effet l'objet à suivre est de forme rudimentaire (rectangulaire) : la boîte englobante offre donc un excellent compromis entre simplicité de mise en œuvre, de calcul et de précision. De plus, comme un véhicule ne subit pas de rotation d'axe orthogonal au plan image, la coordonnée d'orientation n'a pas besoin d'être prise en compte. Dans toute la suite du mémoire, on désignera par le vecteur  $X_t$  l'ensemble des paramètres codant la représentation de l'objet dans l'image à l'instant  $t$ . Ce vecteur définit la "position actuelle" de l'objet.

### 1.2.8 De la représentation vers le modèle d'objet

Nous avons pu voir qu'il existe des équivalences entre les différentes représentations, bien qu'elles possèdent certains avantages les unes par rapport aux autres. Étant donné une représentation choisie de l'objet, et étant donné des informations *a priori* sur l'objet, il est possible de construire un modèle complet d'objet. Cette construction se fait en appliquant des méthodes d'extraction sur la base de données d'information *a priori*, générant ainsi des descripteurs ou attributs de l'objet. Plus la base est grande plus le modèle est robuste. Si elle est suffisamment grande, des méthodes d'apprentissage peuvent être employées.

Le couple représentation/descripteurs forme donc le modèle de l'objet. Le choix des méthodes d'extraction ou de caractérisation est crucial dans l'élaboration de modèles fiables. La section suivante présente différentes méthodes pouvant faire l'objet de ce choix. Notons toutefois que les choix des méthodes de caractérisation doivent se faire en harmonie avec la représentation de l'objet sélectionnée.

## 1.3 Extraction de caractéristiques couleur et texture

Quand on demande à un être humain de distinguer deux objets, la couleur s'impose dans la majorité des cas comme le premier critère discriminant. Il est donc logique de chercher à exploiter la couleur dans un algorithme de suivi. La couleur dépend des valeurs prises par les pixels. A cet effet, nous la rapprochons d'une seconde notion qui repose également en partie sur ces valeurs pixellaires : la texture. La texture ajoute une information spatiale par rapport à la couleur. Par exemple, une texture de base, comme celle présentée en figure 1.10 peut être décrite comme une succession de pixels blancs et noirs, les valeurs des pixels et leur voisinage sont donc pris en compte.

Étant donné la façon dont doit être traitée l'information dans ces deux domaines, nous développons conjointement dans cette partie des méthodes de caractérisation couleur, texture,

**Figure 1.10** : Texture très simple avec deux valeurs pixellaires : blanc et noir.

et texture-couleur.

### 1.3.1 Méthodes colorimétriques

La couleur correspond à un spectre fréquentiel d'une onde lumineuse. Les physiciens préfèrent représenter ces fréquences en longueurs d'onde. Si le spectre colorimétrique ne possède qu'une seule composante, on parle de composante monochromatique. La lumière de manière générale, et plus particulièrement la lumière naturelle (celle présente dans nos séquences), est un mélange de différentes composantes monochromatiques.

Toutefois, il faut bien noter que la couleur est avant tout une notion humaine, car la couleur est une perception, une sensation que nous éprouvons au contact de la lumière. Physiquement parlant, il existe des "couleurs invisibles" (infra-rouges, et ultra-violets) qui en réalité possèdent des fréquences que notre récepteur, l'œil, ne peut capter.

#### Espaces de représentation de la couleur

Les physiciens ont montré que trois composantes monochromatiques suffisent pour générer l'ensemble du spectre visible (théorie trichrome des récepteurs visuels de Young-Helmholtz) : le rouge, le vert, et le bleu. En informatique, une couleur est donc représentée par trois variables discrètes. La valeur prise par ces variables correspond à l'amplitude de la composante dans l'élaboration de la couleur. Cette valeur discrète est très généralement comprise entre 0 et 255 et donc stockée sur un octet. Ce système de représentation colorimétrique de base est appelé RVB ou RGB en anglais.

En analysant l'information de couleur, les traiteurs de signaux se sont aperçus que l'espace tridimensionnel RGB posait certains problèmes de représentation de l'information. Ils ont alors eu l'idée d'utiliser des méthodes mathématiques pour changer d'espace de représentation. Pour ne pas perdre d'information, tout espace colorimétrique doit être de dimension 3.

Nous n'avons pas pour ambition de présenter l'ensemble des espaces colorimétriques conçus à ce jour, mais simplement de présenter les principales catégories d'espaces. Pour plus de détails, nous renvoyons les lecteurs vers une étude de Vandenbroucke et Macaire [Vandenbroucke 05]. Outre l'espace primaire RGB, il existe donc :

- Les espaces Luminance-Chrominance : ces espaces font en sorte que l'information de luminance soit ramenée sur une seule composante et l'information chromatique sur les deux

- autres. Parmi ces espaces, on trouve entre autres XYZ et YCbCr (Y est la composante de luminance dans les deux cas).
- Les espaces perceptuels : ces espaces sont conçus pour qu'une distance colorimétrique perçue par l'homme se rapproche de la distance euclidienne dans le nouvel espace. Lab et Luv sont de tels espaces. Notons qu'eux aussi ont ramené l'information de luminance sur une seule composante (composante L). Un autre espace appelé HSV (hue, saturation, value) tente de reproduire respectivement des notions telles que la teinte, la saturation et l'éclat d'une couleur. Notons que la composante V est très proche d'une composante de luminance.
  - Les espaces à axes indépendants : les axes de ces espaces sont obtenus par analyse en composante principale (ACP) [Ohta 80] ou en composantes indépendantes (ACI) [Liu 91]. Ces espaces sont plus complexes à obtenir donc plus rarement utilisés.

Il est intéressant de se pencher sur le choix d'un espace adapté avant d'analyser une séquence d'images couleur. Un bon espace peut faciliter la tâche des algorithmes utilisés par la suite. En revanche pour ne pas trop alourdir le processus il est préférable de maintenir ce choix pour toute la séquence. Dans nos expériences, nous utiliserons seulement deux espaces : RGB et HSV. Si l'espace n'est pas précisé, il s'agira par défaut de RGB. Dans le cas contraire l'utilisation de HSV sera précisée. La transformation RGB/HSV est peu coûteuse et principalement utilisée en cas de changements d'illumination (*cf.* 2.7.4), car ces changements n'affectent alors qu'une seule composante au lieu des trois.

### Histogrammes couleur et autres attributs statistiques

De même que pour les images en niveaux de gris, l'histogramme est une façon simple et efficace d'obtenir une caractérisation d'un objet. L'histogramme revient à dénombrer l'ensemble des pixels de même couleur, il s'agit donc d'une caractérisation statistique. Si l'on applique la même méthode qu'en niveaux gris il faut représenter l'ensemble des couleurs sur un axe, ce qui physiquement n'a pas de sens. Il est alors possible de choisir un ordre partiel ou quelconque pour les couleurs, car d'un point de vue statistique l'ordre n'est pas fondamental. Quel que soit cet ordre en effet, il est toujours possible de comparer deux histogrammes différents.

Une autre solution plus adaptée consiste à construire un histogramme 3D, c'est-à-dire qu'un histogramme devient une fonction de  $\mathbb{R}^3$  dans  $\mathbb{R}$ . Pour chaque triplet  $(r, g, b)$ , on associe une valeur dénombrant les pixels dont la couleur  $c = (r, g, b)$ .

Il faut également noter qu'en général il est préférable de sous-échantillonner l'espace couleur. En utilisant les données brutes,  $255^3$  couleurs sont définies, ce qui donne des histogrammes trop lourds à gérer en mémoire. Une quantification des couleurs est donc nécessaire, et tant qu'à faire on peut également effectuer cette quantification de façon à privilégier les couleurs dominantes de l'objet à caractériser comme proposé par Lin *et al.* dans [Lin 01] ou par Heckbert [Heckbert 82] (approche par *median cut*).

De nombreux chercheurs ont développé en parallèle des algorithmes de suivi à partir des informations fournies par un histogramme couleur. L'approche de Yoo et Oh [Yoo 99] est un exemple de travaux parmi les premiers en la matière. Les auteurs rappellent que l'histogramme est par nature robuste aux rotations d'axe orthogonal au plan image et aux changements d'échelle de l'objet (si l'histogramme est normalisé). Ils utilisent également le principe de rétropropagation

(backprojection) de l'histogramme dans l'image.

Introduite par Swain et Ballard [Swain 91], la rétroprojection permet pour chaque pixel de l'image de définir un poids correspond à l'adéquation de la couleur du pixel par rapport à un histogramme couleur donné. Cette rétroprojection est obtenue en comparant l'histogramme modèle à un histogramme local centré sur le pixel jugé. Le ratio entre les histogrammes donne un critère de correspondance. Cette opération facilite la tâche de localisation de l'objet et peut, aussi selon la procédure de suivi, permettre d'épargner des calculs répétitifs.

Ce procédé repose en partie sur une approche locale. D'autres auteurs ont aussi remarqué que la localité joue un rôle important en terme de robustesse. Ainsi, il est par fois préférable de calculer de nombreux histogrammes locaux sur un seul objet plutôt que de calculer un seul histogramme global. On peut calculer ces histogrammes de manière quasi-exhaustive en définissant une grille sur l'objet, ou bien en se focalisant sur des points d'intérêt.

Notons enfin que pour réduire la taille des attributs, il est possible de ne retenir que des statistiques générales comme les moments mais en général cette compression d'information est trop destructrice et l'objet n'est alors plus suffisamment bien caractérisé.

## Densité de couleur

La densité de couleur est un attribut indexé dans la littérature en tant que méthode couleur, mais elle se situe déjà à la limite d'un attribut texture-couleur, car outre l'information couleur, de l'information spatiale y est également incorporée. A la base, une densité de couleur est un histogramme 3D, mais lors de la construction de cet histogramme le vote d'un pixel pour la couleur à laquelle il appartient est pondéré. Cette pondération s'effectue en fonction de la distance qui sépare le pixel votant du centre de l'objet caractérisé. C'est en cela que de l'information spatiale est injectée. Plus formellement, une densité  $q$  s'écrit :

$$q_u = C \sum_{i=1}^{N_{support}} k(\|\vec{p} - \vec{p}_i\|) \delta(I_t(\vec{p}_i) - u) \quad (1.1)$$

avec  $u$  une couleur donnée,  $k$  une fonction de type noyau,  $\vec{p}$  le pixel central de l'objet,  $\vec{p}_i$  un pixel appartenant au support de  $k$ ,  $N_{support}$  le nombre de pixel contenus dans le support, et  $C$  une constante de normalisation. La pondération liée à la distance entre le pixel central et les autres est effectuée par la fonction  $k$ . Cette approche par densité a été principalement utilisée par Comaniciu *et al.* [Comaniciu 00] (*cf.* 2.3.5).

### 1.3.2 Méthodes texture

Les méthodes texture sont traditionnellement appliquées sur des images en niveaux de gris afin de caractériser la répartition spatiale de ces niveaux de gris. La texture est une notion liée à la perception visuelle d'une surface, notamment sa rugosité et la périodicité de son rendu. Gagalowicz [Gagalowicz 83] propose la définition suivante : "Il faut considérer une texture comme une structure spatiale constituée par l'organisation de primitives (ou motifs de base) ayant chacune un aspect aléatoire. Une texture est alors une structure hiérarchique à deux niveaux, dans le cas général." L'information extraite par l'analyse des textures renseigne sur l'agencement spatial des pixels d'une région de l'image.

## Primitives et Textons

Pour étudier ces propriétés spatiales certains auteurs se sont d'abord intéressés aux agencements de primitives géométriques selon des règles de placement [Zucker 76, Fu 82]. Ces primitives sont aussi parfois appelées textons, terme désignant l'élément constitutif de base d'une texture. La figure 1.11 donne un exemple de texture construite à base de textons. Ces méthodes sont in-

**Figure 1.11 :** Texture rudimentaire construite avec deux textons.

téressantes pour comprendre les mécanismes de base servant à constituer une texture. Elles sont néanmoins considérées désuètes aujourd'hui, car elles ne peuvent caractériser que des textures grossières et régulières.

## Méthodes statistiques

De biens meilleurs résultats en termes de caractérisation ont été obtenus par les méthodes statistiques, deuxième génération de méthode succédant aux primitives. Gagalowicz [Gagalowicz 83] a montré à l'aide d'expériences psychovisuelles qu'une texture peut être caractérisée par son histogramme et son autocovariance sous réserve que ces attributs soient invariants par translation. Ces travaux mettent donc en avant la bonne capacité des approches statistiques à analyser les textures.

Galloway [Galloway 75] présente cinq attributs issus de matrices statistiques dites matrices des longueurs de plage. L'idée est assez simple : l'image est vue comme un ensemble de régions de niveaux de gris quasi-constant, et ces régions sont appelées plages. En traçant une ligne dans l'image, cette dernière est parcourue de part en part en traversant les plages. Il est alors possible de mesurer le nombre de pixels qu'il faut pour parcourir chaque plage. Les statistiques sur ces parcours sont ensuite rassemblées dans une matrice dont les cellules sont indexées par le niveau de gris d'une plage et la longueur de la plage. Ces matrices sont définies pour plusieurs orientations et positions de la ligne coupant l'image analysée.

Une autre technique assez similaire, mais plus simple et efficace, a été proposée par Haralick [Haralick 73]. Cette technique rassemble également des statistiques dans des matrices appelées matrices de cooccurrences. Ces statistiques s'intéressent directement à des couples de pixels reliés par une relation de voisinage donnée. En parcourant l'image, on dénombre chaque paire de niveaux de gris observée, d'où le terme de cooccurrence. Cette approche caractérise par définition

la topologie des niveaux de gris, et donc une texture. Le calcul de ces cooccurrences est donné par la formule :

$$M_{\vec{d}}(i, j) = \frac{\#(p \in D/I(\vec{p}) = n_1, I(\vec{p} + \vec{d}) = n_2)}{\#(D)} \quad (1.2)$$

où  $I$  est une image définie sur  $D$ ,  $\vec{d}$  un vecteur de translation,  $(n_1, n_2)$  un couple de niveaux de gris donné, et  $\#$  le cardinal d'un ensemble. Les vecteurs  $\vec{d}$  sont généralement choisis de sorte à observer les occurrences sur des pixels 8-voisins. Les matrices étant symétriques, quatre orientations suffisent pour examiner ce 8-voisinage :  $\vec{d} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$  et  $\theta \in \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$ . Le nombre de lignes ainsi que le nombre de colonnes de la matrices est égal au nombre de niveaux de gris pris en compte.

Les cooccurrences peuvent se calculer plus rapidement que les longueurs de plages, et n'ont pas besoin de définir des régions. Haralick définit 14 attributs sur ces matrices, qui ont été utilisés dans de très nombreux travaux. Les matrices de cooccurrences restent aujourd'hui une des méthodes les plus utilisées, car elles sont un bon compromis entre simplicité et efficacité. Leur principal défaut est une forte sensibilité aux variations d'illumination, car ces variations entraînent un changement des niveaux de gris et conséquemment un même objet éclairé différemment aura des matrices très différentes.

L'étude de l'autocorrélation d'une texture a également permis de mettre en évidence des attributs texturaux discriminants. Lecomte [Lecomte 99] a défini une méthode permettant de déterminer l'amplitude d'un changement d'échelle entre deux images d'une même texture. Brochard *et al.* [Brochard 01] ont pris la suite de ces travaux pour proposer des attributs invariants aux mouvements dans le plan image. Toutefois, l'autocorrélation reste assez peu fréquemment utilisée du fait d'un coup calculatoire plus élevée que les approches concurrentes.

Ojala *et al.* [Ojala 96] ont introduit l'analyse par schémas binaires locaux (*Local binary pattern*, LBP). L'idée est de prendre un pixel et ses huit voisins. Les voisins sont seuillés par rapport à la valeur du pixel central, ce qui donne des valeurs binaires. Chaque voisin est repéré par un code (puissance de 2). Les pixels seuillés à 1 sont multipliés par leur code, puis sont additionnés pour former au final une signature unique. la signature dépendant à la fois des niveaux de gris (seuillage) et de la topologie (codes), elle caractérise bel et bien la texture analysée. Les LBP sont une des rares méthodes à pouvoir résister à des changements d'illuminations d'amplitude raisonnable, si ces changements sont uniformes dans l'image. En revanche, le changement d'échelle de la texture pose problème, mais une approximation des LBP [Mäenpää 04] permet de contourner le problème.

### Méthodes fréquentielles, d'échelle et échelle-fréquence

L'analyse fréquentielle possède de très nombreuses applications en traitement du signal, mais elle peut être également vue comme une méthode de caractérisation des textures. En effet la répétition périodique d'un schéma textural est naturellement caractérisable par sa fréquence. C'est pourquoi il n'est pas du tout surprenant que l'analyse fréquentielle des images donne de bons résultats en matière de classification ou d'indexation d'images texturées. Campbell et Robson [Campbell 68] proposaient dès 1968 d'appliquer la transformée de Fourier, méthode d'analyse

spectrale par excellence, à des problématiques d'images.

L'analyse fréquentielle occupe toujours à l'heure actuelle une place prépondérante en analyse de texture. Josso *et al.* [Josso 05] proposent une approche visant à mieux exploiter les données texturales du domaine fréquentiel. Le module de la transformée de Fourier subit une analyse en composantes principales autour des basses fréquences. Cette analyse permet de déterminer s'il existe une orientation privilégiée dans le schéma textural. Dans le cas contraire la texture est anisotrope. Cette information permet une meilleure caractérisation des familles de textures.

De manière analogue aux fréquences, l'analyse des échelles permet de révéler certains processus périodiques propres aux textures. Dans cette veine, l'analyse fractale trouve toute sa place. En analyse d'images, Pentland [Pentland 84] fut un des premiers à utiliser les fractales qui, par définition, caractérisent un schéma se répétant à différentes échelles. Toutefois, en pratique, seul un attribut fractal est fréquemment utilisé dans la littérature : la dimension fractale. Cette dimension a montré un réel pouvoir discriminant concernant les textures. Il est en revanche difficile de ne se baser que sur un seul attribut pour analyser les textures, c'est pourquoi la dimension fractale est souvent associée à d'autres attributs texturaux de natures différentes. Par exemple, Quevedo *et al.* [Quevedo 02] unifient méthodes spectrales et fractales dans l'approche qu'ils nomment transformée de Fourier fractale.

Échelles et fréquences sont le plus souvent conjointement analysées par la famille des transformées en ondelettes et des filtres de Gabor, dont la définition est très proche de celle des ondelettes. Cette analyse conjointe permet une analyse plus précise des textures. Sur des images, les ondelettes se présentent comme des filtres 2D devant être convolués à l'image. Ces ondelettes dépendent de paramètres que l'on fait évoluer pour constituer un banc de filtres. L'ensemble des réponses à ces filtres forme une nouvelle représentation de la texture analysée. Bovik [Bovik 91] et Unser [Unser 95] furent parmi les premiers à montrer l'efficacité de ces méthodes sur des images de textures.

Wolf et Husson [Wolf 93] discutent de l'intérêt des ondelettes pour caractériser les surfaces texturées. Ils préconisent l'utilisation d'une base d'ondelettes orthogonales<sup>1</sup>. Ce type d'ondelettes permet un traitement de l'information non-redondant. À l'inverse Kyrki *et al.* [Kyrki 04] préfèrent les filtres de Gabor qui sont typiquement redondants. Les auteurs étudient l'invariance aux rotations selon l'axe orthogonal à l'image, aux translations dans l'image et aux changements d'échelle de l'objet. En permutant les lignes et les colonnes des images filtrées, une méthode de mise en correspondance invariante à tous ces mouvements est obtenue.

Nous classons également les attributs de Viola et Jones [Viola 04] dans la catégorie textures, car ils sont issus de convolution avec des fonctions s'apparentant aux ondelettes de Haar. L'extraction de ces attributs est rapide grâce à l'utilisation d'images intégrales. La méthode comprend également une phase d'apprentissage via une variante de l'algorithme Adaboost. Ces attributs sont reconnus pour leur excellente capacité à détecter des visages. La détection proposée par les auteurs se fait à l'aide de classificateurs en cascade. Ces attributs peuvent être employés sur des véhicules, mais une base de données d'image-exemple est nécessaire pour la phase d'ap-

---

<sup>1</sup>Orthogonalité au sens du produit scalaire entre fonction de carrés sommables.

prentissage.

La plupart des études comparatives désignent les ondelettes et les filtres de Gabor comme les meilleures méthodes de caractérisation de textures [Whelan 01]. Notons que les critères de jugements de ces études sont toujours des critères orientés classification, qui ne correspondent pas forcément entièrement aux critères pertinents pour un suivi d'objets. Parmi les défauts notables des approches échelle/fréquence, il est à souligner qu'une masse importante de données est obtenue en sortie des filtres. Il faut donc calculer des attributs tels que des moments statistiques sur ces images filtrées. De plus, la paramétrisation des filtres est peu évidente. Cheng *et al.* [Cheng 06] proposent d'utiliser l'algorithme Adaboost pour apprendre les meilleurs paramètres des filtres de Gabor pour une application de suivi de véhicules. Enfin, le temps de calcul nécessaire au bon fonctionnement des filtres est assez élevé en raison des nombreuses convolutions à effectuer.

### 1.3.3 Méthodes texture-couleur

La définition d'une texture ne repose pas sur le fait que l'image soit en niveaux de gris, il est donc naturel de vouloir appliquer les méthodes texture à des images en couleurs. L'intérêt d'un extracteur mixte texture-couleur va au delà de ce simple constat, car ces deux modes de caractérisation dépendent tous deux de la valeurs des pixels. Il existe donc une redondance d'information ou une forme de corrélation entre un extracteur texture et un extracteur couleur. Un extracteur mixte offre une caractérisation plus discriminante.

Une distinction est parfois faite entre techniques spatiales-couleur et techniques texture-couleur, mais la nuance est faible, surtout en prenant la définition d'une texture au sens large. En extrapolant la définition d'une texture, on peut dire par exemple que le fait qu'une région soit entièrement rouge et dépourvue de texture est aussi une forme de caractérisation texturale. Autrement dit, l'absence de texture est une caractéristique texture.

Il existe plusieurs manières d'extraire des caractéristiques mixtes. La plus simple consiste à appliquer les méthodes textures évoquées dans la sous-section précédente directement aux trois plans couleur. Par exemple Jung dans [Jung 06] applique, entre autres, des ondelettes aux trois plans couleur, de même avec les cooccurrences pour Palm [Palm 03] ou les LBP pour Pietikäinen [Pietikäinen 02].

La seconde méthode consiste à considérer la couleur comme une grandeur continue, par analogie aux niveaux de gris qui sont eux issus d'une grandeur réellement continue : la luminosité. Chang *et al.* [Chang 99] calculent les cooccurrences sur les occurrences de voisinage de couleurs à la place de voisinage de niveaux de gris. Notons que certains auteurs parlent de corrélogramme couleur ou encore d'histogramme de cooccurrences couleur, dont les définitions sont quasi-identiques à celle des matrices de cooccurrences couleurs. D'un point de vue efficacité, il est assez difficile de dire lequel des deux schémas de passage à la couleur est le meilleur, mais le premier schéma se contente de générer plus d'attributs sans prendre en compte la sémantique de la couleur.

Concernant les matrices de cooccurrences couleur, nous avons proposé une amélioration du calcul de ces matrices destinée à rendre les matrices invariantes à des changements d'illuminations

d'amplitude raisonnable [Klein 07a, Klein 07b]. Notre méthode s'applique si quelques couleurs de référence d'un objet à suivre sont connues *a priori*. Seules ces couleurs sont alors utilisées pour constituer les cellules des matrices. Pour chaque couleur observée dans notre objet, la distance la plus faible par rapport à l'ensemble des couleurs de référence est retenue ainsi que l'étiquette de la couleur  $c$  la plus proche. Ensuite, une fonction de type noyau est appliquée sur cette distance produisant un scalaire  $w$ . Le vote des couples de couleur en relation de voisinage observés lors du dénombrement est pondéré par  $w$ . Autrement dit, une couleur observée est associée à la couleur de référence  $c$  avec un score  $w$ , ce qui permet de quantifier les couleurs de manière plus subtile. Des couleurs proches correspondent à la même cellule de la matrice et la taille des matrices est réduite.

Afin de ne pas entraver la lisibilité de l'état de l'art, les justifications et expérimentations concernant cette contribution sont renvoyées en annexe A.

D'autre part, il existe des méthodes un peu plus spécifiques pour combiner l'information couleur et l'information spatiale, qui viennent en marge de ces deux schémas. Parmi ces méthodes on peut citer, les vecteurs de cohérences couleur [Pass 96] où la cohérence est jugée en fonction de la taille d'une région appartenant à une classe de couleur, les histogrammes chromatiques spatiaux [Cinque 99], qui partent d'un histogramme couleur, couplé avec la distance des pixels par rapport au barycentre de la région d'intérêt. Cette définition est relativement similaires à celles des densités de couleur [Comaniciu 00] qui, comme nous l'évoquions précédemment, peuvent aussi être considérées comme une méthode texture-couleur. Enfin les travaux de Paschos [Paschos 00] sont à noter. Il passe dans un premier temps dans l'espace CIE XYZ, puis analyse d'une part la répartition dans le plan  $(x, y)$  des pixels de l'image, et d'autre part le nombre de pixels de l'image correspondant à un  $(x, y)$  donné.

Récemment, une méthode spécifiquement applicable aux images couleur a été développée : la transformée de Fourier quaternionique [Moxey 03]. Un quaternion est en quelque sorte l'extension d'un nombre complexe. Il possède une composante réelle et trois composantes complexes pures. Il est alors possible de représenter une image couleur en associant chaque plan couleur avec une composante complexe. La transformée quaternionique permet alors une représentation plus variée du signal image que la transformée de Fourier classique. Cette approche reste néanmoins encore peu utilisée car plus difficilement interprétable que la transformée de Fourier ou les ondelettes, mais des résultats encourageant ont été obtenus [Denis 07].

Il est difficile de donner un classement concernant l'efficacité de l'ensemble de ces méthodes. Elles offrent un compromis entre temps de calcul et précision différent. Tout dépend donc de l'application à laquelle la caractérisation est destinée. Notons tout de même qu'en suivi d'objets, les densités de couleur sont largement plus utilisées que la moyenne.

#### 1.3.4 Bilan sur l'information texture-couleur

De notre point de vue, l'intérêt d'un extracteur mixte texture-couleur est net. Il se résume par l'ensemble de ces différents points discutés dans cette partie :

- caractérisation des valeurs prises par les pixels de l'objet,
- prise en compte de la répartition spatiale de ces mêmes valeurs,

- minimisation de la redondance entre attributs de texture et attributs de couleur,
- concision des données extraites,
- augmentation du pouvoir discriminant des extracteurs.

Le dernier point de nos conclusions est contesté par une étude de Mäenpää et Pietikäinen [Mäenpää 04]. Les auteurs montrent, via une problématique de type classification, que certes un extracteur mixte offre de meilleures performances qu'un extracteur texture ou un extracteur couleur isolé, mais qu'une combinaison d'un extracteur texture et d'un extracteur couleur peut surpasser un extracteur mixte. Des combinaisons assez simples (max, min et moyenne) sont testées. Cette approche est tout à fait valable, même si la gain en précision se fait au détriment du temps de calcul. Nous estimons cependant que la différence entre un extracteur mixte et une combinaison d'extracteurs n'est pas fondamentale, car sémantiquement parlant la volonté d'extraire des caractéristiques à la fois texture et à la fois couleur est la même, seule la méthode diffère.

Par rapports à nos données vidéos, les méthodes de types ondelettes et filtres de Gabor semblent légèrement trop coûteuses en temps de calcul. Les LBP montrent un certain intérêt même si le choix des paramètres (nombre de schémas, position du schéma et taille du schéma) n'est pas évident à déterminer. L'efficacité des densités de couleur a été éprouvée et est reconnue par la communauté en suivi d'objets. Les matrices de cooccurrences pondérées ont montré des résultats encourageants (*cf.* annexe A) et peuvent donc être elles aussi sélectionnées.

Un problème majeur pour ces extracteurs est la présence de changements d'illumination dans la scène, car ces changements modifient les valeurs des pixels, or c'est sur ces valeurs que sont construites les caractéristiques texture-couleur. La partie 2.7.4 donne plus de détails sur les enjeux liés aux changements d'illumination. Seuls les LBP et les matrices de cooccurrences couleur pondérées prennent en compte la variation possible de luminosité dans la conception des extracteurs.

## 1.4 Extraction de caractéristiques de forme

La forme est un attribut caractéristique d'un objet très différent de la couleur ou de la texture. Elle permet soit de caractériser un objet en particulier, soit une classe d'objets. Dans les séquences routières, la forme permet de distinguer une voiture, d'un camion et du fond de la scène. En revanche pour distinguer une voiture  $A$  d'une voiture  $B$ , la forme est généralement insuffisante. La forme est de plus indépendante de l'intensité lumineuse. Ses propriétés sont stables au cours du temps (sauf cas des objets déformables non traités dans ce mémoire).

L'information de forme repose, d'un point de vue des données, beaucoup plus sur les variations spatiales des valeurs des pixels et la topologie de ces variations plutôt que sur les valeurs elles-mêmes des pixels. Elle est donc relativement disjointe de l'information texture-couleur.

Certaines méthodes de caractérisation de forme essaient de résumer la forme de l'objet à l'aide primitives graphiques : contours, squelettes ou graphes. D'autres traitent l'objet dans sa globalité et extraient des mesures notamment sur la géométrie de l'objet. Dans cette partie, nous présentons les principales façons dont est traitée l'information de forme. Néanmoins dans les catégories évoquées nous ne rentrerons pas dans le détail des méthodes ni de leurs évolutions.

Cette partie a pour but de mettre en avant des outils simples et suffisamment efficaces pour pouvoir être exploités dans un algorithme de suivi. Une étude proposée par Zhang *et al.* [Zhang 04] offre des compléments d'information sur le contenu de cette partie.

### 1.4.1 Détection de contours

La base morphologique d'un objet peut être caractérisée par son contour. Il existe deux grandes catégories de méthodes de détection de contours : les méthodes différentielles et les méthodes utilisant des modèles déformables de contours.

– **Méthodes différentielles** Les méthodes différentielles<sup>2</sup> reposent sur le principe qu'un contour est présent dans l'image si deux pixels voisins ont une différence de niveau de gris (ou de couleur) marquée. A ce titre les méthodes dérivatives ont été les premières à être utilisées. La dérivée du premier ordre est approximée par le filtre de Sobel ou de Prewitt [Prewitt 70], et celle du second ordre par le filtre de Laplace.

Canny [Canny 86] lui aussi est parti de la même idée, mais pose trois contraintes d'optimalité. Ces trois critères sont :

- critère de bonne détection : faible taux d'erreur dans la signalisation des contours,
- critère de bonne localisation : minimisation des distances entre les contours détectés et les contours réels,
- critère d'unicité de la réponse : une seule réponse par contour et pas de faux positifs.

En posant ces contraintes, il obtient une équation différentielle résoluble. Le filtre de Canny reste un des plus employés en détection de contours car il est un bon compromis entre rapidité et efficacité. Deriche [Deriche 90] a proposé la version à réponse impulsionnelle infinie de ce filtre, permettant une synthétisation récursive efficace du filtre.

Miché et Debrie [Miche 95] ont néanmoins mis au point un algorithme de complexité équivalente mais donnant dans des images non-compressées de meilleures détections. Cette approche est appelée déclivité. Elle est également basée sur des différences entre valeurs de pixels. Le déclivité se définit comme un ensemble de pixels consécutifs dont les intensités sont une fonction monotone de leurs positions. Cette définition permet un traitement plus global de la notion de contour.

De très nombreuses méthodes de détections de contours sont disponibles dans la littérature. Dans le cadre du suivi de véhicule, nous ne présentons que les approches simples et rapides classiquement utilisées, car notre objectif n'est pas de détecter parfaitement le contour du véhicule. Une approximation de dernier est suffisante pour alimenter la procédure de suivi.

– **Méthodes utilisant des modèles déformables** Dans ces méthodes, le contour correspond à une courbe fermée  $\Gamma$ . Un modèle déformable est soit paramétrique (contours actifs [Kass 88]) soit géométrique (contours actifs géodésiques [Caselles 93, Malladi 95]). Si le modèle est paramétrique la courbe  $\Gamma$  est obtenue en minimisant une fonctionnelle d'énergie  $E$  :

$$E(\Gamma) = \int_0^1 E_{int} + E_{ext} ds \quad (1.3)$$

<sup>2</sup>Le terme différentiel n'est pas un terme courant dans la littérature, nous désignons par ce terme les méthodes qui détectent les contours en s'appuyant sur des différences de niveaux de pixels.

La fonctionnelle se répartit en deux termes. Le premier  $E_{int}$  est un terme de régularisation intérieur qui comprend en général une contrainte sur la dérivée du premier et du second ordre du contour. Le second terme  $E_{ext}$  peut être calculé de manières différentes mais il représente la cohérence de l'apparence de l'objet détourné. On parle de forces externes, car ce terme dépend des données image.

Si le modèle est géométrique, la courbe correspond à l'intersection du plan d'altitude zéro et d'une surface (méthode des courbes de niveaux, ou *level-sets* en anglais). On parle de méthode d'évolution de la courbe.

Les méthodes de modèles déformables sont itératives afin de faire converger le modèle déformable vers le contour réel. C'est le principal défaut de ces méthodes, car il les rend lentes et dépendantes du point de départ des itérations. L'initialisation du processus joue donc un rôle crucial dans l'obtention du bon contour. En revanche ces méthodes donnent en général une meilleure approximation du contour que les méthodes différentielles. Notons que les contours actifs géodésiques peuvent gérer des changements topologiques du contour (séparation en deux), alors qu'un modèle paramétrique ne le peut pas.

Dans le cadre du suivi d'objets, elles sont très rarement utilisées comme méthodes d'extraction de contour, en revanche elles sont relativement souvent utilisées en tant que méthodes mêmes de suivi. En effet les contours actifs peuvent être vus à la fois comme de l'extraction de contour et à la fois comme du suivi de contour (*cf.* section 2.3.3).

### 1.4.2 Squelettisation

Les méthodes à base de squelettisation constituent la deuxième grande famille de méthodes pour l'extraction de primitives graphiques avec celles à base de détection de contours. Plusieurs états de l'art y sont en majeure partie consacrés [Tombre 00, Delalandre 05]. La squelettisation est la procédure qui partant de l'image de l'objet ou de l'image en entier, construit le squelette de ce dernier. Les méthodes classiquement utilisées sont à base de filtre de morphologie mathématique (amincissements successifs) ou à base de distance au contour le plus proche. Des travaux récents permettent notamment de réduire la variabilité des squelettes obtenus en résistant mieux au bruit [Morrison 06] ou bien en éliminant des barbules<sup>3</sup> [Couprie 07].

Dans [Siddiqi 98], une méthode de mise en correspondance est proposée entre squelettes et permet donc d'utiliser le squelette pour caractériser une forme. Néanmoins, il est assez rare d'utiliser le squelette à cet effet sauf dans le cas de suivi de piétons ou d'objet articulés. Il est de plus difficile de mettre en correspondance des squelettes issus de différents algorithmes de squelettisation, car le squelette peut être représenté sous différentes formes, ces formes étant un attribut des méthodes de squelettisation.

Enfin la complexité combinatoire des squelettisations est souvent trop élevée pour les problématiques de vision par ordinateur.

---

<sup>3</sup>Un barbule est une branche du squelette indésirable, souvent liée au bruit de l'image squelettisée.

### 1.4.3 Méthodes structurelles

Les méthodes structurelles ont pour but de résumer la forme d'un objet en une succession d'éléments de formes prédéterminées, appelées primitives  $\{s_1, \dots, s_{N_{\text{primitives}}}\}$ . Ces primitives sont alors regroupées dans une chaîne  $S = \{s_{i_1}, s_{i_2}, \dots, s_{i_n}\}$ , si la chaîne comporte  $n$  primitives. Généralement, les primitives sont des segments, des arcs, ou des splines. Mises bout à bout, elles forment en général une approximation du contour externe de l'objet, bien que les contours internes peuvent être également analysés. La primitive s'accompagne d'attributs tels que la longueur, l'orientation ou la courbure.

Freeman [Freeman 61] a introduit la représentation par chaîne codée. Il utilise des vecteurs unitaires qui peuvent être définis dans les directions du 4-voisinage ou du 8-voisinage. En partant d'un point du contour réel, et en choisissant à chaque fois le vecteur unitaire qui suit le mieux le contour, on obtient une approximation grossière du contour. Il est alors préférable de construire un histogramme à partir du code obtenu, comme proposé par Iivarinen [Iivarinen 96].

Dans [Mehrotra 95] Mehrotra et Gary détectent des points d'intérêt sur le contour de l'objet, grâce à une approximation polygonale. La forme peut être alors caractérisée par les coordonnées normalisées des points détectés. Les auteurs y adjoignent des paramètres globaux d'échelle, de translation et d'orientation, afin de gagner en indépendance par rapport aux mouvements de l'objet.

L'approximation du contour est souvent meilleure si les primitives sont des courbes lisses au lieu de segments. Berreti *et al.* [Berretti 00] étendent une méthode polygonale [Gorskey 90] à cet effet. Les primitives sont obtenues en lissant le contour par une gaussienne et en conservant l'intersection avec le plan d'altitude zéro. Une information de courbure est intégrée aux paramètres caractérisant les primitives. La figure 1.12 montre l'exemple d'une forme de cheval compressée et caractérisée par cette méthode. Il est en outre possible d'utiliser l'analyse syntaxique [Fu 74]

**Figure 1.12 :** Contour d'objet résumé par un ensemble de primitives de types courbes.

pour caractériser un ensemble de primitives. S'inspirant des sciences du langage, les primitives sont interprétées comme des mots de code, et une suite de mots forme comme une phrase. Les formes sont appareillables par comparaison de chaînes de caractères. Il est en général préférable d'avoir à disposition une grammaire, c'est-à-dire un ensemble de règles régissant la construction de phrases. L'approche syntaxique nécessite donc un *a priori* élevé.

Nous ne développons pas plus en détails les méthodes structurelles, bien qu'elles constituent un pan important des méthodes de caractérisation de forme. Elles sont en effet inadaptées à

notre problématique pour de nombreuses raisons. Elles s'appliquent à des formes simples déjà disponibles sous un format d'images binaires, ce qui présupposerait l'application de méthodes de détection de contour très performantes. Elles sont souvent très lentes, et n'offrent que peu de résistance au bruit. Étant donné l'état dégradé de certaines images issues de séquences routières, ces méthodes doivent être écartées.

#### 1.4.4 Points d'intérêt et graphes

Cette sous-section rejoint la partie 1.2.6. Nous donnons ici des détails supplémentaires sur l'obtention d'un nuage de point ou d'un graphe à partir de l'image d'un objet.

– **Nuage de points** Les nuages de points sont dans la très grande majorité des cas obtenus par des techniques de détection de points d'intérêt. Un point d'intérêt peut se définir comme un pixel présentant une position stratégique par rapport à la forme de l'objet, par exemple un coin pour un objet rectangulaire. L'ensemble des points d'intérêt permet ainsi de décrire efficacement une forme.

Les points d'intérêt sont situés sur des zones de fort contraste, c'est-à-dire que l'intensité des pixels voisins varie fortement. Cette propriété est un atout pour la détection des points d'intérêt.

La méthode la plus courante de détection de point d'intérêt est la méthode de Harris [Harris 88]. Harris reprend l'idée de Moravec [Moravec 80] de fenêtrer l'image et de mesurer le changement  $E_{i,j}$  engendré par un petit déplacement  $(i, j)$ . Un développement de Taylor du second ordre est utilisé pour estimer ce changement. Le fenêtrage utilisé  $win()$  est une gaussienne.  $E_{i,j}$  est alors approximée de la manière suivante :

$$E_{i,j} = (i, j) M \begin{pmatrix} i \\ j \end{pmatrix} \quad (1.4)$$

$$M = \begin{pmatrix} \frac{\partial^2 I}{\partial i^2} * win & \frac{\partial^2 I}{\partial i \partial j} * win \\ \frac{\partial^2 I}{\partial i \partial j} * win & \frac{\partial^2 I}{\partial j^2} * win \end{pmatrix} \quad (1.5)$$

avec  $*$  l'opérateur de convolution. En observant les valeurs propres de  $M$ , on peut établir une première classification des points en deux classes : point normal et point d'intérêt. Si les deux valeurs sont faibles, le contraste autour du point est faible. Pour les points d'intérêt une deuxième classification est possible : point de contour et coin. Si l'une des deux valeurs est élevée, alors il existe un fort contraste dans une seule direction, il s'agit donc d'un point de contour, sinon le contraste est important dans les deux directions des vecteurs propres, il s'agit d'un coin.

Le détecteur de Harris est encore à l'heure actuelle un des plus utilisés. Simple à mettre en œuvre, rapide, il offre de plus un bon taux de détection des points d'intérêt. La figure 1.13 montre un exemple de détecteur de Harris appliqué à une image de véhicule. La fiabilité de la méthode de Harris peut être encore améliorée par des techniques multi-échelles [Lowe 99, Lowe 04]. En effet le contraste autour d'un point ressort de manières différentes selon l'échelle à laquelle il est examiné.

Une autre méthode classiquement utilisée est celle dite du laplacien de Gauss (Laplacian of Gaussian, LoG) ou de différence de gaussiennes (Difference of Gaussian, DoG) qui est une

**Figure 1.13 :** Détection de points d'intérêt de Harris multi-échelles : points détectés en jaune.

approximation du LoG. Cette méthode utilise une convolution par une gaussienne. Cela revient en réalité à lisser l'image et permet de lutter contre l'apparition de points d'intérêt fictifs, comme ceux dus, par exemple, au bruit de quantification de l'image. Il est ensuite naturel de penser à examiner le laplacien du résultat de la convolution précédente. Les zéros du laplacien caractérisent dans une certaine mesure les points d'inflexion de l'intensité des pixels, et donc la présence de coins ou de contours. Comme la dérivation d'une gaussienne possède certaines propriétés, le problème se simplifie grandement et le LoG revient à une convolution avec un filtre prédéfini.

Pour les DoG, la méthode est adaptée au multi-échelle. Comme leur nom l'indique, les DoG calculent des différences entre des gaussiennes avant de convoluer cette différence avec l'image. Les gaussiennes différenciées possèdent des écart-types différents correspondant à différentes échelles.

Les DoG sont un des piliers de la méthode proposée par Lowe [Lowe 99, Lowe 04]. Cette méthode, appelée communément SIFT (Scale Invariant Feature Transform) se décompose en deux étapes : une étape de détection de points d'intérêt et une étape de caractérisation des points détectés. L'étape de détection repose sur une pyramide de DoG calculée sur plusieurs octaves et à pas d'échelle constant. Les extrema locaux à la structure tridimensionnelle de la pyramide constituent les points d'intérêt. Plusieurs améliorations sont apportées afin d'éliminer des faux positifs. Une méthode itérative est employée pour faire converger des points d'intérêt qui ne se situeraient pas exactement au niveau de l'extremum local (ce phénomène est assez fréquent avec des approches multi-échelles). Les points de contraste trop faibles sont purement éliminés. Enfin de manière analogue à Harris, la matrice hessienne permet de différencier les points de contour des coins.

La seconde partie de SIFT consiste à calculer de très nombreux attributs sur les points détectés. Ces attributs dépendent de l'intensité de l'orientation des gradients locaux à chaque point. Bien que SIFT soit considéré comme une des meilleures méthodes de détection/caractérisation de points d'intérêt actuelle [Moreno 07], elle est difficilement applicable à nos ambitions, car la quantité d'attributs fournie en sortie de l'algorithme est trop importante. Qui plus est, la partie purement détection de SIFT s'est révélée peu adaptée à nos images. La répétabilité<sup>4</sup> de la détection des points est sur nos images inférieure à celle de la méthode d'Harris. Il semblerait donc que SIFT doive être employée dans son intégralité pour livrer tout son potentiel.

Une fois les points d'intérêt obtenus, les couples de coordonnées de ces différents points

---

<sup>4</sup>La répétabilité est la capacité d'un point d'intérêt à être localisé de manière stable temporellement et topologiquement.

constituent déjà une caractérisation de forme. Néanmoins, cette caractérisation n'est pas bonne car beaucoup trop dépendante des mouvements de l'objet. Il convient en général de normaliser les coordonnées, calculer des orientations ou faire des statistiques globales sur le nuage de points.

– **Graphes** Les graphes constituent un mode de représentation et d'analyse de l'information naturelle et doté d'un formalisme unifié particulièrement adapté notamment à la caractérisation d'objets. En vision par ordinateur, les nœuds sont souvent constitués par des points d'intérêt, plus rarement par les centroïdes des régions qui le compose. En ce qui concerne la construction des arcs, plusieurs techniques peuvent être utilisées. Nous reprenons la présentation de ces méthodes proposée par Barbu [Barbu 07] :

- Construction par adjacence de régions (à supposer donc que de telles régions existent). Si deux régions ont une frontière commune, les nœuds correspondants sont reliés par un arc.
- Construction par triangulation de Delaunay. Elle repose sur la méthode de pavage du plan de Voronoï.
- Construction par  $k$  plus proches voisins. Pour un nœud donné, ses  $k$  plus proches voisins sont mis en relation avec lui.
- Construction à partir d'un seuil de distance. Une distance est calculée entre chaque paire de nœuds, ces nœuds sont mis en relation si la distance est inférieure à un seuil donné.
- Construction à l'aide de l'arbre couvrant de longueur minimale du graphe complet<sup>5</sup>. En partant du graphe complet on attribue un poids à chaque arc. Ce poids peut par exemple être un fonction de la distance entre les deux nœuds mis en relation. Il faut ensuite supprimer le maximum d'arcs du graphe complet, tout en conservant au moins un arc par nœud. L'arbre couvrant minimal est alors le sous-graphe qui est tel que la somme totale des poids des arcs restants soit minimale.

Notons qu'en vision, les méthodes les plus simples sont à privilégier car moins coûteuses en temps de calcul (adjacence, plus proches voisins et seuil de distance). De même que pour les nuages de points, des attributs supplémentaires peuvent être ajoutés au nœud et au arc du graphe.

Zass et Shashua [Zass 08] utilisent la notion d'hypergraphe pour caractériser un objet. Dans un hypergraphe, un arc peut relier plus de deux nœuds. Cela permet entre autres de hiérarchiser le graphe et de rendre la mise en correspondance plus robuste.

Le principal défaut des graphes est qu'ils sont difficiles à comparer. Or comme nous le verrons dans le chapitre suivant la mise en correspondance, modèle/observation est primordiale en suivi d'objets. Les graphes sont comparés grâce à des similarités, métriques ou pseudo-métriques, qui imposent le recours à des méthodes extrêmement coûteuses, telles que, par exemple, la recherche d'isomorphismes de graphes ou de sous-graphes (problème NP complet).

#### 1.4.5 Attributs statistiques

Hu [Hu 62] a introduit des attributs invariants, appelés invariants des moments géométriques. Ils sont construits directement sur les moments de l'image de l'objet à l'aide de combinaisons non-linéaires. C'est par rapport aux mouvements rotationnels, translationnels et d'échelle,

<sup>5</sup>Le graphe complet est celui qui relie tous les nœuds entre eux sans distinction.

que les attributs sont invariants. Ces statistiques sont d'autant plus performantes que des moments de rangs élevés sont calculés, au détriment du temps de calcul. Zhang [Zhang 02] a montré que ces invariants étaient surtout performants dans le cadre de formes peu complexes.

Taubin et Cooper [Taubin 92] ont mis en place d'autres invariants dits invariants des moments algébriques. Des matrices sont prédéfinies et leurs éléments dépendent des facteurs d'échelle des moments centrés de l'image-objet. Les attributs sont alors tirés des valeurs propres de ces matrices. L'invariance est réalisée pour des transformations affines. Dans [Scassellati 94] des expériences montrent que ces invariants ont des performances relativement instables, et qu'il est donc difficile de porter un jugement global les concernant.

Le principal problème imputable aux invariants algébriques provient de leur forte dépendance mutuelle. Par ailleurs, les moments d'ordres supérieurs donnent lieu à des valeurs de trop grandes amplitudes. Pour pallier ces inconvénients, les moments orthogonaux leur sont souvent préférés. Les moments orthogonaux comprennent entre autres, les moments de Legendre et de Zernike. L'idée est d'utiliser une fonction de type noyau à l'intérieur de la formule de calcul d'un moment. Teague [Teague 80] propose alors comme fonction, les polynômes de Legendre ou de Zernike. Les moments de Legendre d'ordre  $m$  en ligne et  $n$  en colonne sont donnés par la formule suivante :

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \sum_i \sum_j P_m(i) P_n(j) I_t(i, j) \quad (1.6)$$

$$\forall n, P_n(i) = \frac{1}{2^n n!} \frac{d^n}{di^n} (i^2 - 1)^n \quad (1.7)$$

Ceux de Zernike s'obtiennent par :

$$A_{mn} = \frac{n+1}{\pi} \sum_i \sum_j V_{nm}^*(i, j) I_t(i, j) \quad (1.8)$$

$$V_{mn}(i, j) = V_{nm}(\rho \cos \theta, \rho \sin \theta) = R_{mn}(\rho) \exp(\mathbf{i}m\theta) \quad (1.9)$$

$$R_{mn}(\rho) = \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s} \quad (1.10)$$

avec  $(\rho, \theta)$  les coordonnées polaires correspondantes à  $(i, j)$ , et  $\mathbf{i}$  le complexe tel que  $\mathbf{i}^2 = -1$ . Étant donné que les polynômes utilisés dans les deux cas forment une base orthogonale de l'espace des fonctions, les moments sont qualifiés d'orthogonaux. Ces attributs ont les mêmes propriétés d'invariance aux mouvements que les moments algébriques.

Les moments de Zernike sont souvent préférés à ceux de Legendre car ils résistent mieux au bruit. On remarque toutefois que les moments de Zernike ont une définition proche des opérateurs d'analyse fréquentielle. L'opérateur de Fourier-Mellin entre lui aussi dans cette catégorie et peut offrir de meilleures performances que les moments de Zernike (*cf.* Adam *et al.* [Adam 01]).

Ces méthodes se rapprochent d'un extracteur de texture. Certains auteurs considèrent d'ailleurs d'emblée les méthodes fréquentielles ou échelle/fréquence comme des méthodes de caractérisation de forme. Nous ne suivons pas cette philosophie dans ce mémoire car, comme vu en 1.3.2, nous considérons que la fréquence caractérise la répétition d'un motif, ce qui par

définition correspond à une texture. Nous reconnaissons toutefois les liens entre analyses de formes et de textures.

#### 1.4.6 Cartes de symétrie

Les cartes de symétrie sont principalement utilisées pour caractériser la forme des véhicules [Kuehnle 91, Bensrhair 02, Sun 06]. La symétrie peut être considérée comme une caractéristique morphologique globale. Pour bien caractériser la forme d'un objet il faut des attributs calculés localement et globalement, ce qui est réalisé par les cartes de symétrie. L'algorithme calculant ces cartes est simple :

- 1 : un axe de symétrie (vertical par exemple) est fixé sur une des colonnes de l'image analysée.
- 2 : chaque paire de pixels équidistants à l'axe est examinée.
- 3 : si la différence colorimétrique ou de niveau de gris selon la nature de l'image est suffisamment faible entre les deux pixels examinés, la case correspondante de la carte est alors incrémentée. Les coordonnées de la case sont définie par la position de l'axe de symétrie et la distance entre la paire de pixels examinée et l'axe.

Il suffit ensuite d'itérer pour chaque position de l'axe et chaque distance. En ce sens, les symétries globales et locales sont prises en compte. La figure 1.14 donne un exemple de calcul d'une telle carte. Étant données les bornes de l'image la carte ne peut être calculée en n'importe quelles

**Figure 1.14** : Image d'un véhicule de sa carte de symétrie correspondante. La présence de symétrie se traduit par des raies noires au niveau de la carte.

coordonnées et est donc triangulaire. Il est possible de calculer deux cartes : une pour les symétries verticales et une seconde pour les symétries horizontales. Sur les images routières, comme la route est elle aussi relativement symétrique, certains auteurs préfèrent calculer les cartes sur des images de contours. Notons enfin que le défaut principal des cartes est le temps de calcul relativement élevé dû aux nombreuses itérations nécessaires.

#### 1.4.7 Autres méthodes de caractérisation de forme

Il existe quelques méthodes en marge de celles précédemment évoquées, qui restent fréquemment évoqués dans les états de l'art consacrés à l'extraction d'attributs de forme :

- Les méthodes de grille [Lu 99] : une grille est superposée à la forme, les cases de la grille englobant un contour de l'objet sont mises à 1, les autres à 0. On obtient ainsi une "bitmap", qui caractérise la forme.
- Les matrices de forme [Goshtasby 85] : au lieu d'échantillonner par une grille, l'échantillonnage est réalisé en coordonnées polaires, ce qui donne des arcs de cercles concen-

triques. Les lignes et les colonnes de la matrice correspondent à un rayon  $r$  et à un angle  $\theta$  donné. La valeur binaire de la matrice dépend du fait que le point  $(r, \theta)$  soit à l'intérieur de la région de l'image recouvrant l'objet.

- Les méthodes à base d'enveloppes convexes [Sonka 93] (*convex hull*) : les différences entre l'enveloppe convexe et l'enveloppe réelle permettent de mettre en évidence des régions où la courbure varie plus intensément. Une description arborescente de la forme est ainsi obtenue.
- Les descripteurs géométriques : aire, périmètre, excentricité (longueur de l'axe principal/longueur de l'axe mineur), circularité (périmètre<sup>2</sup>/aire), orientation de l'axe majeur, etc.
- Les signatures de forme [Davies 97] : il s'agit de fonctions réelles d'une seule variable calculées sur le contour de l'objet (par exemple, distance d'un point de contour au centroïde).

Ces méthodes sont jugées trop simplistes pour réellement caractériser une forme surtout dans un contexte aussi dynamique et complexe que les séquences routières. Elles sont de plus souvent dépendantes d'une autres techniques comme la détection de contour.

#### 1.4.8 Bilan sur l'information de forme

Concernant les méthodes d'extraction de forme, nous remarquons que celles produisant des caractéristiques suffisamment discriminantes sont toutes coûteuses en temps de calcul. Bien que le temps de calcul soit un critère relativement important pour notre application, il apparaît difficile pour l'information de forme de se passer d'un certain effort calculatoire. Cet effort garantit un minimum de robustesse indispensable à un algorithme de suivi.

L'information de forme reste précieuse et a été exploitée par de nombreux auteurs en suivi. Trois méthodes reviennent régulièrement dans les approches proposées : les contours, les points d'intérêt et les cartes de symétrie. Les contours sont assez difficiles à exploiter directement sur nos séquences, car ils sont en général assez flous. De plus, le contour des véhicules est relativement simpliste et donc sujet à confusion avec d'autres objets. Les points d'intérêt ont montré une persistance assez faible dans le temps de leurs performances sur nos séquences. Sans doute nos séquences sont elles trop dynamiques pour assurer une répétabilité des détections de points d'intérêt. En revanche les cartes de symétrie ont prouvé leur efficacité sur des séquences similaires aux nôtres [Kuehnle 91, Bensrhair 02, Sun 06], elles ont donc un avantage par rapport aux autres.

### 1.5 Extraction de caractéristiques de mouvement

L'information de type mouvement analyse l'évolution de la valeur des pixels dans le temps. Un mouvement est repéré si des pixels possèdent d'une image à l'autre des valeurs semblables mais des positions différentes. Pour notre application, et plus généralement en suivi, l'information de mouvement est celle qui caractérise le moins un objet. Il est possible de détecter les objets en mouvement dans une image de la scène, mais rarement de déterminer parmi ces détections quel est l'objet recherché. Une véritable caractérisation n'est possible que sous l'hypothèse que le mouvement de l'objet suivi est le seul de la scène à posséder certaines propriétés (direction, sens, orientation ou vitesse par exemple).

Toutefois l'information de mouvement procure un indice important sur la localisation d'un objet. Chez les organismes vivants, le mouvement est un élément essentiel de la perception visuelle. Certains animaux ne voient que par le mouvement, et chez les humains un mouvement attire inconsciemment l'attention visuelle sur l'entité en déplacement. Dans notre contexte, il semble d'emblée qu'on ne puisse utiliser les caractéristiques de mouvement que conjointement à des caractéristiques de nature différente. Cependant il est aussi vrai de dire que l'information de mouvement procure des renseignements que les autres types d'information ne peuvent fournir.

Le mouvement dans les séquences d'images a été naturellement étudié à de nombreuses reprises. Les méthodes de détection de mouvement cherchent à trouver les entités dont le mouvement se détache du reste de la scène. S'il faut de plus que chaque région en mouvement soit labellisée, on parle de segmentation de mouvement. Par ailleurs, on trouve aussi les méthodes dites d'estimation de mouvement qui ont pour but d'estimer les paramètres de mouvement d'un objet correspondant au mieux aux observations.

De même que pour les textures-couleurs ou la forme, nous ne présentons pas ici un état de l'art des méthodes de mouvement, mais simplement les méthodes les plus marquantes. Nous reprenons une classification des méthodes de détection de mouvement en quatre catégories proposées par Bugeau [Bugeau 07a].

### 1.5.1 Méthodes basées sur la différence inter-images

#### Détection de mouvement

La méthode la plus simple visant à détecter des objets dans les séquences vidéo est la différence entre images successives. Elle fut exploitée dès 1979 par Jain et Nagel [Jain 79]. L'hypothèse avancée est que si la caméra est immobile, le fond de la scène l'est aussi et donc seuls les objets en mouvement produisent des changements d'intensité de pixel. La différence s'exprime simplement de la manière suivante :

$$\delta I_t(i, j) = I_t(i, j) - I_{t-1}(i, j) \quad (1.11)$$

Ainsi les pixels pour lesquels  $\delta I$  est éloignée de zéro appartiennent à un objet en mouvement. Par analogie, il est souvent conclu que les pixels proches de zéro sont des pixels correspondant au fond de la scène. En réalité c'est le cas seulement si l'objet n'est pas de couleur uniforme. En effet, même en se déplaçant les pixels intérieurs à l'objet ne produiront pas de différence significatives ; seuls ceux proches du contour de l'objet seront détectés. Toutefois, il existe de nombreuses manières de pallier ce problème en ayant recours à des méthodes peu complexes (morphologie mathématique, croissance de régions).

Un problème plus délicat se pose quant au choix du seuil qui permet de déterminer quels pixels sont proches de zéro et lesquels ne le sont pas. D'une part ce seuil peut être amené à évoluer au cours du temps car la luminosité de la scène peut évoluer. D'autre part, ce seuil n'est pas forcément valable pour toute l'image car l'éclairage varie également localement.

Konrad [Konrad 00] utilise un test d'hypothèse pour classer les pixels dans les catégories "en mouvement" ou "immobiles". Sous chaque hypothèse, les vraisemblances sont calculées

d'après un modèle gaussien. Le rapport des vraisemblances soumis à un seuillage permet la classification. La méthode est étendue au cas où la caméra est mobile, les vraisemblances étant alors modélisées par des lois uniformes.

La détermination du seuil de décision peut être améliorée avec des méthodes de seuillages automatiques. Wu [Wu 05] présente une telle méthode appliquée à la détection de mouvement. Une image de ratios entre les intensités lumineuses des images consécutives est utilisée pour construire un histogramme cumulé. Le point de changement<sup>6</sup> de l'histogramme permet de déterminer la valeur optimale du seuil.

De la cohérence spatiale peut être gagnée en prenant en compte les pixels voisins. Les champs de Markov sont une méthode particulièrement adaptée à une telle prise en compte. Konrad [Konrad 00] propose l'utilisation des champs de Markov et obtient un seuillage adaptatif. Une autre possibilité est d'utiliser des descripteurs de niveau région en plus des descripteurs de niveau pixel. Lisani et Morel [Lisani 03] proposent un modèle dit *a contrario*. Les auteurs analysent des fenêtres carrées en calculant l'orientation du gradient. Les différences d'orientation sont ensuite seuillées, et en dénombrant les pixels noirs, le modèle *a contrario* détecte si la répartition obtenue localement est éloignée d'une répartition globale supposée uniforme, auquel cas il permet de mettre en avant un changement important dans la région examinée.

En outre, la cohérence temporelle peut être renforcée en prenant en compte plus de deux images consécutives. Veit *et al.* [Veit 04] partent eux aussi d'un modèle *a contrario* mais celui-ci repose sur une grandeur appelée résidu du flot normal :

$$\frac{I_{t+1}(\vec{p} + \vec{w}_\theta) - I_t(\vec{p})}{\|\nabla I_t(\vec{p})\|} \quad (1.12)$$

avec  $\vec{w}_\theta$  un déplacement paramétré par un vecteur  $\theta$ , lui-même estimé par une méthode d'optimisation. La cohérence temporelle est obtenue en calculant cette grandeur entre  $t$  et  $t-5$  maximum, la largeur de la fenêtre temporelle étant liée à l'amplitude du mouvement.

### Estimation d'un champ de déplacement ou de vitesse

Une méthode simpliste consiste à appareiller des régions (*block matching*). Cette méthode détermine la zone dans  $I_t$ , la plus proche au sens de l'intensité lumineuse d'une région étudiée dans  $I_{t-1}$ . Le déplacement moyen subi par la région étudiée est ainsi déterminé. L'approche est locale et doit être répétée pour chaque pixel de  $I_{t-1}$  constituant autant de centres de régions. L'appariement des régions peut se faire en minimisant la distance inter-région ou en maximisant la corrélation. Nous ne développons pas plus cette méthode généralement trop lente et peu précise. Elle se rapproche par ailleurs d'une procédure de suivi présentée en 2.3.1.

– **Flot optique** D'autres auteurs ont préféré exprimer le problème sous forme d'équations aux dérivées partielles donnant naissance à la plus connue des méthodes d'analyse du mouvement

<sup>6</sup>Point pour lequel la dérivée passe en dessous de 1.

dans les séquences d'images : le flot optique. La forme différentielle de la fonction image s'écrit :

$$dI = \frac{\partial I}{\partial t} dt + \frac{\partial I}{\partial i} di + \frac{\partial I}{\partial j} dj \quad (1.13)$$

$$\frac{dI}{dt} = \frac{\partial I}{\partial t} + \nabla I \cdot (u, v) \quad (1.14)$$

avec  $(u, v) = \left( \frac{di}{dt}, \frac{dj}{dt} \right)$  le champ des vitesses instantanées que l'on cherche à déterminer. En considérant que pour un laps de temps très court (celui séparant deux images), il y a conservation de l'intensité lumineuse, c'est-à-dire :

$$I(i + di, j + dj, t + 1) \approx I(i, j, t). \quad (1.15)$$

Il vient alors :

$$\frac{\partial I}{\partial t} + \nabla I \cdot (u, v) \approx 0. \quad (1.16)$$

Ce résultat est souvent appelé équation de contrainte du mouvement apparent (ECMA). La dérivée partielle par rapport à la variable temporelle s'exprime simplement par la différence inter-images. Le flot optique est plus précis que les méthodes de détection évoquées en début de cette sous-section, car il fournit en sortie un champ et non une simple classification binaire. Toutefois, sous la forme présentée, l'ECMA n'est pas résoluble (deux inconnues pour une équation), on parle de problème d'ouverture.

Horn et Schunck [Horn 81] réécrivent l'ECMA en l'intégrant par rapport à  $di$  et  $dj$ . Ils obtiennent ainsi une fonctionnelle à minimiser à laquelle ils ajoutent une contrainte dite de lissage. Le problème est cette fois suffisamment contraint pour être résolu. La contrainte impose une certaine régularité dans le mouvement observé et aboutit à une fonctionnelle convexe, ce qui facilite donc la minimisation de la fonctionnelle.

Parallèlement Lucas et Kanade [Lucas 81] proposèrent la même année une autre manière de résoudre l'ECMA. Les auteurs supposent que le champ de mouvement est constant sur un support local  $win(\vec{p})$  où  $\vec{p}$  est le pixel central de ce support. On obtient alors un système linéaire surdimensionné :

$$\begin{pmatrix} \frac{\partial I}{\partial i}(\vec{p}_1) & \frac{\partial I}{\partial j}(\vec{p}_1) \\ \vdots & \vdots \\ \frac{\partial I}{\partial i}(\vec{p}_K) & \frac{\partial I}{\partial j}(\vec{p}_K) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \frac{\partial I}{\partial t}(\vec{p}_1) \\ \vdots \\ \frac{\partial I}{\partial t}(\vec{p}_K) \end{pmatrix} = 0 \quad (1.17)$$

$$\sum_{k=1}^K \frac{\partial I}{\partial i}(\vec{p}_k) u + \frac{\partial I}{\partial j}(\vec{p}_k) v + \delta I(\vec{p}_k) = 0 \quad (1.18)$$

avec  $\{\vec{p}_k\}_{k=1}^K = win(\vec{p})$ . La solution de ce système est directement obtenue par une minimisation de type moindres carrés.

D'autres méthodes existent pour estimer le flot optique, mais sont plus rarement utilisées. Pour n'en citer que quelques unes, nous pouvons parler des approches de Fleet [Fleet 90], Heeger [Heeger 88], et Spinei [Spinei 98] qui sont dites fréquentielles et s'appuient sur la transformée de Fourier en mouvement. Le problème d'ouverture est géré en supposant que le mouvement est constant sur le support des filtres spatio-temporels directionnels utilisés (ex : filtres de Gabor).

L'approche de Weber et Malik [Weber 95] mérite également d'être citée. La séquence d'images y est préalablement convoluée avec un banc de filtres gaussiens, ce qui alimente la résolution du système et aide à faire ressortir des propriétés texturales.

Globalement, on peut dire que les méthodes de type flot optique sont assujetties aux défauts inhérents à l'hypothèse de conservation de la luminance. Cette hypothèse n'étant pas vérifiée en cas d'occultations, de transparences, de réflexions spéculaires et plus généralement dans le cas de changements brutaux de luminosité. Ce genre d'événements sont malheureusement fréquents dans les séquences routières.

La technique du flot optique a néanmoins été améliorée par de très nombreux procédés en vue de la rendre plus robuste. Notons par exemple qu'une dimension multi-échelle de l'algorithme aide à traiter des mouvement d'amplitudes variables. Récemment Ren [Ren 08] s'est attaqué au problème d'ouverture sous un angle nouveau. Un descripteur textural est utilisé pour regrouper des pixels en mouvement, et apporter une condition de cohérence spatiale moins brutale que celle de Lucas et Kanade. Nous renvoyons le lecteur vers Mémmin [Mémmin 03] pour une présentation plus formelle et plus complète du flot optique ainsi que des méthodes s'appuyant sur le flot optique.

– **Méthode d'Odobez et Bouthémy** Nous présentons dans ce paragraphe une méthode proche de celle du flot optique mais dont la résolution du problème s'écrit différemment. Elle a été introduite par Odobez et Bouthemy [Odobez 95], puis adaptée aux séquences routières par Lefaix *et al.* [Lefaix 02]. Le mouvement de l'objet est supposé paramétrable, c'est-à-dire qu'un ensemble fini de paramètres  $\theta = \{a_1, \dots, a_{N_{param}}\}$  permet de décrire le mouvement de l'objet d'après les lois classiques de la mécanique. Selon la complexité du mouvement, le nombre de paramètres  $N_{param}$  varie. Pour un point  $\vec{p}_k = (i_k, j_k)$  appartenant à l'objet, on trouve entre autres :

– la translation :

$$\begin{cases} u_k = a_1 + i_k \\ v_k = a_2 + j_k \end{cases} \quad (1.19)$$

– la transformation affine :

$$\begin{cases} u_k = a_1 + a_2 i_k + a_3 j_k \\ v_k = a_4 + a_5 i_k + a_6 j_k \end{cases} \quad (1.20)$$

– l'homographie :

$$\begin{cases} u_k = (a_1 + a_2 i_k + a_3 j_k) / (1 + a_7 i_k + a_8 j_k) \\ v_k = (a_4 + a_5 i_k + a_6 j_k) / (1 + a_7 i_k + a_8 j_k) \end{cases} \quad (1.21)$$

– le modèle quadratique :

$$\begin{cases} u_k = a_1 + a_2 i_k + a_3 j_k + a_4 i_k j_k + a_5 i_k^2 + a_6 j_k^2 \\ v_k = a_7 + a_8 i_k + a_9 j_k + a_{10} i_k j_k + a_{11} i_k^2 + a_{12} j_k^2 \end{cases} \quad (1.22)$$

Le modèle translationnel est généralement trop simpliste. Un mouvement en trois dimensions d'un objet, projeté dans le plan image, est entièrement décrit par une homographie. Pour plus de commodité dans les calculs, on lui préfère néanmoins souvent la transformation affine, qui possède deux paramètres de moins. De plus, la transformation affine est suffisante pour décrire

des mouvements contraints comme ceux concernant la face arrière d'un véhicule. Enfin le modèle quadratique s'adresse à des mouvements plus complexes d'objets non-rigides ou articulés. Notons que Lefaix *et al.* optent pour un modèle quadratique sous-paramétré (8 paramètres au lieu de 12) dont le degré de liberté est équivalent à l'homographie.

Quel que soit le modèle choisi, il est possible de réécrire l'obtention du champ de vitesse  $vel_\theta$  en un point donné sous forme matricielle :

$$vel_\theta(\vec{p}_k) = (u_k, v_k)^T = B(\vec{p}_k) \theta \quad (1.23)$$

où  $B(\vec{p}_k)$  est une matrice dont les valeurs dépendent uniquement de  $\vec{p}_k = (i_k, j_k)$  et dont la taille varie en fonction du modèle de mouvement. L'idée est alors de trouver le vecteur  $\theta$  qui décrit le mieux le mouvement, ce qui est possible en minimisant l'erreur  $E_\theta$  :

$$\hat{\theta} = \arg \min_{\theta} E_\theta = \arg \min_{\theta} \sum_k \rho \left( I_{t+1} \left( \vec{p}_k + v\vec{el}_\theta(\vec{p}_k) \right) - I_t(\vec{p}_k) \right) \quad (1.24)$$

avec  $\rho(\cdot)$  la fonction de Tukey. La somme est en général effectuée sur l'ensemble des pixels de l'image. La méthode propose une estimation récursive de  $\theta = \hat{\theta}_l + \delta\theta_l$  où  $\hat{\theta}_l$  est l'estimation courante de  $\theta$ . Sur le principe de l'ECMA l'expression d'un résidu  $r_{\delta\theta_l}(\vec{p}_k)$  est obtenue :

$$r_{\delta\theta_l}(\vec{p}_k) = \nabla I_t \left( \vec{p}_k + v\vec{el}_{\hat{\theta}_l}(\vec{p}_k) \right) \cdot v\vec{el}_{\delta\theta_l}(\vec{p}_k) + I_{t+1} \left( \vec{p}_k + v\vec{el}_{\hat{\theta}_l}(\vec{p}_k) \right) - I_t(\vec{p}_k) \quad (1.25)$$

Grâce à la méthode des moindres carrés pondérés itérés, on minimise l'erreur récursivement :

$$E(\delta\theta_l) = \sum_k \rho(r_{\delta\theta_l}(\vec{p}_k)) \quad (1.26)$$

A l'aide de propriétés de la fonction de Tukey, des poids sont obtenus en chaque pixel, traduisant l'adéquation du mouvement local au pixel au mouvement global de la scène. Ainsi les pixels trop éloignés du modèle estimé de la scène appartiennent à des objets ayant un mouvement indépendant. La méthode d'Odobez et Bouthémy possède donc un avantage par rapport à celles précédemment citées, car l'estimation du mouvement global de la scène permet une bonne détection des objets en mouvement même dans des séquences fortement dynamiques, comme les séquences routières. Cette méthode est de plus adaptable au multi-échelle. Demonceau et Kachi-Akkouche [Demonceaux 04] injectent des attributs texturaux obtenus par ondelettes à cet effet.

Une méthode relativement similaire a été proposée par Zhang *et al.* [Zhang 06]. Cette méthode s'applique très bien lorsque la caméra est en mouvement et donc dans les séquences routières. L'approche repose sur l'idée que lorsque la caméra subit une translation parallèle à l'axe optique, alors un point particulier, appelé point de fuite, apparaît. Ce point se caractérise par le fait que le champ de vitesse semble soit converger vers ce point, soit diverger depuis ce point. La position du point est estimée grâce au flot optique et des M-estimateurs [Huber 64]. Le mouvement global de la scène étant lié à la position de ce point, les résidus des M-estimateurs traduisent ici aussi l'adéquation au mouvement global. En revanche, la méthode ne présente plus d'intérêt quand la caméra devient immobile. Dans nos séquences, la caméra est la plupart du temps mobile, mais peut devenir immobile (embouteillage, feu rouge).

### 1.5.2 Méthodes d'analyse du fond de la scène

L'analyse du fond de la scène repose sur trois hypothèses principales :

- la caméra doit être fixe,
- des images du fond doivent être disponibles *a priori*,
- les occultations doivent rester limitées dans le temps.

D'emblée ces hypothèses entrent en contradiction avec nos besoins, mais si l'on veut faire un tour d'horizon des méthodes classiquement utilisées en analyse de mouvement, elles méritent d'être présentées. L'analyse du fond de la scène aboutit à une classification des pixels en deux classes : "fond" et "non-fond". Une phase d'apprentissage des propriétés visuelles du fond précède généralement une phase d'étiquetage des pixels.

Grimson *et al.* [Grimson 98] utilisent un mélange de 3 à 5 gaussiennes, défini pour chaque pixel, et modélisant l'appartenance au fond de la scène. Les paramètres des gaussiennes sont appris à l'aide d'un algorithme de K-moyennes puis par l'algorithme EM<sup>7</sup> sur les premières images (ce qui relâche certaines hypothèses énoncées plus haut). Au fil de la séquence d'images, chaque pixel observé est comparé aux composantes du mélange. Si l'observation correspond à un des modes principaux du mélange, l'observation est utilisée pour mettre à jour les paramètres de la gaussienne correspondant au mode. Dans le cas contraire, la composante ayant le plus faible poids est remplacée par une nouvelle, elle-même centrée sur l'observation et faiblement pondérée. Un seuil sur les poids des composantes permet de savoir si elles décrivent le fond de la scène ou une région d'un objet en mouvement.

Elgammal *et al.* [Elgammal 00] utilisent une méthode similaire avec mélange de gaussiennes, mais les gaussiennes sont centrées sur des vecteurs plus complexes. On définit un vecteur par pixel, et ce vecteur est obtenu par concaténation des valeurs précédentes du pixel examiné. Ainsi un gain en cohérence temporelle est obtenu. Un seuil sur la probabilité estimée par le mélange permet de déterminer si le pixel appartient au fond de la scène ou non.

La notion de cohérence temporelle se rapproche de celle de prédiction. Toyama *et al.* [Toyama 99] utilisent un filtre de Wiener (version simplifiée du filtre de Kalman) pour prédire la valeur d'un pixel sachant ses valeurs précédentes. Ainsi les pixels trop éloignés de la prévision sont étiquetés en tant que pixel d'objet en mouvement.

La cohérence spatiale est un autre axe d'amélioration possible. Sheikh [Sheikh 05] ajoute un noyau spatialement plus étendu que celui de Grimson. L'étiquetage est réalisé par une approche par champs de Markov<sup>8</sup> sous le principe du maximum *a posteriori*.

Des tentatives d'adaptation des méthodes de modélisation du fond à des caméras mobiles ont été effectuées. L'approche d'Elgammal en fait partie, mais cette méthode reste néanmoins limitée à des mouvements de faible ampleur. D'autres approches, comme celle de Ren *et al.* [Ren 03] s'adressent spécifiquement à ce cadre avec un succès limité. De manière générale, il ne semble pas naturel d'utiliser ces méthodes dans un contexte trop dynamique.

<sup>7</sup> cf. 2.4.2 pour plus de détails sur cet algorithme.

<sup>8</sup> cf. partie 2.5.1 pour plus de détails sur les champs de Markov.

### 1.5.3 Extraction en couches de mouvement

Les objets d'une scène ont tous un mouvement propre. Chaque mouvement projeté dans le plan 2D d'une image peut être vu comme une "couche". Extraire toutes les couches de mouvement revient à segmenter l'image sur des critères de mouvement. Les méthodes présentées dans cette sous-partie sont donc plus précises sur l'information fournie en sortie.

Choi et Kim [Choi 96] appliquent par exemple un processus de segmentation directement sur le champ du flot optique. La technique utilisée est classique en segmentation : croissance de régions. Shi et Malik [Shi 98] préfèrent la technique de coupure de graphe pour segmenter le flot optique. D'autres méthodes peuvent être appliquées et complexifiées, d'où de nombreux papiers sur le sujet qu'il ne nous apparaît pas pertinent de citer. Notons tout de même un article de Ross [Ross 03] qui présente différents critères topologiques et de trajectoire pour la segmentation d'objets en mouvement.

L'extraction des couches de mouvement repose plus souvent sur de multiples modèles paramétrés de mouvement (un modèle à estimer par couche). Cette vision pose trois problèmes principaux :

- (i) Combien de modèles sont nécessaires pour décrire les mouvements de la scène ?
- (ii) Quels sont les paramètres des modèles ?
- (iii) Quel est le support spatial relatif à chaque couche ?

Ayer et Sawhney [Ayer 95] définissent un mélange de modèles décrivant les changements d'intensité des pixels conditionnellement aux paramètres de mouvement. Les limites des supports des couches sont obtenues en considérant les probabilités d'appartenance aux couches. Le maximum de vraisemblance du mélange et les probabilités d'appartenance aux couches sont fournis par un algorithme EM. Enfin le nombre adéquat de modèles est donné par le principe d'encodage MDL (Minimum Description Length) appliqué aux paramètres de mouvement et aux résidus de l'estimation du maximum de vraisemblance. Des approches méthodologiques similaires sont présentées dans : [Darrel 91, Jepson 93, Wang 94, Weiss 97].

Plus récemment, Vidal [Vidal 04] a proposé une approche très différente. Le problème est abordé sous un angle algébrique et est unifié au sens où tous les modèles de mouvement sont déterminés conjointement. Les paramètres, pour un type de mouvement donné, appartiennent à un espace de type  $\mathbb{R}^K$  avec  $K$  la dimension de cet espace. Le problème est étendu aux complexes dans  $\mathbb{C}^K$ , espace dans lequel le problème revient à une équation d'hyperplan. Chaque observation de mouvement doit satisfaire une équation d'hyperplan qui peut être vue chacune comme un terme factorisable d'un polynôme. Les paramètres de tous les mouvements correspondent à l'unique racine du polynôme et sont obtenus via une méthode de type moindres carrés. En dérivant ce polynôme les paramètres des modèles sont séparés pour chaque couche. Reste à avoir à disposition ce genre d'observations de mouvement ; les auteurs donnent en exemple le flot optique. La méthode bien qu'efficace semble lourde à mettre en œuvre, les auteurs n'évoquent pas le temps de calcul nécessaire. Notons que la méthode est étendue aux caméras mobiles dans [Vidal 05].

Les méthodes de segmentation en couches de mouvement tendent vers des procédures de suivi. En effet, s'il est possible d'identifier tous les mouvements de la scène, il est clair qu'une grande partie du problème de suivi est résolu. Les procédures du type appariement de détection

sont particulièrement proches de l'extraction en couche de mouvement. Elles seront évoquées dans le chapitre suivant section 2.5.

#### 1.5.4 Méthodes à base de mouvement cohérent

Les méthodes à base de mouvement cohérent sont fondées sur le principe que le mouvement d'un objet possède des propriétés bien définies. Ainsi, les zones de l'image dont les détections de mouvement forment un tout cohérent, correspondent à des objets. L'idée est assez proche de l'extraction en couches de mouvement, mais au lieu d'analyser les mouvements de la scène, puis de s'appuyer sur des critères de cohérence, les méthodes à base de mouvement cohérent s'attaquent directement à caractériser les cohérences avant d'identifier les mouvements.

Wixson et Hansen [Wixson 00] partent de la mesure du flot optique. Celle-ci est accumulée au fur et à mesure de la séquence ; il est ainsi possible de retracer approximativement "le chemin" parcouru par chaque pixel. Le flot est accumulé dans la direction des lignes d'une part et dans la direction des colonnes d'autre part. C'est alors en réalité les incohérences de trajectoire qui sont détectées quand il y a changement de direction sur l'une des deux composantes. Tian et Hampapur [Tian 05] simplifient légèrement le schéma précédent en se focalisant sur le signe des directions, ce qui permet d'atteindre le temps-réel. L'image binaire de cohérence obtenue permet de constituer des graines pour une croissance de région étendue par le flot optique. Au final les objets dont le mouvement se détache nettement dans la scène sont segmentés.

Dans des cadres applicatifs bien précis, le mouvement de l'objet suivi possède certaines caractéristiques qui le rende plus facile à discriminer. Laptev *et al.* [Laptev 05] s'intéressent aux mouvements périodiques, *i.e.* des mouvements se répétant de manière cyclique. Les cycles offrent donc une cohérence temporelle intéressante à exploiter. Le cas typique d'un tel mouvement est par exemple un piéton en train de courir, le mouvement des jambes, du bassin et des bras se répétant périodiquement. Dans les séquences routières, bien que le mouvement d'un véhicule soit contraint, de tels mouvements ne peuvent être exploités, un véhicule pouvant par exemple changer de file ralentir ou accélérer de manière imprévisible.

Les méthodes s'appuyant sur la notion de cohérence ont l'avantage d'être robustes par rapport à des mouvements parasites comme par exemple un faible mouvement de caméra, ou le mouvement de végétaux dû au vent. Ce sont bien les mouvements principaux, c'est-à-dire les plus cohérents qui sont détectés. Néanmoins pour des scènes très dynamiques, comme les scènes routières, cette approche reste d'un intérêt limité.

#### 1.5.5 Bilan sur l'information de mouvement

Notons qu'il existe des méthodes transverses aux classes évoquées dans cette partie. Par exemple Kim *et al.* [Kim 05] proposent une méthode de modélisation du fond et d'extraction de couche. De même, Zhu *et al.* [Zhu 05] introduisent une approche s'appuyant sur les points d'intérêts du descripteur SIFT et combinant un modèle de fond et de la cohérence de mouvement.

De plus, il est en réalité également possible dans des cas bien particuliers d'estimer un mouvement à partir d'une seule image, et non au moins une paire d'images successives. La seule méthode à notre connaissance capable de résoudre ce problème est l'analyse du flou. Quand un

objet se déplace plus vite que ce que le dispositif d'acquisition ne peut tolérer, alors un flou apparaît au niveau de l'objet. En analysant ce flou il est même possible de remonter aux paramètres du mouvement. Dai et Wu ont produit des travaux récents sur ce thème [Dai 08]. Les méthodes d'analyse du flou font figure d'exception, car du coup ce n'est plus l'évolution des pixels qui est analysée, mais elles souffrent d'un contexte applicatif très restreint. En effet les images floutées ne constituent pas la majorité des images acquises lors d'une séquence.

En outre, les méthodes évoquées dans cette partie permettent au mieux de détecter une entité en mouvement dans la scène, entité correspondant à une région segmentée  $R$ . Selon la procédure de suivi employée par la suite, la détection n'est pas toujours suffisante en soi pour caractériser la présence d'un objet suivi à tel ou tel endroit de la scène. Par exemple si deux positions possibles sont examinées, comment valider le choix en s'appuyant sur la région  $R$ . Il faut être capable de mettre en correspondance la représentation de l'objet choisie et la région  $R$ . Pour une boîte englobante, cela peut être la distance entre le centre de la boîte et le centroïde de  $R$ , couplée à un rapport de surface. On peut aussi utiliser des histogrammes de mouvement comme dans Perez *et al.* [Perez 04].

## 1.6 Conclusion

Ce chapitre nous a permis de présenter l'ensemble des représentations d'un objet communément utilisées en informatique, ainsi qu'un grand nombre de méthodes de caractérisation d'objets populaires en traitement d'images ou en vision par ordinateur. La combinaison d'une représentation et d'une méthode de caractérisation forme un modèle d'objet. L'identification de ces modèles était l'objectif de ce chapitre. Il est donc possible de définir de très nombreux modèles d'objets, d'autant plus qu'il est également possible d'utiliser conjointement plusieurs méthodes d'extraction avec une représentation (modèles multi-attributs).

Bien qu'il soit encore trop tôt pour formuler un choix sur le modèle d'objet à utiliser, l'examen de l'état de l'art a déjà permis d'écartier un certain nombre de modèles inadaptés à nos séquences. Des critères de sélection supplémentaires viendront en complément dans la partie 2.7. C'est donc en conclusion du chapitre suivant que modèle d'objet et procédure de suivi prendront une forme plus précise. Une fois toutes les difficultés identifiées, nous verrons comment combler ces défauts et nous proposerons des contributions en ce sens.



# Chapitre 2

## Procédures de suivi d'objets

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>41</b>
<b>2.2</b>	<b>Détection et suivi</b>	<b>42</b>
<b>2.3</b>	<b>Méthodes d'optimisation d'une grandeur scalaire</b>	<b>43</b>
2.3.1	Maximisation de l'inter-corrélation	43
2.3.2	Minimisation de l'erreur	43
2.3.3	Minimisation d'une fonctionnelle d'énergie	45
2.3.4	Minimisation d'un coût	45
2.3.5	Minimisation d'une distance ou d'une mesure de dissimilarité	46
2.3.6	Remarques sur les méthodes d'optimisation	48
<b>2.4</b>	<b>Méthodes probabilistes</b>	<b>49</b>
2.4.1	Filtrage Optimal	49
2.4.2	Algorithme EM	54
<b>2.5</b>	<b>Méthodes d'appariements de détection</b>	<b>55</b>
2.5.1	Segmentation séquentielle	55
2.5.2	Appariements probabilistes et crédibilistes	57
<b>2.6</b>	<b>Remarques supplémentaires</b>	<b>58</b>
2.6.1	Mise à jour des modèles	58
2.6.2	Suivi d'objets multiples	59
2.6.3	Évaluation des performances	60
<b>2.7</b>	<b>Les principales difficultés en suivi d'objets</b>	<b>61</b>
2.7.1	Les occultations	61
2.7.2	Le fouillis	62
2.7.3	Les mouvements particuliers	63
2.7.4	Les changements d'illumination	64
<b>2.8</b>	<b>Conclusion et approche retenue</b>	<b>65</b>

---

### 2.1 Introduction

Quel que soit le modèle d'objet, les procédures de suivi reposent sur deux hypothèses principales :

- l'hypothèse de persistance de l'apparence. Cette hypothèse traduit le fait que les propriétés relatives à l'apparence de l'objet restent valables au moins pendant une certaine période.
- l'hypothèse de cohérence de la trajectoire. Cette hypothèse traduit le fait que le mouvement de l'objet suivi ne subit pas de changements de trajectoire trop brutaux, et que leur amplitude est limitée entre deux images consécutives de la séquence traitée.

D'une manière très générale, une procédure de suivi génère des hypothèses sur la localisation de l'objet, et valide ou invalide ces hypothèses en mettant en correspondance le modèle d'objet avec l'observation relative à une hypothèse. Notons que dans les procédures itératives, la mise en correspondance observation/modèle peut conditionner la génération d'hypothèses.

Nous entamons ce chapitre par quelques mots sur la différence entre détection et suivi d'objets. Nous proposons par la suite une classification des procédures de suivi selon la façon dont est posé le problème de suivi et non selon la façon dont il est résolu :

- méthodes d'optimisation d'une grandeur scalaire,
- méthodes probabilistes,
- méthodes d'appariements de détection.

Certaines méthodes classées dans des catégories différentes pourront utiliser des concepts très similaires, comme la théorie des probabilités ou des algorithmes d'optimisation. Notre classification est orientée sur la philosophie des approches de la littérature. Cette classification est discutable et ne satisfera sans doute pas tous les experts du domaine, mais nous espérons qu'elle permette de mettre en avant les inconvénients et avantages des méthodes discutées par rapport à nos ambitions.

En outre, nous pourrions voir dans ce chapitre que les techniques de mise en correspondance observation/modèle sont transverses aux trois classes de méthodes proposées. Elles sont une sorte de trait d'union entre modèle d'objet et procédure de suivi. Par exemple, une distance est une forme de mise en correspondance, en s'appuyant sur cette distance on peut alors générer une densité de probabilité ou bien utiliser une technique de minimisation. Nous ne détaillerons pas les méthodes de mise en correspondance au sein d'une partie spécifique, elles seront évoquées ponctuellement si elles présentent un intérêt ou un inconvénient dans certains cas particuliers.

## 2.2 Détection et suivi

D'une manière générale, une procédure de suivi ne peut être utilisée sans une procédure de détection de l'objet qui permet d'enclencher le suivi. En réalité, les méthodes de détection sont souvent très proches des méthodes de suivi, à ceci près qu'elles ne prennent pas en compte la dimension temporelle des séquences traitées. Dans ce mémoire, nous nous focalisons sur les méthodes de suivi et faisons l'hypothèse qu'une position de l'objet est disponible à l'instant où le suivi démarre.

On peut en outre considérer que la détection séquentielle est une forme particulière de suivi. Un algorithme de détection séquentielle se contente de rechercher image par image la position de l'objet sachant notre modèle. Il est très rare que les auteurs ne prennent pas un minimum en compte les informations passées, ne serait-ce que pour limiter la zone de recherche de l'objet dans l'image actuelle.

A l'inverse, certaines approches utilisent l'ensemble des positions précédemment calculées afin d'accélérer le suivi et d'améliorer la robustesse en éliminant des trajectoires aberrantes. Toutefois, plus le nombre d'informations passées prises en compte est important, plus la fiabilité de ces mêmes informations s'avère indispensable.

La grande majorité des auteurs optent pour une approche intermédiaire en exploitant une à deux positions précédentes. Nos travaux s'inscrivent dans cette lignée.

## 2.3 Méthodes d'optimisation d'une grandeur scalaire

Dans le cadre des méthodes à base de minimisation, le problème de suivi se pose à l'aide d'une fonction à valeur dans  $\mathbb{R}$ , qui représente l'adéquation ou respectivement l'antinomie entre le modèle d'objet et une observation relative à une position supposée de l'objet. En recherchant le maximum, respectivement le minimum, de cette fonction, on obtient ainsi directement la position réelle de l'objet dans l'image  $I_t$ .

Dans cette partie plus que dans les autres, des méthodologies de mise en correspondance objet/observations seront évoquées (distances, similarité, coût, etc.), car nombre de ces méthodes définissent une grandeur optimisable. Elles pourront être utilisées dans un contexte de suivi différent, auquel cas elles ne seront pas de nouveau discutées.

### 2.3.1 Maximisation de l'inter-corrélation

Quand l'objet est caractérisé par une simple imagerie  $W$ , cette dernière peut être mise en correspondance avec des images  $W_t^{(i)}$  extraites de l'image courante  $I_t$ , on parle alors en anglais de *template matching*. C'est la méthode de suivi la plus simpliste et aussi une des premières à avoir été testées. L'inter-corrélation est en général utilisée dans le cas de ces méthodes ; elle se définit par :

$$d_{inter-corr}(W, W_t^{(i)}) = \frac{\sum_{i,j} W(i,j) W_t^{(i)}(i+di, j+dj)}{\sum_{i,j} W(i,j)^2} \quad (2.1)$$

L'idée est de trouver le déplacement  $(di, dj)$  qui maximise l'inter-corrélation et donc la ressemblance entre le modèle et une observation. Cette recherche n'est en général pas effectuée par un algorithme d'optimisation, mais de façon exhaustive sur une région de l'image conditionnellement à la position précédente de l'objet. Ce genre de méthodes est rarement employé de nos jours car le temps de calcul est souvent exorbitant et la qualité du résultat limitée.

Velduis et Brodland [Velduis 99] ont proposé un modèle déformable pour pouvoir gérer des mouvements plus complexes de l'objet suivi. Récemment Hii *et al.* [Hii 06] ont utilisé avec succès la corrélation sur des images médicales. Les auteurs passent dans le domaine fréquentiel et proposent une optimisation de l'algorithme permettant de réduire le temps de calcul.

L'inter-corrélation mesure la ressemblance entre le modèle et l'observation ; à l'inverse les autres grandeurs à optimiser évoquées dans ce mémoire en mesurent l'écart.

### 2.3.2 Minimisation de l'erreur

Un grand nombre de méthodes déterministes repose sur le principe de minimisation d'une erreur  $Err \in \mathbb{R}$ . Cette erreur représente la différence entre le modèle et les observations locales à la position supposée de l'objet. Plus on minimise cette erreur, plus on se rapproche naturellement

de la position réelle de l'objet.

Il existe de multiples façons de définir l'erreur. Dans Gérard et Gagalowicz [Gerard 00], l'erreur est calculée entre l'image observée et une image synthétisée  $I_{synth,t}$  en utilisant un modèle type maillage 3D et des attributs de textures. Elle est exprimée par l'équation :

$$Err = \frac{1}{\#(\mathcal{P})} \sum_{\vec{p} \in \mathcal{P}} |I_t(\vec{p}) - I_{synth,t}(\vec{p})| \quad (2.2)$$

L'erreur est calculée en utilisant la norme 1 entre les valeurs des pixels, ensuite l'erreur est cumulée sur l'ensemble des pixels. Ce calcul a l'avantage d'être rapide et relativement précis. Plus l'image synthétisée est proche de l'observation, plus l'erreur est petite. Les auteurs trouvent un minimum à cette erreur grâce à l'algorithme des recuits simulés [Kirkpatrick 83, Cerny 85] qui est, lui, beaucoup plus coûteux en temps de calcul.

Georgescu et Meer [Georgescu 04] calculent deux erreurs : l'une s'appuyant sur une technique à base de flot optique pour trouver des paramètres de mouvement, l'autre reposant sur des critères de couleur. Les deux erreurs sont conjointement minimisées en vue de combiner les intérêts des deux approches.

S'il existe suffisamment de points pour lesquels il est possible de confronter une observation  $Y_i$  à un modèle  $X_i$ , l'erreur peut être calculée au sens des moindres carrés :

$$Err_{\vec{\theta}} = \sum_{i=1}^{N_{points}} (Y_i - X_i)^2 \quad (2.3)$$

avec  $N_{points}$  le nombre de points disponibles et  $\vec{\theta}$  un vecteur de paramètres servant à coder la position de l'objet. La méthode des moindres carrés, sorte de régression linéaire, donne directement la valeur de  $\vec{\theta}$  minimisant  $Err_{\vec{\theta}}$  si le modèle est linéaire. Il suffit alors de calculer une pseudo-inverse pour trouver où s'annule la dérivée de  $Err_{\vec{\theta}}$ .

L'algorithme le plus utilisé reposant en partie sur les moindres carrés est le filtre de Kalman. Toutefois, il est également possible de présenter ce filtre sous un angle probabiliste, c'est pourquoi il sera évoqué plus en détail dans la section 2.4.

Dans [Huang 02] les auteurs préfèrent utiliser une extension des moindres carrés, appelée M-estimateurs [Huber 64]. On parle alors de fonction d'énergie robuste [Georgescu 04]. L'idée des M-estimateurs est d'appliquer une fonction de coût aux erreurs, qui élimine les mesures aberrantes du processus itératif.

Dans [Feris 04, Koller 93, Regh 94] l'algorithme de Levenberg-Marquardt [Marquardt 63] est utilisé. Cette approche de type moindres carrés propose une solution hybride entre l'algorithme de Gauss-Newton et la descente de gradient. L'importance accordée à l'un des deux algorithmes est régulé par un terme d'amortissement.

Shi et Tomasi [Shi 94] parlent de dissimilarité plutôt que d'erreur. Cette dissimilarité est, dans l'esprit, similaire aux moindres carrés à ceci près qu'elle est locale à une fenêtre de l'image.

La norme  $L_2$  est appliquée entre les pixels de l'image observée  $I_t$  et celle d'une image obtenue par distorsion de  $I_{t-1}$ . La distorsion est effectuée conditionnellement au mouvement à estimer. En appliquant la formule de Taylor, les auteurs utilisent l'algorithme de Newton-Raphson afin de minimiser la dissimilarité et déterminer les paramètres de mouvement.

### 2.3.3 Minimisation d'une fonctionnelle d'énergie

Certains auteurs parlent parfois plutôt d'énergie que d'erreur. On peut interpréter cette grandeur comme l'effort à fournir pour passer du modèle à l'observation. La différence avec une erreur n'est pas fondamentale. Dans Feris *et al.* [Feris 04] une fonctionnelle d'énergie est définie comme une distance entre des attributs texturaux (coefficients d'ondelettes) plutôt qu'entre des valeurs de pixels.

La majorité des méthodes minimisant une fonctionnelle d'énergie sont celles reposant sur la théorie des contours actifs (*cf.* 1.4.1). Plutôt que de chercher à extraire un contour, la même méthode peut être employée pour suivre le contour au fil de la séquence, il suffit de faire le lien entre l'estimation du contour obtenu par convergence dans l'image précédente et l'initialisation de la recherche du nouveau contour dans l'image actuelle. Certains auteurs ont recours au flot optique pour faire ce lien. C'est le cas notamment dans [Ha 04].

Concernant les contours actifs géodésiques, l'énergie à minimiser se définit comme l'intégrale sur une courbe d'une fonction décroissante du gradient de la courbe et de la norme de la dérivée de la courbe. Utilisant les principes de la géométrie Riemannienne, il s'agit de trouver la courbe géodésique qui soit la plus proche des contours suivis. Cette énergie est minimisée en résolvant l'équation d'Euler-Lagrange. La théorie des courbes de niveaux fait évoluer la courbe vers la solution. Paragios et Deriche [Paragios 00] modifient la définition de l'énergie pour y intégrer un terme spécifique au suivi.

### 2.3.4 Minimisation d'un coût

La notion de coût est sémantiquement quasi-équivalente à celle d'énergie. Quel coût a la mise en correspondance du modèle avec une observation? Les méthodes de minimisation d'un coût doivent offrir une réponse à cette question et, bien sûr, rechercher l'observation de coût minimal, qui permettra la prise de décision quant à la position de l'objet ciblé.

Une fonction de coût est souvent définie pour mettre en correspondance des ensembles de points ou des pseudo-graphes. Parmi les premiers travaux ayant utilisé ce genre de caractérisation appliquée à la vision par ordinateur, ceux de Sethi et Jain sont particulièrement cités [Sethi 87]. Leur approche suppose une détection préalable des points d'intérêt dans les images  $I_{t-2}$ ,  $I_{t-1}$  et  $I_t$ . Si on appelle pour  $I_t$  cet ensemble  $P_t = \{p_i^t\}_{i=1}^{N_t}$ , la fonction de coût  $c_t$  est alors donnée par :

$$c_t(p_i^{t-2}, p_j^{t-1}, p_k^t) = w_1 \left( 1 - \frac{\overrightarrow{p_i^{t-2} p_j^{t-1}} \cdot \overrightarrow{p_j^{t-1} p_k^t}}{\| \overrightarrow{p_i^{t-2} p_j^{t-1}} \| \| \overrightarrow{p_j^{t-1} p_k^t} \|} \right) + w_2 \left( 1 - 2 \frac{\sqrt{\| \overrightarrow{p_i^{t-2} p_j^{t-1}} \| \| \overrightarrow{p_j^{t-1} p_k^t} \|}}{\| \overrightarrow{p_i^{t-2} p_j^{t-1}} \| + \| \overrightarrow{p_j^{t-1} p_k^t} \|} \right) \quad (2.4)$$

avec  $w_1$  et  $w_2$  des poids. La fonction est définie suivant une trajectoire temporelle du point  $p_i^{t-2}$  vers  $p_k^t$ . Le premier terme pénalise un changement de direction et le second un changement de la vitesse. La minimisation est obtenue par une méthode d'échange itérative gloutonne (*iterative exchange greedy method*). Une méthode gloutonne cherche à optimiser la fonction de coût par des recherches optimales locales successives. Cette recherche ici est basée sur des critères des plus proches voisins et de lissage. Le processus est répété d'avant en arrière (*forward and backward*) jusqu'à l'équilibre. Leur approche repose sur une hypothèse restrictive : aucun point ne disparaît ni n'apparaît dans la scène.

De nombreuses modifications ont été apportées à cette approche [Salari 90, Rangarajan 91]. Veenman *et al.* [Veenman 03] partent également de cette même approche mais ajoutent une contrainte de cohérence globale du mouvement et utilisent la méthode hongroise [Kuhn 55] pour optimiser la recherche du coût minimal. Cette méthode est très populaire pour la mise en correspondance de graphe.

La méthode hongroise pose le problème du coût sous forme matricielle, dont chaque élément de la matrice correspond à l'appariement d'un point du modèle avec un point observé. La valeur d'un élément de la matrice est liée au coût donné par l'équation 2.4. Le coût minimal est recherché grâce à des manipulations sur les lignes et les colonnes de la matrice, dans le but de valider les assignations les moins coûteuses, avant de réitérer le processus restreint au sous-ensemble de points qu'il reste à assigner. La méthode hongroise est en revanche excessivement coûteuse en temps de calcul, et paraît donc peu appropriée dans notre cadre applicatif.

Il existe des exemples d'approches où le coût n'est pas défini par rapport à un appariement de points. Li *et al.* [Li 02] calculent une fonction de coût selon trois critères différents prenant en compte respectivement la mise en correspondance, la disparition et l'apparition d'objets. La fonction de coût est minimisée par un algorithme de relaxation.

### 2.3.5 Minimisation d'une distance ou d'une mesure de dissimilarité

Sur la forme, une erreur n'est pas très différente d'une distance, car elles traduisent toutes les deux la proximité de deux jeux de caractéristiques. Sur le fond, la distance correspond à une définition mathématique précise, comportant des propriétés qu'une erreur ne possède pas forcément. Dans la pratique, la minimisation d'une distance est la procédure de suivi la plus répandue dans la littérature. Il est en effet naturel de vouloir localiser un objet en mesurant une distance entre le modèle et une observation. De plus, la plupart des caractérisations d'objets existantes peuvent être utilisées au travers de distances.

Si une mesure ne possède pas toutes les propriétés de la distance, mais seulement quelques unes d'entre elles, le terme de distance ne peut être employé. On parle alors de divergence ou de similarité.

Quand l'objet est caractérisé par un histogramme, de nombreuses distances ou dissimilarités peuvent être employées. Les plus fréquemment utilisées sont :

– la divergence de Kullback-Leibler :

$$\sum_{i=1}^{N_{histo}} h_1(i) \ln \left( \frac{h_1(i)}{h_2(i)} \right) \quad (2.5)$$

– les distances de type Minkowski :

$$\left( \sum_{i=1}^{N_{histo}} (h_1(i) - h_2(i))^p \right)^{1/p} \quad (2.6)$$

qui regroupe la norme  $L_1$ , la distance euclidienne, et le max quand  $p$  tend vers l'infini.

– la distance du  $\chi^2$  :

$$\sum_{i=1}^{N_{histo}} \frac{(h_1(i) - h_2(i))^2}{(h_1(i) + h_2(i))^2} \quad (2.7)$$

– la distance de Bhattacharyya :

$$\sum_{i=1}^{N_{histo}} \sqrt{1 - \sqrt{\frac{h_1(i) h_2(i)}{\|h_1\| \|h_2\|}}} \quad (2.8)$$

avec  $h_1$  et  $h_2$  deux histogrammes de taille  $N_{histo}$ . La littérature regorge de distance d'histogrammes, le chapitre 7.3 de [Trémeau 04] constitue une des listes les plus complètes de distances d'histogrammes. Si dans certains cas une distance est plus adéquate qu'une autre [Kang 04], le choix de la distance est secondaire par rapport au choix de la méthode permettant de minimiser cette distance.

L'approche de Comaniciu *et al.* [Comaniciu 00] minimise la distance de Bhattacharyya  $d_{Bhatta}(\cdot, \cdot)$  entre une densité de couleur apprise  $q$  et une densité de couleur observée  $p(y)$  en un point  $y$  de l'image  $I_t$  grâce à l'algorithme du *Mean-Shift*. Si une densité de couleur est proche d'un histogramme (cf.1.3.1), les auteurs utilisent la distance sous la forme :

$$d_{Bhatta}(p(y), q) = \sqrt{1 - \rho(p(y), q)} \quad (2.9)$$

$$\text{avec } \rho(p(y), q) = \sum_{u=1}^{N_{dist}} \sqrt{p_u(y), q_u} \quad (2.10)$$

$\rho$  est appelé coefficient de Bhattacharyya.  $u$  est une couleur donnée,  $N_{dist}$  le nombre de couleurs quantifiées. Minimiser la distance est équivalent à maximiser le coefficient. De plus, ce coefficient peut être estimé par l'approximation suivante :

$$\rho(p(y), q) \approx \frac{1}{2} \sum_{u=1}^{N_{dist}} \sqrt{p_u(y) q_u} + \frac{C_h}{2} \sum_{i=1}^{N_{spatial}} w_i k \left( \left\| \frac{y - x_i}{h} \right\|^2 \right) \quad (2.11)$$

$N_{spatial}$  est le nombre de pixels contenus dans le support du noyau  $k(\cdot)$ , ce support correspond à l'ensemble des points  $x_i$ ,  $C_h$  est une constante relative à la distribution  $p(y)$  et  $h$  est un paramètre du noyau. Les poids  $w_i$  représentent le degré de ressemblance locale au pixel  $x_i$  entre les deux distributions. Le premier terme de l'équation 2.11 est fixe ; il suffit de maximiser le second terme. Pour ce faire, les auteurs utilisent une remontée de gradient par l'algorithme du

*Mean-Shift*<sup>1</sup>. Les auteurs montrent que dans le contexte du Mean-Shift le noyau d'Epanechnikov est optimal. Le vecteur du mean-shift  $\vec{m}(y)$  s'écrit :

$$\vec{m}(y) = \frac{\sum_{i=1}^{N_{spatial}} x_i w_i g\left(\left\|\frac{y-x_i}{h}\right\|\right)}{\sum_{i=1}^{N_{spatial}} w_i g\left(\left\|\frac{y-x_i}{h}\right\|\right)} - y \quad (2.12)$$

avec  $g(\cdot)$  la dérivée de  $k(\cdot)$ . Ce vecteur est dans la direction du gradient de la distribution. En suivant ce vecteur on remonte donc vers le mode principal de la distribution du coefficient de Bhattacharyya. La procédure itérative de suivi est alors extrêmement simple :

- 1. calcul de vecteur de mean-shift  $\vec{m}(y)$ ,
- 2. translation de la position de la densité  $p(y)$ ,  $y \leftarrow y + \vec{m}(y)$ .

La norme du vecteur d'adapte à la proximité du maximum du coefficient de Bhattacharyya ce qui assure la convergence du procédé (*cf.* [Comaniciu 00]). La popularité du Mean-Shift a été acquise sur deux aspects principaux : tout d'abord sur la procédure du Mean-Shift qui est simple et rapide (complexité en  $O(N_{spatial})$ ), et d'autre part sur la modélisation de l'objet par une distribution de couleur, qui permet un gain en robustesse.

D'autres techniques d'optimisation de type descente de gradient peuvent être utilisées quand on cherche à minimiser une distance entre histogrammes, mais le Mean-Shift reste de loin le plus utilisé.

Quand l'objet est caractérisé par un ensemble de points ou de contours, la distance de Hausdorff peut s'avérer un outil efficace de mise en correspondance. Soient  $A$  et  $B$  deux ensembles de points, la distance  $H(A, B)$  entre ces deux ensembles se définit de la manière suivante :

$$H(A, B) = \max\{h(A, B), h(B, A)\} \quad (2.13)$$

$$h(A, B) = \sup_{a \in A} \inf_{b \in B} \|a - b\| \quad (2.14)$$

où  $\|\cdot\|$  est une norme donnée. La distance de Hausdorff mesure la plus mauvaise correspondance entre les points de  $A$  et  $B$ . Dans le cas de contours, la distance s'applique de façon similaire, il suffit à partir d'une image binaire de contours de considérer que l'ensemble des points blancs définit un ensemble. Une des qualités principales de la distance de Hausdorff est sa capacité à bien gérer les occultations partielles (*cf.* 2.7.1 pour plus d'éléments sur les occultations). En effet la distance est proche voire équivalente même si un morceau de contour est absent. Néanmoins cette distance est lourde à calculer. Huttenlocher *et al.* [Huttenlocher 93] utilisent cette distance. Les auteurs génèrent des hypothèses de localisation en appliquant une série de transformations à la fenêtre correspondant à l'objet dans l'image précédente. Ce groupe de transformations peut être par exemple un ensemble de translations, la recherche du minimum est donc quasi-exhaustive. De plus cette approche nécessite un modèle de contour mis à jour efficacement.

### 2.3.6 Remarques sur les méthodes d'optimisation

Les méthodes d'optimisation d'une grandeur scalaire sont toutes itératives. Grâce à un algorithme d'optimisation, une recherche du minimum ou du maximum global de cette grandeur est possible. La faiblesse de ces approches réside souvent au niveau de cette étape d'optimisation,

<sup>1</sup>Le terme de *Mean-Shift* est relatif à la technique de remontée de gradient, toutefois dans la littérature l'approche de Comaniciu *et al.* est très souvent appelée dans sa globalité *Mean-Shift*.

car d'une part la convergence peut être longue et d'autre part l'algorithme peut converger vers un minimum local et non global. Le minimum global permet théoriquement de remonter à la meilleure estimation de la position de l'objet. C'est tout du moins l'hypothèse qui est faite dans l'utilisation de ces algorithmes. On peut estimer en effet que si le minimum global ne renvoie pas à la position de l'objet, alors celui-ci a été mal modélisé.

Comme on peut le constater dans les paragraphes précédents, les méthodes les plus souvent utilisées pour l'optimisation sont du type descente de gradient au sens large. Parmi ces techniques, l'approche par gradient stochastique possède un avantage important, en ce que l'aspect aléatoire permet d'explorer dans une certaine mesure la fonction à minimiser. Cette exploration permet d'éviter de rester bloqué dans un minimum local. Bray *et al.* [Bray 07] ont récemment appliqué ce gradient stochastique pour minimiser une fonction de coût et suivre des objets articulés.

On utilise plus rarement la famille des méthodes de relaxation, bien que celles-ci soient très efficaces dans la recherche d'un minimum global. La convergence de ces méthodes est en effet trop longue pour traiter une séquence d'images dans son intégralité.

Le principal défaut à retenir de ces méthodes à base déterministe est qu'étant donné leur lenteur à converger, la zone de l'image actuelle  $I_t$  examinée est souvent restreinte ou fortement contrainte par les positions passées de l'objet. Ainsi, si l'objet possède un mouvement trop rapide, il peut sortir de la zone de recherche entre deux images et être perdu par l'algorithme.

## 2.4 Méthodes probabilistes

Les procédures de suivi dites probabilistes tendent à répondre à la question : quelle est la position de l'objet la plus probable dans l'image actuelle ? Notons que c'est parce que le problème est posé en terme de probabilité que les méthodes évoquées dans cette section s'y trouvent.

La réponse à la question se fait grâce à des probabilités conditionnelles, autrement dit "sachant" d'autres événements. Ces événements sont d'une part *toutes les observations disponibles depuis le départ*  $Y_{0:t}$  et d'autre part *les positions précédentes*  $X_{0:t-1}$ . Il s'agit donc d'un problème de type maximum *a posteriori* (MAP), c'est-à-dire de maximisation de  $P(X_t|Y_{0:t}, X_{0:t-1})$ .

### 2.4.1 Filtrage Optimal

Parmi les méthodes probabilistes, les techniques à base de filtrage optimal sont les plus fréquemment utilisées. Nous reprenons ici une présentation issue du tutoriel de Arulampalam [Arulampalam 02] et des travaux de thèse d'Arnaud [Arnaud 04a].

Dans le cadre du filtrage optimal le problème se pose sous forme d'état. Le vecteur  $X_t$  codant la position de l'objet est appelé vecteur d'état. Le vecteur d'observation  $Y_t$  regroupe l'ensemble des mesures et analyses effectuées sur l'image actuelle  $I_t$ . Les filtres optimaux reposent sur l'hypothèse suivante : l'état présent  $X_t$  dépend des états passés  $X_{0:t-1}$  et des observations passées et présentes  $Y_{0:t}$ <sup>2</sup>.

<sup>2</sup>L'observation  $Y_0$  à l'instant nul peut être vue comme l'ensemble des données connue *a priori*, elle sert ici surtout à simplifier les notations par rapport à la variable  $X$ .

La représentation du problème sous forme d'état suppose que soient connues une relation entre l'état présent et l'état précédent d'une part, et une relation entre l'état présent et l'observation présente d'autre part :

$$X_t = f(X_{t-1}, b_t) \text{ équation de transition} \quad (2.15)$$

$$Y_t = g(X_t, c_t) \text{ équation d'observation} \quad (2.16)$$

avec  $f$  et  $g$  deux fonctions *a priori* quelconques et  $b_t$  et  $c_t$  deux bruits *a priori* quelconques.

Le but du filtrage optimal est de trouver l'expression de la densité  $p(X_{0:t}|Y_{0:t})$ , appelée densité de filtrage. La connaissance de cette densité revient à une connaissance totale du problème du problème de suivi.

Le filtre de Bayes permet d'obtenir directement une densité,  $p(X_t|Y_{0:t})$ , voisine de la densité de filtrage sous trois conditions :

- la densité à l'instant 0 est disponible (*a priori*) :  $p(X_0)$
- l'état présent  $X_t$  dépend uniquement de l'état qui le précède (hypothèse de Markovianité) :  $p(X_t|X_{0:t-1}, Y_{0:t}) = p(X_t|X_{t-1})$
- l'observation présente  $Y_t$  ne dépend que de l'état présent  $X_t$  (autre volet de Markovianité) :  $p(Y_t|X_{0:t}, Y_{0:t-1}) = p(Y_t|X_t)$

Ces hypothèses sont résumées au travers de la figure 2.1 Le filtre procède alors en deux étapes :

**Figure 2.1 :** Chaîne reliant les états et les observations.

une étape de prédiction et une étape de mise à jour. La prédiction se fait en utilisant le formule de Chapman-Kolmogorov :

$$\begin{aligned} p(X_t|Y_{0:t-1}) &= \int p(X_t|X_{t-1}, Y_{0:t-1}) p(X_{t-1}|Y_{0:t-1}) dX_{t-1} \\ &= \int p(X_t|X_{t-1}) p(X_{t-1}|Y_{0:t-1}) dX_{t-1} \end{aligned} \quad (2.17)$$

$p(X_t|X_{t-1})$  est supposée connue et est appelée densité de transition.  $p(X_{t-1}|Y_{0:t-1})$  provient de l'itération précédente. La seconde étape est obtenue en appliquant la formule de Bayes :

$$\begin{aligned} p(X_t|Y_{0:t}) &= \frac{p(Y_t|X_t) p(X_t|Y_{0:t-1})}{p(Y_t|Y_{t-1})} \\ &= \frac{p(Y_t|X_t) p(X_t|Y_{0:t-1})}{\int p(Y_t|X_t) p(X_t|Y_{0:t-1})} \end{aligned} \quad (2.18)$$

Le dénominateur peut être considéré comme une constante de normalisation et ne pose donc pas de problème. Le terme  $p(Y_t|X_t)$  est supposé connu, on parle de vraisemblance par léger abus de langage. Malheureusement l'intégrale de l'équation 2.17 n'est pas résoluble dans le cas général. Les filtres de Kalman et particuliers présentés ci-après proposent alors respectivement une solution dans un cas particulier et une solution sous-optimale.

### Filtre de Kalman

Quand les fonctions  $f$  et  $g$  sont linéaires, on peut réécrire les équations 2.15 et 2.16 sous forme matricielle :

$$X_t = F_t X_{t-1} + b_t \quad (2.19)$$

$$Y_t = G_t X_t + c_t \quad (2.20)$$

avec  $F_t$  et  $G_t$  deux matrices et  $b_t$  et  $c_t$  cette fois deux vecteurs de bruits. Si de plus les bruits sont Gaussiens, alors le filtre de Kalman [Kalman 60] donne une réponse optimale au problème. Le cœur de la procédure du filtre est donné par les équations suivantes<sup>3</sup> :

$$\tilde{X}_t = F_t \hat{X}_{t-1} \text{ prédiction} \quad (2.21)$$

$$\hat{X}_t = \tilde{X}_t + K_t (Y_t - G_t \tilde{X}_t) \text{ correction} \quad (2.22)$$

$\tilde{X}_t$  est une estimation intermédiaire prédite par l'équation de transition,  $K_t$  est un scalaire appelé gain de Kalman, il permet de favoriser la prédiction ou l'observation selon sa valeur. Le facteur qui le suit, noté  $\nu_t = Y_t - G_t F_t \hat{X}_{t-1}$  est appelé innovation. Le formalisme complet décrivant le filtre Kalman est donné en annexe B.

Le filtre de Kalman a été adapté à de très nombreuses reprises pour le suivi d'objets. Nguyen *et al.* [Nguyen 04] obtiennent des résultats particulièrement intéressants. Au niveau de l'équation d'observation, ils utilisent une fonction robuste (similaire à celles employées dans les M-estimateur) au sein du modèle gaussien de la vraisemblance. Le filtre est rebaptisé filtre de Kalman robuste. Dans Weng *et al.* [Weng 06] l'équation de transition est définie par un modèle de mouvement de l'objet. Quant à l'étape de correction, elle fonctionne avec un module de détection d'objets entre deux images d'une séquence. Le filtre proposé est adaptatif car les paramètres du filtre évoluent en fonction du contexte.

### Filtre de Kalman étendu

Une des principales limitations du filtre de Kalman est qu'il ne peut être appliqué que dans le cas linéaire. Pour pouvoir traiter les cas non-linéaires, une des premières idées fût de partir du filtre Kalman et d'utiliser un développement de Taylor, méthode classique pour approximer un cas non-linéaire par un cas linéaire. Cette approche est appelée filtre de Kalman étendu et par

<sup>3</sup>Dans le cas présenté, les bruits sont supposés centrés et statistiquement indépendants.

conséquent est sous-optimale.

$$F_t \approx \left. \frac{df}{dX_t} \right|_{X_t=\hat{X}_{t-1}} \quad (2.23)$$

$$G_t \approx \left. \frac{dg}{dX_t} \right|_{X_t=\hat{X}_{t-1}} \quad (2.24)$$

$F_t$  et  $G_t$  sont ici les matrices jacobiennes des fonctions  $f$  et  $g$ . Rosales et Sclaroff [Rosales 99] ont proposé un exemple de filtre de Kalman étendu appliqué au suivi d'objets. Un modèle de projection non-linéaire est employé pour modéliser le mouvement de l'objet.

Notons que les reproches classiquement adressés au filtre de Kalman étendu sont d'une part que le calcul des matrices jacobiennes est coûteux et d'autre part que l'approximation de Taylor n'est pas toujours très satisfaisante.

### Filtre de Kalman sans parfum

Le filtre de Kalman sans parfum [Julier 97, Wan 00] tente de surmonter les défauts du filtre de Kalman étendu tout en s'attaquant à la même problématique. Le filtre sans parfum approxime la distribution de filtrage à l'aide d'échantillons, appelés sigma-points. Chaque point de ce nuage de points  $\left\{ X_t^{(i)} \right\}_{i=1}^{N_{points}}$  est associé à un poids  $w_t^{(i)}$ . Les points sont échantillonnés par un algorithme déterministe à chaque instant  $t$  à partir de l'estimation précédente de l'état. Ils sont propagés par la fonction non-linéaire  $f : \tilde{X}_t^{(i)} = f(\tilde{X}_{t-1}^{(i)})$ . On peut alors atteindre la moyenne  $\bar{X}_t$  et la covariance  $P_{X_t X_t}$  de la variable  $X_t$  :

$$\bar{X}_t = \sum_{i=1}^{N_{points}} w_t^{(i)} \tilde{X}_t^{(i)} \quad (2.25)$$

$$P_{X_t X_t} = \sum_{i=1}^{N_{points}} w_t^{(i)} \left[ \left( \tilde{X}_t^{(i)} - \bar{X}_t \right) \left( \tilde{X}_t^{(i)} - \bar{X}_t \right)^t \right] + B_t \quad (2.26)$$

avec  $B_t$  la covariance attachée au bruit  $b_t$ . En supposant que  $X_t|X_{t-1}$  suit une loi Gaussienne, ces paramètres suffisent pour décrire entièrement la loi, l'approximation proposée est alors judicieuse et efficace. L'ensemble de cette procédure est appelée transformation sans parfum<sup>4</sup>. La transformation est également appliquée par rapport à l'équation d'observation.

A complexité calculatoire équivalente, le filtre sans parfum offre une approximation du second ordre, comparé au filtre étendu qui se cantonne au premier ordre. Des problèmes de stabilité ont néanmoins été recensés. Merve [Van der Merwe 01] propose une amélioration du filtre sans parfum destinée à gérer cette instabilité.

Li *et al.* [Li 04] ont appliqué le filtre de Kalman sans parfum au suivi d'objets par contour. Leurs expériences montrent sans surprise que l'approche sans parfum procure une meilleure

<sup>4</sup>Le terme "sans parfum" n'a pas de justification précise, nous le traduisons directement de l'anglais "unscented".

estimation de l'état que le filtre de Kalman. En revanche, leurs conclusions montrent qu'une approche concurrente par filtre particulaire donne des résultats encore meilleurs. Nous présentons cette approche ci-après.

### Filtre particulaire

Les filtres particuliers [Gordon 93, Kitagawa 96], aussi appelés filtres à particules ou méthodes de Monte Carlo séquentielles, reposent principalement sur deux principes : l'approximation de Monte Carlo et l'échantillonnage d'importance. L'approximation de Monte Carlo permet de représenter la densité de filtrage par un ensemble d'échantillons, appelés particules, tirés selon cette même loi de filtrage. À l'inverse du filtre de Kalman sans parfum, l'échantillonnage est donc probabiliste. Comme il n'est pas possible d'échantillonner selon  $p(\cdot|Y_{0:t})$  dans le cas général, la technique d'échantillonnage d'importance est utilisée. Elle permet d'échantillonner par une autre loi  $q(\cdot)$ , choisie par l'utilisateur, et appelée loi d'importance. L'estimation de Monte Carlo reste valable en pondérant les échantillons par les poids  $w_t^{(i)} = \frac{p(\cdot)}{q(\cdot)}$ . Grâce au contexte markovien, l'estimation de ces poids peut se faire de manière récursive, la procédure est alors applicable.

Le formalisme des filtres particuliers est présenté plus en détail en section 3.1. Le grand avantage des filtres particuliers, est qu'ils ne posent aucune restriction sur les fonctions  $f$  et  $g$  ou les bruits  $b_t$  et  $c_t$ . L'utilisateur a donc beaucoup plus de liberté pour modéliser les relations entre états et observations. En revanche les filtres particuliers restent une approche sous-optimale.

L'algorithme CONDENSATION de Isard et Blake [Isard 98a] est un des premiers filtres particuliers adaptés au suivi d'objets et le plus cité dans la littérature. Utilisant une représentation de type contour de l'objet et des caractéristiques de forme, CONDENSATION s'est montré capable de suivre un objet dans des conditions plus difficiles que les approches de l'état de l'art d'alors. Cet algorithme a ensuite évolué dans une forme que ces mêmes auteurs ont appelé ICONDENSATION [Isard 98b]. La fonction d'importance dans ICONDENSATION permet d'allier une approche locale et une approche globale du suivi. Perez *et al.* [Perez 04] ont eux choisi une représentation de type boîte englobante et une caractérisation d'objets similaire à celle du Mean-Shift, c'est-à-dire par densité de couleur. Cette approche très facile à mettre en œuvre est appliquée au suivi de visage en temps réel.

### Variante du filtre particulaire

La littérature regorge de variantes du filtre particulaire, nous ne citerons que quelques unes des plus marquantes. Le filtre particulaire auxiliaire proposé par Pitt [Pitt 99] ajoute une étape d'échantillonnage intermédiaire, dit échantillonnage auxiliaire. En pratique, cela permet d'orienter les filtres particuliers vers des zones pertinentes de l'image avant même que la propagation classique ait lieu. En théorie, cela permet de se rapprocher de la fonction d'importance optimale. Ce filtre a été mis en œuvre pour le suivi par Brèthes *et al.* [Brèthes 06] et Mac Kenna *et al.* [Kenna 07].

Toujours dans l'idée d'optimiser la fonction d'importance, Arnaud et Mémin [Arnaud 04b] ont proposé un algorithme de suivi utilisant directement un modèle pour la fonction optimale d'importance (plus de détails sur cette fonction en 3.1). L'image  $I_t$  est considérée comme une variable aléatoire et est incorporée dans la chaîne de Markov. Dans le cas de bruits gaussiens, les auteurs montrent que la fonction d'importance optimale est également gaussienne et que sa moyenne et sa covariance peuvent être obtenues.

Enfin le principe de la Rao-Blackwellisation est également fréquemment utilisé. Ce principe s'applique si une partie de l'état évolue de façon déterministe, l'idée est alors de ne propager que la partie non-déterministe de l'état. Ce principe est moins intéressant car il n'est pas applicable en général.

Un article comparatif de Brèthes [Brèthes 06] tend à montrer que le filtre particulière auxiliaire donne de meilleurs résultats. Néanmoins les résultats de cette étude sont aussi corrélés à une certaine forme de fusion d'informations implicite dans les filtres étudiés.

## 2.4.2 Algorithme EM

L'algorithme EM (Espérance-Maximisation) [Dempster 77] est un algorithme probabiliste qui se divise en deux phases : l'étape E et l'étape M. L'étape E consiste à calculer l'espérance de la vraisemblance. L'étape M optimise cette vraisemblance par rapport aux paramètres recherchés. La procédure est itérée jusqu'à convergence.

Jepson et Black [Jepson 93] partent d'un modèle de mouvement par mélange de distributions. Différents vecteurs de mouvement sont observés localement. L'étape E cherche à classifier ces vecteurs de mouvement en couches de mouvement. Selon cette classification, l'étape M maximise la log-vraisemblance par rapport aux paramètres de mouvement global de l'objet. Une fois obtenus, ces paramètres permettent de déterminer la position de l'objet suivi.

Plus récemment, Jepson *et al.* [Jepson 03] ont proposé une extension de cet algorithme gérant les changements d'apparence de l'objet et adaptant le modèle en conséquence. La représentation de l'objet est une ellipse englobante. Le modèle est divisé en trois parties : une première partie est définie par une composante stable, une seconde pour les erreurs aberrantes et une dernière pour un suivi temporellement local (d'une image à la suivante). La vraisemblance d'une observation se définit comme un mélange de ces trois parties. Cinq paramètres de cette vraisemblance sont optimisés grâce à l'algorithme EM. L'étape E permet de classifier les observations selon les trois composantes du modèle. L'étape M fait évoluer les paramètres conditionnellement à la classification. Pour pouvoir appliquer EM à la volée, les classes des observations passées ne sont pas mises à jour, et l'étape M est présentée sous une forme récursive. En utilisant cette vraisemblance et en s'inspirant des travaux précédents, les paramètres de mouvement de l'objet sont obtenus par un autre algorithme EM.

EM permet de modéliser certains problèmes relatifs au suivi d'objets, mais impose de présenter le problème de manière itérative. Ces itérations ont de plus un coût en temps de calcul, dans [Jepson 03] le traitement d'une image demande 10 secondes avec un processeur de 400Mhz.

## 2.5 Méthodes d'appariements de détection

Les méthodes dites d'appariement posent le problème de suivi de la manière décrite ci-après. Soient  $n$  objets ou régions l'image  $I_{t-1}$ , et  $m$  autres dans l'image  $I_t$ . Quel est alors l'appariement des  $n$  régions avec les  $m$  nouvelles régions qui traduit l'évolution de la scène ? Sachant quelles régions appartenaient à un objet suivi dans  $I_{t-1}$  et sachant cet appariement, il est désormais possible de savoir quelles régions correspondent à mon objet dans  $I_t$ .

Ces méthodes s'appuient en général sur une méthode de détection, qui permet de générer des hypothèses, ou des positions candidates de l'objet. La méthode de détection ne répondant pas au problème de suivi, c'est l'appariement qui exécute cette tâche.

### 2.5.1 Segmentation séquentielle

L'approche d'appariement la plus répandue est la segmentation séquentielle. Les méthodes de segmentation séquentielle utilisent des images segmentées de la scène analysée. On obtient ainsi  $n_{t-1}$  régions  $\{\mathfrak{A}_{t-1}\}_{i=1}^{n_{t-1}}$  dans l'image  $I_{t-1}$  et  $n_t$  dans l'image  $I_t$ . La procédure de suivi correspond alors à une question d'appariement des régions  $\mathfrak{A}_{t-1}$  avec les régions  $\mathfrak{A}_t$ . Il reste ensuite à identifier quelles régions forment l'objet suivi.

Izquierdo dans ses travaux de thèse propose une méthode de mise en correspondance de régions [Izquierdo 04] utilisant l'algorithme EM. La figure 2.2 montre un processus de segmentation et d'appariement. Notons que la technique employée (algorithme EM) est à base probabiliste,

**Figure 2.2 :** Segmentation de deux véhicules, puis labélisation et estimation des régions globales correspondant à chaque véhicule.

toutefois le problème se pose sous forme d'appariement des régions.

Concernant la méthode de segmentation<sup>5</sup> à utiliser, les champs de Markov (*Markov Random Fields*, MRF) [Geman 84] restent une des méthodes les plus efficaces. Bien que dans la théorie des champs de Markov le problème se pose comme un problème de maximisation de probabilité, dans la pratique le problème posé est résolu par des méthodes itératives déterministes. De plus, les champs de Markov ne posent pas le problème sous la forme : "quelle est la position de l'objet la plus probable ?", mais plutôt sous la forme : "quelle est la segmentation la plus probable de mon image par rapport aux données image ?". Les champs de Markov ne donnent pas en sortie la position de l'objet suivi, une étape d'identification de l'objet parmi les régions segmentées est nécessaire. C'est pourquoi nous présentons cette approche dans cette sous-section. Les travaux concernant un étiquetage haut-niveau des données (par exemple une étiquette "mon objet" et une étiquette "fond de l'image") restent anecdotiques et coopèrent avec

<sup>5</sup>Nous faisons ici l'amalgame entre méthode de segmentation et méthode d'étiquetage.

d'autres méthodes.

Si on appelle  $S_t$  l'ensemble des sites dont l'étiquette de région est à déterminer et  $Y_t$  les données images observées, les MRF ont pour but de maximiser la probabilité suivante :

$$P(S|Y) \propto P(Y|S) P(S) \quad (2.27)$$

La probabilité est souvent répartie en deux termes : un terme de régularité ( $P(S)$ ) et un terme d'attache aux données ( $P(Y|S)$ ). Sous hypothèse de markovianité et d'indépendance des observations, le terme d'attache aux données s'écrit :

$$P(Y = y|S = s) = \prod_{s' \in \mathfrak{V}(s)} P(y_{s'}|s_{s'}) \quad (2.28)$$

$\mathfrak{V}(s)$  étant le voisinage du site  $s$ . Les termes  $P(y_{s'}|s_{s'})$  sont souvent modélisés par des mélanges de gaussiennes dont les paramètres sont appris. D'après le théorème d'Hammersley-Clifford [Hammersley 71], le terme de régularité s'exprime lui de la manière suivante :

$$P(S = s) \propto \exp\left(-\sum_{c \in \mathcal{C}} V_c(s)\right) \quad (2.29)$$

avec  $c$  une clique, c'est-à-dire un couple de sites en relation de voisinage, et  $V_c$  une fonction de potentiel à définir. Les méthodes employées pour trouver le minimum sont en général des méthodes de relaxation : algorithme ICM [Besag 86], HCF [Chou 90], algorithmes génétiques [Kim 00], ou recuits simulés [Geman 84].

Kamijo et Sakuchi [Kamijo 02] ont apporté une modification importante aux MRF dans le cadre du suivi d'objets. Ils introduisent le concept de champs de Markov aléatoire spatio-temporel. Cette fois-ci les cliques ne sont plus seulement définies dans une image donnée (information spatiale) mais aussi entre des pixels de différentes images (information temporelle). La figure 2.3 illustre cette idée. La différence principale avec l'approche classique réside dans la défini-

**Figure 2.3 :** Champ de Markov aléatoire spatio-temporel.

tion de la fonction d'énergie qui dépend de l'image précédente et de la segmentation précédente.

Les champs de Markov sont aujourd'hui dépassés par une approche similaire : les champs conditionnels aléatoires (Conditional random fields, CRF). Les CRF cherchent à modéliser directement la probabilité recherchée  $P(S|Y)$  sans passer par la règle de Bayes. Les travaux de Nicolas [Nicolas 06] illustrent en quoi les CRF donnent de meilleures performances dans une application de classification. Toutefois, à notre connaissance les CRF n'ont pas encore été utilisés dans le cadre du suivi d'objets.

Les MRF comme les CRF souffrent en général des mêmes critiques : un apprentissage est requis, et la complexité algorithmique est prohibitive.

Toujours dans une vue graphique de l'image, la technique des coupures de graphes (*graph-cuts*) a donné des résultats intéressants en terme de segmentation et de suivi. Si l'image est représentée par un graphe global  $\mathcal{G}$ , chaque région ou objet de l'image est un sous-graphe de  $\mathcal{G}$ . On attribue un poids aux arcs du graphe en fonction de la cohérence qui lie les deux sommets du graphe. Cette cohérence peut être jaugée sur la couleur, la luminosité ou la texture par exemple. En coupant un arc du graphe, le poids associé à cet arc sera affecté comme un coût imputable à la coupe. Le coût total d'une coupe est appelé capacité. Il faut alors trouver la coupe de capacité minimale, segmentant l'image  $I_t$ . Shi et Malik [Shi 98] couplent leur méthode de segmentation de mouvement (*cf.* 1.5.3) par coupes de graphes à un algorithme de suivi. Plus récemment, Bugeau [Bugeau 07b] propose de suivre et segmenter des objets en minimisant une énergie par graph-cuts. L'information de mouvement est utilisée sur les images segmentées pour suivre pas-à-pas l'évolution des régions segmentées. Cette approche permet de plus de segmenter efficacement des objets ayant fusionnés, c'est-à-dire qui occupent une zone de l'image commune.

Les méthodes de segmentation séquentielle sont sans doute celles qui nécessitent le moins d'*a priori* sur l'objet. La segmentation s'opère à partir du moment où une frontière existe entre l'objet et le fond de l'image. En revanche, au moment de l'appariement des régions, des descripteurs sont généralement utilisés pour une meilleure robustesse. De plus, le résultat d'une telle procédure de suivi dépend fortement de la qualité de la segmentation.

### 2.5.2 Appariements probabilistes et crédibilistes

Une fois des positions candidates définies (observations ou prédictions), les appariements probabilistes cherchent naturellement à choisir la position qui maximise la probabilité d'être la position réelle de l'objet. Deux de ces techniques sont principalement présentes dans la littérature : le MHT ("Multiple Hypothesis Tracking") et le PDAF ("Probability Data Association Filtering").

Le MHT a été introduit par Reid [Reid 79], puis a été adapté à la problématique du suivi d'objets par Cox [Cox 93]. Il a été ensuite repris par de nombreux auteurs, notamment Cham et Rehg [Cham 99]. Le MHT part du constat qu'une mise en correspondance modèle/observation peut être erronée à un moment donné de la procédure de suivi, si la correspondance ne se fait qu'entre deux images consécutives. Le MHT répertorie au fil du temps les observations donnant un score de correspondance pour différentes positions prédites par un algorithme sous-jacent. Dans l'approche de Cox, cet algorithme est un filtre de Kalman. Le MHT au final analyse les données répertoriées pour définir la trajectoire maximisant les associations modèle/observation sur toute la période de temps. C'est en quelque sorte un lissage de la trajectoire. Le MHT est en

revanche complexe à mettre en œuvre et ne s'utilise que pour des vidéos "off-line" (il ne traite pas les images à la volée), ce qui ne correspond pas à notre cadre d'étude.

Le PDAF [BarShalom 88, Cox 93] se présente comme une extension du filtre de Kalman et utilise une approche bayésienne par rapport au problème d'association de données. Pour une cible donnée et un ensemble de positions candidates d'après les observations, le PDAF attribue un poids à ces observations selon deux postulats. Le premier est que seule une position candidate est la bonne, bien qu'elle puisse occasionnellement disparaître lors d'une occultation par exemple. Le second est que toutes les autres positions sont des "faux-positifs" dus à un bruit uniforme. Sous cette hypothèse d'uniformité, il est possible de calculer l'innovation de Kalman (*cf.* equation 2.22), comme la somme des innovations de chaque candidat :

$$\nu_i = \sum_{i=1}^{N_{candidates}} \beta_i \nu_i \quad (2.30)$$

avec  $\beta_i$  la probabilité d'association de l'observation candidate  $i$  avec la cible, et  $N_{candidates}$  le nombre d'observations candidates retenues. Un exemple de PDAF pour le suivi d'objets est donné par Rasmussen [Rasmussen 01]. Le PDAF peut aussi être utilisé avec un filtre particulière (*cf.* Schultz *et al.* [Schulz 01])

Toujours dans le cadre d'un filtre particulière, Gidel *et al.* [Gidel 08] ont récemment proposé deux alternatives non-paramétriques pour le problème d'appariement des observations avec les prédictions. Les méthodes proposées sont fondées soit sur un fenêtrage de Parzen soit sur les k-plus-proches voisins. Le fenêtrage de Parzen consiste en l'application d'une fonction de type noyau à support fini sur les distances entre prédictions et observations. L'approche des plus proches voisins repose sur le dénombrement d'échantillons dans la projection sur l'image d'un volume donné de la scène. Les deux approches réduisent le nombre de paramètres à régler, mais l'utilisation de la fenêtre de Parzen s'avère plus efficace du point de vue des résultats.

Dans le domaine des fonctions de croyance, autre théorie de l'incertain, des travaux relatifs à l'association de données donnent également de très bons résultats. Royère *et al.* [Royère 00] ont dans un premier temps proposé un mécanisme d'appariement dans le cadre des fonctions de croyance, puis une version crédibiliste du MHT a été proposée dans [Mourllion 05]. Parallèlement aux travaux de Mourllion *et al.*, Megherbi *et al.* [Megherbi 05c, Megherbi 05b] ont également mis au point un modèle permettant d'opérer un suivi multi-cibles, multi-capteurs. Les auteurs utilisent des informations sonores et visuelles pour effectuer l'association de données. Dans [Megherbi 05a], le problème de l'apparition et de la disparition d'objets est traité à l'aide d'une grandeur représentant le conflit généré par les appariements et les sources. Cette grandeur n'a pas d'équivalent en probabilités d'où l'intérêt, entre autres, des modèles crédibilistes.

## 2.6 Remarques supplémentaires

### 2.6.1 Mise à jour des modèles

La mise à jour du modèle est une option possible pour gagner en robustesse si le milieu dans lequel évolue l'objet est amené à subir de fortes modifications visuelles. C'est notamment le cas des séquences routières. Il paraît difficile de pouvoir mettre au point un système ayant appris

*a priori* l'ensemble des modifications possibles. En mettant à jour les caractéristiques de l'objet ciblé, l'algorithme s'adapte aux modifications.

Bien qu'à première vue cette solution paraisse simple et attirante, elle est en pratique très délicate à mettre en œuvre. En effet, la mise à jour repose sur l'hypothèse que les estimations précédentes de la position de l'objet sont correctes. En utilisant ces positions précédentes de nouvelles caractéristiques sont extraites de la zone correspondant à l'objet. Ainsi si cette zone ne couvre pas parfaitement l'objet ou couvre une partie du reste de la scène, des caractéristiques erronées sont apprises. Cela aboutit alors à une divergence de l'algorithme de suivi, appelée "dérive".

Un module de mise à jour est toutefois utilisé avec un certain succès par exemple dans les approches de Jacquot *et al.* [Jacquot 05] ou de Handmann *et al.* [Handmann 00]

### 2.6.2 Suivi d'objets multiples

Nous n'aborderons que très brièvement le suivi d'objets multiples dans ce mémoire, car il ne s'agit pas d'un des objectifs fixés par le sujet de cette thèse. Au contraire nous cherchons plutôt à suivre un objet particulier dans une scène.

Les méthodes de type appariement de détections sont celles qui s'adaptent le plus facilement au suivi d'objets multiples, car les méthodes de détections sous-jacentes détectent en général tous les objets de la scène ou au moins partiellement. Il est donc relativement aisé d'étendre le mécanisme d'appariement pour passer au suivi de plusieurs objets. Dans Kamijo et Sakauchi [Kamijo 02] l'algorithme peut suivre d'emblée un grand nombre d'objets. Le filtre PDAF peut lui être étendu au multi-objets, dans sa version "joint" (JPDAF) (*cf.* Rasmussen [Rasmussen 01]).

Les algorithmes probabilistes s'adaptent plutôt bien au contexte multi-objets. Il est possible de définir une variable aléatoire par objet suivi et de faire interagir ces variables à l'aide de probabilités jointes. Isard et Mac Cormick [Isard 01] utilisent un filtre particulière et un vecteur d'état global à tous les objets est alors défini. Il est la concaténation des états des objets et du nombre d'objets. Chaque sous-état individuel est doté d'une variable correspondant à l'étiquette de l'objet. Toutes les particules sont alors exploitées pour chaque suivi engagé. Czyz *et al.* [Czyz 05] se contentent d'ajouter le nombre d'objets suivis comme variable aléatoire supplémentaire. Des densités jointes et conditionnelles par rapport à cette variable sont proposées et permettent entre autres de gérer l'apparition et la disparition d'objets dans le champ de la caméra. Khan *et al.* [Kahn 05] adoptent une méthode similaire mais ajoutent une étape d'échantillonnage par chaîne de Markov Monte Carlo (Markov Chain Monte Carlo, MCMC). Cette étape permet une meilleure diffusion des particules à l'aide de densités intermédiaires. Pour pouvoir traiter un nombre variable d'objets, c'est l'algorithme MCMC à sauts réversibles (Réversible Jump MCMC) qui est utilisé. Leurs expériences montrent qu'un grand nombre de cibles sont suivies avec succès.

Enfin, les algorithmes minimisant une grandeur scalaire en revanche sont plus délicats à transposer au contexte multi-objets. La solution naïve consiste à lancer autant de "trackers"<sup>6</sup> que d'objets à suivre au prix d'un temps de calcul démesuré. Il est aussi possible de définir le

<sup>6</sup>Nous entendons par tracker une instance d'un algorithme de suivi.

problème de manière plus globale, mais le critère d'optimalité devient beaucoup plus difficile à optimiser. Notons que cette remarque n'est pas valable pour les contours actifs géodésiques, qui de par leur nature peuvent détourner directement plusieurs objets à la fois (*cf.* Paragios et Deriche [Paragios 00]).

### 2.6.3 Évaluation des performances

L'évaluation des performances des algorithmes de suivi comme dans la majorité des disciplines informatiques se fondent sur des critères de précision, de robustesse et de vitesse d'exécution. L'algorithme parfait est celui qui détermine en temps réel la position de l'objet précisément, quelle que soit les situations rencontrées dans la séquence. Néanmoins cette évaluation pose de grandes difficultés par rapport aux autres disciplines en raison de deux points en particulier :

- la subjectivité de la qualité des résultats : difficile de définir ce qu'est un "bon suivi". Cette notion est laissée à l'appréciation de l'utilisateur humain.
- l'hétérogénéité des algorithmes : si deux algorithmes utilisent une représentation différente de l'objet difficile de comparer leurs performances. Il en va de même pour la nature des séquences traitées, la plupart des algorithmes étant relativement *ad hoc*.

Dans la pratique, une majorité d'auteurs démontre l'efficacité d'une approche sur une vidéo en montrant certaines images sélectionnées à l'intérieur de la vidéo auxquelles la position estimée de l'objet est superposée. Il est possible de sélectionner des images correspondant à certains événements particuliers dans la séquence pour prouver la robustesse de l'algorithme. Il est également possible de sélectionner les images à intervalles de temps réguliers pour prouver la constance des résultats.

Pour comparer deux approches, une même séquence est utilisée pour les deux algorithmes et il convient naturellement de présenter les résultats sur les mêmes images de la séquence. Nous aurons recours cette manière classique et visuellement parlante de présenter les résultats pour différentes séquences.

Certains auteurs ont tenté de mettre au point des procédés plus objectifs pour évaluer les algorithmes de suivi [Bashir 06]. Pour plus de transparence, nous aurons également recours à la mesure dite du "taux de suivi" inspirée de [Bashir 06]. Cette mesure permet de suivre au fil de la séquence à la fois la précision et la robustesse d'un algorithme. Son principal défaut est la nécessité d'une vérité-terrain, c'est-à-dire d'un ensemble de données codant image par image la position réelle de l'objet (en fonction de la représentation objet choisie). Cette vérité-terrain n'est obtensible que par l'expertise d'un agent humain déterminant image après image la position de l'objet. La mesure du taux de suivi, noté  $r_t$ , se définit alors de la manière suivante :

$$r_t = \frac{2 \times S(A_t \cap B_t)}{S(A_t) + S(B_t)} \quad (2.31)$$

où  $S$  est l'opérateur donnant la surface en pixels d'une région,  $A_t$  est la boîte englobante estimée par l'algorithme jugé dans l'image  $I_t$ ,  $B_t$  est la boîte fournie par la vérité-terrain sur cette même image. Le taux  $r_t$  vaut 1 si et seulement si  $A_t = B_t$ , des situations telles que  $A_t \subset B_t$  ou  $A_t \supset B_t$  tendent toute deux à faire tendre  $r_t$  vers 0. Plus  $r_t$  tend vers 1 plus l'algorithme est précis. Si en moyenne le taux est relativement élevé et que les variations de  $r_t$  au cours du temps sont limitées, l'algorithme est robuste. On parlera de *suivi partiel* quand seulement une partie

de l'objet est suivi, cette situation correspondant à un taux de suivi compris entre 0,4 et 0,6. Au delà de 0,7 le suivi peut être considéré comme globalement satisfaisant. Au delà de 0,9 le suivi est d'excellente qualité. Notons qu'au delà de cette valeur la précision de la mesure est limitée par la précision de la vérité terrain elle-même. Enfin en dessous de 0,4 le suivi est médiocre.

Étant donné, la difficulté d'obtenir une vérité-terrain, nous serons pas en mesure de fournir le taux de suivi pour chacune des séquences présentées dans les chapitres suivants.

## 2.7 Les principales difficultés en suivi d'objets

Les principales difficultés rencontrées en suivi d'objets dans les séquences vidéo sont inhérentes au caractère dynamique du support à traiter. Les événements se déroulant dans les séquences d'images sont de natures très variées et difficilement prévisibles. Une catégorie d'événements a néanmoins été identifiée par la communauté au fur et à mesure des expérimentations. Nous appellerons ces événements, des événements perturbateurs, et nous rappelons dans cette partie quels sont les événements perturbateurs les plus fréquents :

- les occultations,
- le fouillis,
- les mouvements particuliers,
- les changements d'illumination.

### 2.7.1 Les occultations

Une occultation se définit par le fait qu'un objet tiers vienne s'interposer entre l'objet suivi et l'objectif de la caméra, cachant ainsi une partie ou l'intégralité de l'objet suivi. La figure 2.4 est un exemple d'occultation dans une scène naturelle. Dans les séquences routières les occulta-

**Figure 2.4 :** Exemple d'occultation.

tions sont fréquentes. Une occultation est particulièrement pénalisante pour les algorithmes de suivi car elle supprime tout accès à l'information visuelle de l'objet suivi. Il est impossible de plus de connaître le déplacement de l'objet quand celui-ci a totalement disparu de l'image. Le comportement attendu d'un algorithme de suivi est alors de détecter l'occultation et de geler le suivi en attendant la réapparition de l'objet.

Pour reprendre le suivi après l'occultation, la plupart des auteurs font l'hypothèse que durant l'occultation, l'objet n'a pas parcouru une distance trop importante. Comme pour l'utilisation des positions précédentes de l'objet, cette hypothèse permet de rechercher l'objet dans une zone restreinte de l'image. Rechercher l'objet dans toute l'image est souvent beaucoup trop

couteux en temps de calcul.

La gestion des occultations fait souvent l'objet d'un module spécifique dans les algorithmes de suivi. Ces modules ne sont pas spécifiques aux procédures et relèvent plus d'une volonté des auteurs de modéliser les occultations et d'inclure ce modèle supplémentaire dans les procédures. Nguyen et Smeulders [Nguyen 04] détectent les occultations en seuillant l'erreur du filtre de Kalman dans un premier temps, puis en seuillant le nombre d'erreurs dans un second temps. Les occultations partielles et de courte durée (inférieures à 1s) sont gérées mais pas celles de trop longue durée. Les auteurs font remarquer que si une occultation dure trop longtemps des données non-relatives à l'objet sont petit à petit intégrées au modèle d'objet, d'où la difficulté de mettre à jour les modèles (*cf.* section 2.6.1).

Yin et Collins [Yin 08] remarquent très justement qu'en cas d'occultation, une recherche globale dans toute l'image sera bien plus efficace qu'une recherche locale autour de la dernière détection afin de retrouver l'objet. Comme nous l'avons évoqué auparavant, les méthodes de recherche globale sont extrêmement coûteuses en temps de calcul et rarement utilisées, c'est pourquoi les auteurs préfèrent une approche hybride avec une approche globale par recuits simulés adaptatifs, qui au bout de quelques itérations alimente une approche locale par Mean-Shift.

Dans le cas d'occultations, une approche de suivi multi-objets confère un avantage indéniable, car le fait de détecter plusieurs objets permet d'anticiper leurs interactions et conséquemment l'occultation d'un objet par un autre. Rasmussen et Hager [Rasmussen 01] montrent bien l'intérêt du JPDAF par rapport au PDAF en ce sens. Kamijo et Sakaushi [Kamijo 02] utilisent le champ d'étiquette de leur approche markovienne pour détecter les occultations, dont 95% sont correctement traitées. Toutefois leur approche utilise une caméra fixe et un apprentissage de la scène, ce qui ne cadre pas avec notre problématique.

Il est à noter que la représentation des objets par régions ainsi que la procédure de suivi par filtrage particulière présentent un avantage en cas d'occultation. Si, par exemple, une seule région de l'objet sort du champ de la caméra, le reste du modèle à base de régions reste valable pour maintenir le suivi sur la partie toujours visible. Le filtre particulière grâce à l'échantillonnage d'importance permet de tirer au hasard des positions éventuelles de l'objet dans une région plus importante pour un temps de calcul équivalent. Même si la position tirée au hasard ne permet que de retrouver une partie de l'objet, la convergence du filtre après quelques itérations assure la reprise complète du suivi.

### 2.7.2 Le fouillis

Une situation de fouillis se traduit par le fait que plusieurs objets aux propriétés visuelles similaires sont présents en même temps dans la scène. L'algorithme de suivi est alors confronté à une ambiguïté et "hésite" entre les deux objets. Plus la caractérisation de l'objet est précise, moins le risque de confusion est grand, mais plus le modèle est rigide et plus il aura de difficultés à surmonter d'autres événements comme les changements d'illumination. La figure 2.5 montre une situation de fouillis. On distinguera le cas du fouillis total où deux objets parfaitement identiques se retrouvent dans la scène. L'information temporelle et la cohérence de la trajectoire sont alors les seuls éléments qui puissent faire la différence. Les situations de fouillis sont fréquentes dans

**Figure 2.5 :** Exemple de situation de fouillis, où trois gardes nationaux grecques présentent des propriétés visuelles similaires.

les scènes routières, on peut en effet considérer que deux voitures de couleurs proches, mais de marques différentes forment une situation de fouillis.

Dans les algorithmes de suivi probabiliste, certains auteurs définissent une variable aléatoire contextuelle traduisant la présence d'un fouillis. Des marginales sont définies selon les cas, et permettent de prendre des décisions plus prudentes le cas échéant. CONDENSATION [Isard 98a] propose un tel mécanisme.

De même que pour les occultations, les approches multi-objets sont mieux adaptées aux situations de fouillis. Si deux objets très proches du modèle se trouvent dans la scène, il est naturellement préférable de lancer deux suivis propres à chaque objet au lieu d'un seul suivi constamment tiraillé entre deux solutions. Si les approches multi-objets permettent une meilleure localisation, l'identification des objets reste difficile. Il est en effet assez fréquent que les étiquettes nominatives des objets soient interverties lorsque les objets sont très proches.

L'approche de Kahn *et al.* [Kahn 05], évoquée précédemment en 2.6.2, permet un suivi en situation de fort fouillis (suivis de fourmis) mais sans occultations. Betke *et al.* [Betke 07] présentent un algorithme d'association de données permettant également de suivre un grand nombre de petites cibles (chauve-souris nord-américaines) même en cas d'occultation. Une approche probabiliste conditionnelle aux événements permet de gérer les apparitions, disparitions et situations de fouillis pour les objets.

### 2.7.3 Les mouvements particuliers

On désigne par mouvements particuliers les mouvements propres d'un objet pouvant amener à un changement radical de son aspect visuel. Parmi la gamme de mouvements posant classiquement un problème on trouve :

- la rotation d'axe, l'axe orthogonal au plan image. De nombreuses méthodes d'extraction sont sensibles à cette rotation.
- le changement d'échelle. Il s'agit d'une translation dont le vecteur mouvement correspondant n'appartient pas au plan image. L'objet change alors de "taille". De même certaines méthodes d'extraction supportent mal ce changement.

- le mouvement de forte amplitude. La plupart des algorithmes de suivi font l'hypothèse d'un mouvement d'amplitude limité afin de gagner en temps de calcul.
- une rotation d'axe, un axe différent de celui orthogonal à l'image. C'est le mouvement le plus pénalisant, car une face cachée de l'objet apparaît alors, tandis qu'une partie préalablement visible disparaît. Si ce type de mouvement se produit, il est impératif d'apprendre au préalable les différentes vues de l'objet, comme par exemple dans l'approche Eigentracking de Black et Jepson [Black 96]. On parle aussi parfois de changement de pose ou de vue de l'objet.

Dans les séquences routières, seul le changement d'échelle intervient. C'est par chance le plus facile à traiter des quatre évoqués. Sa prise en compte s'effectue principalement au niveau de la modélisation de l'objet. Au niveau de la procédure, il suffit que cette dernière laisse un degré de liberté suffisant pour changer d'échelle.

#### 2.7.4 Les changements d'illumination

Les changements d'illumination posent d'importants problèmes à toutes les méthodes de caractérisation reposant sur les valeurs prises par les pixels. La vision, de manière générale, est un triptyque : **source, objet, récepteur**. Dans le cas des changements d'illumination, la source lumineuse est modifiée ce qui entraîne *ipso facto* une modification de la perception de l'objet par le récepteur. Le changement d'illumination est un problème très présent dans les scènes routières, car d'une part elles sont tournées en extérieur et d'autre part car elles sont dynamiques. La figure 2.6 illustre à quel point l'aspect visuel d'un véhicule peut être modifié au cours du temps. De nombreux travaux ont été entrepris afin de lutter contre les changements d'illumi-

**Figure 2.6 :** Exemple d'apparences d'un véhicule à l'intérieur d'une séquence routière.

nation. L'approche principalement retenue consiste à modéliser les changements d'illumination (modèle de Blinn-Phong [Phong 75, Blinn 77]). Silveira et Malis [Silveira 07] utilise un tel modèle conjointement à un suivi de type minimisation d'erreur. L'approche est efficace sur des objets planaires en niveaux de gris. Bien qu'une attention particulière soit accordée à la complexité de la méthode, un étage supplémentaire de modélisation de l'illumination alourdit l'algorithme de suivi. Sans rejeter l'intérêt de ces méthodes, nous faisons le choix dans ce mémoire de ne pas développer de module explicitement dédié à l'invariance aux changements d'illumination. Ce problème sera traité en ayant recours à des informations peu sensibles à ces changements (informations de forme ou de mouvement). Ce choix se justifie par ailleurs du fait qu'à l'heure actuelle les techniques d'invariance ne permettent pas de couvrir l'étendue des changements rencontrés dans les séquences routières conjointement au dynamisme de ces mêmes séquences. Nous renvoyons le lecteur aux travaux de Muselet [Muselet 05] pour plus de précisions sur l'invariance aux changements d'illumination.

## 2.8 Conclusion et approche retenue

La gestion des événements perturbateurs constitue un des principaux verrous scientifiques à lever. L'approche que nous désirons mettre au point devra prendre en compte ces événements afin de répondre le plus largement possible au problème de suivi d'objets. La prise en compte de ces événements forme des critères de choix supplémentaires par rapport à ceux déjà présentés et relatifs à l'application sur des séquences routières. Le choix de l'approche ne peut se faire en se focalisant sur la procédure de suivi ou sur la modélisation de l'objet. Il faut s'assurer que le modèle et la procédure collaborent harmonieusement. Nous utilisons donc les conclusions et réflexions des deux premiers chapitres pour proposer notre solution.

Nous avons déjà justifié du choix d'une représentation de l'objet par boîte englobante. Il reste donc à choisir un mode de caractérisation de l'objet. Le tableau 2.1 résume les avantages et limites des trois grandes catégories d'extracteurs présentés dans le chapitre précédent. Les

**Tableau 2.1** : Comparaison des extracteurs d'information.

Critères	Extracteurs texture/couleur	Extracteurs de formes	Extracteurs de mouvement
occultation	sensible	sensible	sensible
Fouillis	sensible (peu fréquent)	sensible (fréquent)	sensible (très fréquent)
Mouvements particuliers	potentiellement sensible	potentiellement sensible	potentiellement sensible
Changements d'illumination	très sensible	quasi-invariant	peu sensible
Intérêt pour les séquences routières	permet d'identifier un véhicule en particulier	permet d'identifier une catégorie de véhicule	permet de séparer le fond de la scène des véhicules

occultations endommageant directement les observations, tous les extracteurs y sont sensibles. C'est d'ailleurs pour cela que les auteurs traitent les occultations au niveau procédural plutôt qu'au niveau du modèle.

La même remarque est valable pour le fouillis, mais dans le cadre des séquences routières cet événement n'a pas une même fréquence d'apparition pour tous les types d'extracteurs. En effet les extracteurs de mouvement vont détecter tous les véhicules en mouvement, il reste donc une imprécision en sortie de ces extracteurs par rapport au choix parmi les véhicules en mouvement de celui qui est réellement suivi. Il en va de même, mais dans une moindre mesure tout de même, pour les extracteurs de formes, car les formes des véhicules sont relativement semblables. Les extracteurs texture-couleur ont en ce sens un avantage, car il est moins fréquent d'avoir deux véhicules de couleurs et de textures totalement identiques.

Concernant les mouvements particuliers, il n'y a pas de catégorie d'extracteurs qui se distingue par rapport aux autres. La gestion des mouvements particuliers doit être prise en compte au niveau de la conception de l'extracteur. A ce titre, chacune des catégories citées possède des extracteurs robustes aux divers mouvements. Comme évoqué précédemment, dans notre cadre applicatif, seule l'invariance aux changements d'échelle est impérative.

Les changements d'illumination sont clairement beaucoup plus dommageables pour les extracteurs texture-couleur qui reposent beaucoup sur la valeur des pixels. A l'inverse, ces mêmes extracteurs texture-couleur sont plus efficaces pour identifier un véhicule particulier dans les

scènes routières. Cette remarque rejoint d'ailleurs celle sur le fouillis.

En résumé, nous voyons qu'il n'y a pas une catégorie d'extracteurs qui puisse allier tous les avantages évoqués. A ce titre, et comme une majorité d'auteurs, nous soutenons dans ce mémoire qu'il est impératif de faire collaborer ces différentes catégories d'extracteurs afin de pouvoir combiner leurs avantages respectifs. Il s'agit donc d'opérer une fusion de données en sortie des extracteurs. Pour reprendre la terminologie du domaine de la fusion de données, les extracteurs correspondent à des sources d'information. Nous confondrons ces deux termes dans la suite du document.

Pour obtenir une fusion efficace et une diversité dans l'information, il faut au moins un extracteur par catégorie d'information. Ces extracteurs pourront se relayer en cas de défaillance. Les choix concrets des extracteurs employés sont très liés à l'application de suivi de face arrière du véhicule précédent. En particulier, nous avons retenu :

- comme extracteurs de texture-couleur : les matrices de cooccurrences pondérées et les densités de couleur
- comme extracteurs de forme : les cartes de symétrie et cartes de symétrie des contours
- comme extracteurs de mouvement : l'algorithme de Lefaix *et al.* [Lefaix 02]

Plus de justifications sur ces extracteurs en particulier sont disponibles dans le chapitre applicatif (chap 5).

En ce qui concerne la procédure de suivi retenue, le contexte impose un temps de calcul "raisonnable". Cela exclut d'emblée la plupart des approches ayant recours à une méthode d'optimisation à l'exception notable du Mean-Shift. Il en va de même pour les méthodes nécessitant un apprentissage élaboré, car nous n'avons pas de bases de données suffisantes à cet effet, et de plus le contexte des séquences routières donne des séquences disparates, qui rendent l'apprentissage complexe. En outre, comme il a été choisi d'opérer une fusion de données pour caractériser l'objet suivi, il faut que la procédure soit compatible avec un processus de fusion. A ce titre, les approches probabilistes ont un avantage, car il est assez aisé d'étendre les modèles aux multi-sources. Enfin, parmi les méthodes probabilistes, le filtrage particulaire permet de lutter plus facilement contre les occultations et offre plus de possibilités en terme de modélisation. En vertu de cette analyse, le filtrage particulaire a été retenu comme procédure de suivi pour nos travaux.

L'analyse de l'état de l'art nous pousse à présent à examiner les procédés de fusion pouvant être intégrés à un filtre particulaire. Le chapitre suivant s'attachera dans un premier temps à présenter l'ensemble de ces méthodes, puis le formalisme de la méthode retenue et enfin les outils de fusion d'informations répondant à nos exigences.

# Chapitre 3

## Processus de fusion de données applicable au filtrage particulaire

### Sommaire

---

<b>3.1</b>	<b>Rappel sur les filtres particulaires . . . . .</b>	<b>68</b>
3.1.1	Principe de Monte Carlo . . . . .	68
3.1.2	Échantillonnage d'importance . . . . .	69
3.1.3	Estimation récursive des poids . . . . .	70
3.1.4	Problème de dégénérescence . . . . .	71
3.1.5	Autres éléments sur les filtres particulaires . . . . .	75
<b>3.2</b>	<b>Filtre particulaire adapté au problème du suivi . . . . .</b>	<b>76</b>
<b>3.3</b>	<b>Processus et stratégies de fusion dans les filtres particulaires . . . . .</b>	<b>78</b>
3.3.1	Approches existantes . . . . .	79
3.3.2	Méthodes à apprentissage . . . . .	82
3.3.3	Théorie des possibilités . . . . .	83
3.3.4	Théories des fonctions de croyance . . . . .	87
3.3.5	Justification du choix des fonctions de croyance comme méthode de fusion	88
<b>3.4</b>	<b>Formalisme des fonctions de croyance . . . . .</b>	<b>90</b>
3.4.1	Masse de croyance élémentaire . . . . .	90
3.4.2	Le principe du minimum d'information . . . . .	92
3.4.3	Règles de combinaison . . . . .	92
3.4.4	Autres représentations des croyances . . . . .	94
3.4.5	Notion de distinction . . . . .	95
3.4.6	Opération d'affaiblissement . . . . .	95
3.4.7	Transformation pignistique . . . . .	96
3.4.8	Autres processus crédibilistes . . . . .	97
<b>3.5</b>	<b>Approche proposée . . . . .</b>	<b>97</b>
3.5.1	Définition du cadre de discernement $\Omega$ . . . . .	97
3.5.2	Calcul des masses de croyance pour chaque source . . . . .	100
3.5.3	Fusion des FDM et intégration de la masse fusionnée au filtre particulaire	104
3.5.4	Tests de suivi génériques . . . . .	105
<b>3.6</b>	<b>Conclusion . . . . .</b>	<b>119</b>

---

Nous avons conclu suite aux chapitres précédents à la nécessité de combiner différents extracteurs d'information afin d'obtenir un modèle d'objet robuste. Nous avons également choisi d'utiliser le filtrage particulaire, ce qui implique d'utiliser une technique de fusion qui soit compatible avec la théorie des filtres particulaires. Dans ce chapitre, nous commencerons par présenter quelques rappels sur cette théorie. Pour plus de clarté, nous expliciterons ensuite comment le filtre particulaire sera exploité dans le cadre du suivi d'objets. Nous serons alors en mesure d'observer à quelles étapes du filtre il est possible d'opérer une fusion d'informations. Les travaux existants dans ce domaine seront présentés, afin de déterminer leurs limites. L'analyse de ces méthodes nous conduira à opter pour une fusion par fonctions de croyance. Nous présenterons donc également les fondamentaux de la théorie des fonctions de croyance, avant de finalement proposer un algorithme complet. Cet algorithme reposera sur la base d'un filtre particulaire tout en opérant la fusion des données fournies par des extracteurs de texture-couleur, forme et mouvement.

### 3.1 Rappel sur les filtres particulaires

En traitement de l'information, les filtres particulaires se sont imposés progressivement comme la technique la plus populaire en terme de traitement non-linéaire du signal. Les articles fondateurs sont notamment l'œuvre de Gordon *et al.* [Gordon 93] et Kitagawa [Kitagawa 96], mais la technique de l'échantillonnage d'importance, cœur des filtres particulaires, a commencé à être exploitée en physique statistique dès les années 50 [Hammersle 54, Rosenbluth 55]. A cause d'un phénomène de dégénérescence, l'algorithme est resté marginal pendant de longues années. Le concept de rééchantillonnage introduit par Gordon a permis de contourner le problème et a ouvert de nouvelles voies à de très nombreux développements.

Comme vu en 2.4.1, le filtre particulaire propose une solution sous-optimale à un problème de filtrage non-linéaire. Ce problème s'écrit sous la forme de deux équations, que nous redonnons ici pour plus de commodité :

$$X_t = f_t(X_{t-1}, b_t) \text{ équation de transition} \quad (3.1)$$

$$Y_t = g_t(X_t, c_t) \text{ équation d'observation} \quad (3.2)$$

avec  $f_t()$  et  $g_t()$  deux fonctions *a priori* quelconques et  $b_t$  et  $c_t$  deux bruits *a priori* quelconques.  $X_t$  est le vecteur stochastique qu'il faut estimer sachant les estimations passées  $X_{0:t-1}$  et les observations présentes et passées  $Y_{0:t}$ . Ce vecteur est appelé vecteur d'état. En suivi, il code la réponse au problème c'est-à-dire la position actuelle de l'objet dans la séquence d'images.

L'estimation  $\hat{X}_t$  peut s'obtenir via l'expression de la densité  $p(X_{0:t}|Y_{0:t})$ , appelée densité de filtrage. La connaissance de cette densité revient à une connaissance totale du problème. La densité de filtrage n'étant en général pas connue, plusieurs procédés sont utilisés afin de l'estimer elle aussi. Ces procédés sont entre autres le principe de Monte Carlo et l'échantillonnage d'importance.

#### 3.1.1 Principe de Monte Carlo

Le principe des méthodes de Monte Carlo (MC) propose d'estimer une densité de probabilité  $p(X)$  caractérisant la loi suivie par la variable  $X$  définie sur  $D_X$ , domaine de la variable  $X$ , sachant un ensemble de réalisations  $X^{(i)}$ , ou échantillons, tirées selon la densité  $p(X)$ . Dans le

cadre des filtres particulaires, ces échantillons sont souvent appelés particules. L'approximation discrète suivante est alors obtenue :

$$\hat{P}(X \in dX) = \frac{1}{N} \sum_{i=1}^N \delta_{X^{(i)}}(X) \quad (3.3)$$

où  $\delta_{X^{(i)}}$  est la distribution de Dirac centrée en  $X^{(i)}$  et  $dX$  le différentiel de  $X$ . Cette estimation est un cas particulier de la loi forte des grands nombres, qui stipule que quelle que soit  $f$   $p(\cdot)$ -intégrable, la relation suivante est vérifiée :

$$\hat{Int}_N(f(X)) = \frac{1}{N} \sum_{i=1}^N f(X^{(i)}) \xrightarrow[N \rightarrow \infty]{p.s.} Int(f(X)) = \int_{R_X} f(X) p(X) dX \quad (3.4)$$

où  $Int$  symbolise l'opérateur d'intégration. Le principe de Monte Carlo offre une approximation non biaisée en convergence presque sûre de la densité de filtrage. Pour plus de détail sur le comportement asymptotique de ces estimateurs, nous invitons le lecteur à se reporter aux travaux de Doucet [Doucet 98].

Tout le problème est alors de pouvoir échantillonner selon la distribution  $p(X)$  sans pour autant avoir accès directement à cette dernière. Plusieurs méthodes permettant d'atteindre cet objectif existent, comme l'algorithme d'acceptation/rejet par exemple. Dans le cadre des filtres particulaires, une autre technique plus évoluée, appelée échantillonnage d'importance, est utilisée.

### 3.1.2 Échantillonnage d'importance

Supposons connue une distribution  $q$  liée à  $p$  de la manière suivante :  $p > 0 \Rightarrow q > 0$ . Le calcul suivant est alors valable :

$$Int_N(f(X)) = \int_{R_X} f(X) \frac{p(X)}{q(X)} q(X) dX \quad (3.5)$$

Supposons de plus qu'il soit possible d'échantillonner  $N$  réalisations  $X^{(i)}$  selon  $q$ , l'estimateur de Monte Carlo peut alors s'écrire :

$$\hat{Int}_N(f(X)) = \frac{1}{N} \sum_{i=1}^N f(X^{(i)}) \frac{p(X^{(i)})}{q(X^{(i)})} \quad (3.6)$$

On pose alors  $w^{(i)} = \frac{p(X^{(i)})}{q(X^{(i)})}$ . Les  $w^{(i)}$  sont appelés poids d'importance. En effet, plus  $w^{(i)}$  est grand, plus  $X^{(i)}$  "suit" la loi  $p$ , et donc plus il est "important".  $q(\cdot)$  est logiquement appelée densité d'importance. Cet estimateur est également non biaisé et possède une variance moindre que l'algorithme acceptation/rejet. De plus, les hypothèses sont beaucoup moins contraignantes. Notons par ailleurs que les poids sont normalisés selon la formule  $\tilde{w}^{(i)} = \frac{w^{(i)}}{\sum_{j=1}^N w^{(j)}}$ , et l'estimation devient  $\hat{Int}_N(f(X)) = \sum_{i=1}^N f(X^{(i)}) \tilde{w}^{(i)}$ .

En appliquant ce principe au cas particulier de la densité de filtrage, il vient :

$$\hat{p}(X_{0:t}|Y_{0:t}) = \frac{1}{N} \sum_{i=1}^N w_t^{(i)} \delta_{X^{(i)}}(X) \quad (3.7)$$

$$w_t^{(i)} = \frac{p(X_{0:t}^{(i)}|Y_{0:t})}{q(X_{0:t}^{(i)}|Y_{0:t})} \quad (3.8)$$

Le problème est maintenant reporté sur l'estimation des poids. Cette estimation est possible en utilisant la théorie des probabilités, comme le montre la sous-section suivante.

### 3.1.3 Estimation récursive des poids

La méthode de l'échantillonnage d'importance peut être mise en œuvre de manière séquentielle pour chaque arrivée d'une nouvelle observation  $Y_t$ . En partant de la densité de filtrage, la règle de Bayes permet d'écrire :

$$\begin{aligned}
p(X_{0:t}|Y_{0:t}) &= p(X_{0:t}|Y_{0:t-1}, Y_t) \\
p(X_{0:t}|Y_{0:t}) &= \frac{p(X_{0:t}, Y_t|Y_{0:t-1})}{p(Y_t|Y_{0:t-1})} \\
p(X_{0:t}|Y_{0:t}) &= \frac{p(Y_t|X_{0:t}, Y_{0:t-1}) p(X_{0:t}|Y_{0:t-1})}{p(Y_t|Y_{0:t-1})} \\
p(X_{0:t}|Y_{0:t}) &= \frac{p(Y_t|X_{0:t}, Y_{0:t-1}) p(X_t|X_{0:t-1}, Y_{0:t-1}) p(X_{0:t-1}|Y_{0:t-1})}{p(Y_t|Y_{0:t-1})} \quad (3.9)
\end{aligned}$$

En se plaçant dans le cadre markovien (*cf.* Figure 2.1), on peut alors émettre les hypothèses simplificatrices suivantes :

$$p(X_t|X_{0:t-1}, Y_{0:t-1}) = p(X_t|X_{t-1}) \quad (3.10)$$

$$p(Y_t|X_{0:t}, Y_{0:t-1}) = p(Y_t|X_t) \quad (3.11)$$

L'expression de la densité de filtrage devient :

$$\begin{aligned}
p(X_{0:t}|Y_{0:t}) &= \frac{p(Y_t|X_t) p(X_t|X_{t-1}) p(X_{0:t-1}|Y_{0:t-1})}{p(Y_t|Y_{0:t-1})} \\
p(X_{0:t}|Y_{0:t}) &\propto p(Y_t|X_t) p(X_t|X_{t-1}) p(X_{0:t-1}|Y_{0:t-1}) \quad (3.12)
\end{aligned}$$

En réinjectant ce résultat dans la formule de calcul des poids on obtient :

$$w_t^{(i)} = \frac{p(X_{0:t}^{(i)}|Y_{0:t})}{q(X_{0:t}^{(i)}|Y_{0:t})} \quad (3.13)$$

$$w_t^{(i)} \propto \frac{p(Y_t|X_t^{(i)}) p(X_t^{(i)}|X_{t-1}) p(X_{0:t-1}^{(i)}|Y_{0:t-1})}{q(X_{0:t}^{(i)}|Y_{0:t})} \quad (3.14)$$

En supposant que  $q$  est choisie de sorte à respecter  $q(X_{0:t}^{(i)}|Y_{0:t}) = q(X_t^{(i)}|X_{0:t-1}^{(i)}, Y_{0:t}) q(X_{0:t-1}^{(i)}|Y_{0:t-1})$ , on a finalement :

$$w_t^{(i)} \propto \frac{p(Y_t|X_t^{(i)}) p(X_t^{(i)}|X_{t-1}) p(X_{0:t-1}^{(i)}|Y_{0:t-1})}{q(X_t^{(i)}|X_{0:t-1}^{(i)}, Y_{0:t}) q(X_{0:t-1}^{(i)}|Y_{0:t-1})} \quad (3.15)$$

$$w_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(Y_t|X_t^{(i)}) p(X_t^{(i)}|X_{t-1})}{q(X_t^{(i)}|X_{0:t-1}^{(i)}, Y_{0:t})} \quad (3.16)$$

Ce résultat définit la procédure de base des filtres particulières. Cette procédure est appelée Sequential Importance Sampling (SIS) :

---

 Algorithme : Filtre Particulaire de base/ SIS
 

---

**Initialisation**,  $t = 0$  :

 Échantillonner  $N$  particules selon une distribution  $q(X_0)$  et poser  $\forall i w_0^{(i)} = \frac{1}{N}$ 
**pour**  $t \geq 1$  **faire**
**pour**  $i = 1$  à  $N$  **faire**

 Échantillonner les particules selon  $X_t^{(i)} \sim q(X_t | X_{t-1}^{(i)}, Y_t)$ 

 Mettre à jour les poids selon  $w_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(Y_t | X_t^{(i)}) p(X_t^{(i)} | X_{t-1}^{(i)})}{q(X_t^{(i)} | X_{t-1}^{(i)}, Y_t)}$ 

 Normaliser les poids selon  $\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^N w_t^{(j)}}$ 
**fin pour**

 Estimer l'état :  $\hat{p}(X_t | Y_{0:t}) = \sum_{i=1}^N \delta_{X_t^{(i)}}(X_t) \tilde{w}_t^{(i)}$ 
**fin pour**
**Fin**


---

Algorithme 1: SIS.

### 3.1.4 Problème de dégénérescence

Dans la pratique, l'algorithme SIS est sujet à un phénomène de dégénérescence important au sens où après quelques itérations, une seule particule possèdera un poids proche de 1, tandis que les  $N - 1$  autres auront un poids proche de 0. Une partie de la capacité de calcul est alors gâchée, et l'intérêt du filtre qui consiste à explorer l'espace d'état est perdu. Dans [Doucet 98] Doucet montre que la variance des poids ne peut qu'augmenter avec le temps, et que donc conséquemment le problème ne peut être contourné. Une mesure de la dégénérescence est proposée dans [Liu 98] par Liu et Chen. Cette mesure appelée taille efficace du  $N$ -échantillon s'écrit :

$$N_{eff} = \frac{N}{1 + \text{var}(w_t^{*(i)})} \quad (3.17)$$

avec  $w_t^{*(i)} = p(X_t^{(i)} | Y_{0:t}) / q(X_t^{(i)} | X_{t-1}^{(i)}, Y_t)$ . Une approximation est disponible sous la forme :

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^N (\tilde{w}_t^{(i)})^2} \quad (3.18)$$

Plusieurs méthodes pour pallier le phénomène de dégénérescence ont été mises au point. Nous n'évoquons ici que brièvement les deux plus importantes, d'autres méthodes sont explicitées dans [Arumpalam 02].

#### Choix de la densité d'importance

Il est clair que le choix de la densité d'importance a une influence directe sur la variance des poids. En prenant trivialement  $q(X_{0:t} | Y_{0:t}) = p(X_{0:t} | Y_{0:t})$ , on constate que ce choix est le meilleur possible, car les poids non normalisés valent alors en moyenne 1 et ont une variance nulle. Mais ce choix n'a pas vraiment de sens puisqu'il revient à ne pas se servir d'une densité d'importance. Hormis ce cas de figure, il a été montré que le choix optimal revient à prendre  $q$  telle que  $q(X_t | X_{0:t-1}, Y_{0:t}) = p(X_t | X_{t-1}^{(i)}, Y_t)$ . Une preuve est disponible dans les travaux de thèse de Doucet [Doucet 97].

Malheureusement, il est en pratique peu évident d'échantillonner selon  $p(\cdot | X_{t-1}^{(i)}, Y_t)$ . Comme nous l'évoquions dans le chapitre précédent (cf. 2.4.1), Arnaud [Arnaud 04b] a néanmoins réussi un tel échantillonnage dans un contexte particulier : équation d'observation linéaire et bruits gaussiens. Les auteurs montrent que la densité optimale est alors une gaussienne dont les paramètres peuvent être directement obtenus en connaissant les covariances des bruits, la fonction de transition  $f_t$  et la matrice reliant l'état aux observations. L'extension à une vraisemblance multi-modale est possible.

Dans le cas général, une solution approximative peut être fournie par des méthodes MC. Puisque  $p(Y_t | X_t) = \frac{p(X_t | X_t^{(i)}, Y_t)}{p(X_t | X_{t-1}^{(i)})}$  et qu'il est possible d'échantillonner selon  $p(X_t | X_{t-1}^{(i)})$ , alors si un majorant  $M_t$  de  $p(Y_t | X_t)$  est connu, il est possible d'utiliser l'algorithme d'acceptation/rejet pour obtenir des échantillons  $X_t \sim p(X_t | X_{t-1}^{(i)}, Y_t)$ . D'autre part, pour évaluer  $p(Y_t | X_t^{(i)})$ , en échantillonnant  $N'$  réalisations  $X_t^{(i),(j)} \sim p(X_t | X_{t-1}^{(i)})$ , une approximation MC donne  $\hat{p}(Y_t | X_{t-1}^{(i)}) = \frac{1}{N'} \sum_{j=1}^{N'} p(Y_t | X_t^{(i),(j)})$ . Ces méthodes présentent toutefois dans la pratique peu d'intérêt, car elles possèdent un coût de calcul élevé, et l'approximation n'est valable que pour  $N'$  très grand.

Une autre solution consiste à utiliser un développement de Taylor à l'ordre 2 de  $l(X_t) = \ln p(X_t | X_{t-1}^{(i)}, Y_t)$ . Doucet [Doucet 98] montre qu'autour d'un point quelconque  $X_t$ , la densité  $p(X_t | X_{t-1}, Y_t)$  est approximable par une gaussienne plus un terme négligeable devant  $X_t^2$ , d'où l'idée de choisir également une gaussienne pour la densité  $q(\cdot)$ . Les deux premiers moments de la distribution sont accessibles par les dérivées de  $l(\cdot)$ .

Pour contourner ces difficultés, il est possible de faire le choix d'une fonction d'importance sous-optimale. Il s'agit de la densité d'importance *a priori* :  $q(X_{0:t} | Y_{0:t}) = p(X_t | X_{t-1}^{(i)})$ . Dans ces conditions l'équation de mise à jour des poids prends une forme très simple :  $w_t^{(i)} = w_{t-1}^{(i)} p(Y_t | X_t^{(i)})$ . La modélisation de  $p(Y_t | X_t^{(i)})$  n'est pas un obstacle au sens où la connaissance de l'équation d'observation nous renseigne sur cette question.

## Rééchantillonnage

Pour lutter contre le phénomène de dégénérescence, l'idée la plus efficace consiste à chaque itération à éliminer les particules de poids faibles, et démultiplier celles de poids forts. Cette technique est connue sous le nom de rééchantillonnage. Elle répond directement au problème en forçant la répartition des poids des particules.

Si on choisit de rééchantillonner systématiquement, le filtre particulière est alors communément appelé Sampling Importance Resampling (SIR) et prend la forme de l'algorithme 2.

Si le rééchantillonnage est conditionné par un seuil sur la mesure  $\hat{N}_{eff}$ , on obtient l'algorithme générique du filtre particulière (cf. algorithme 3).

Un autre critère de redistribution existe [Pham 01]. Il est basé sur l'entropie des particules  $E = \frac{1}{N} \sum_{i=1}^N w_t^{(i)} \ln w_t^{(i)}$  et s'utilise lui aussi par seuillage. L'algorithme du filtre particulière dit "générique" est le plus répandu dans la littérature.

---

 Algorithme : Filtre SIR
 

---

**Initialisation**,  $t = 0$  :Échantillonner  $N$  particules selon une distribution  $q(X_0)$  et poser  $\forall i w_0^{(i)} = \frac{1}{N}$ **pour**  $t \geq 1$  **faire****pour**  $i = 1$  à  $N$  **faire**Échantillonner les particules selon  $X_t^{(i)} \sim q(X_t^{(i)} | X_{t-1}^{(i)}, Y_t)$ Mettre à jour les poids selon  $w_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(Y_t | X_t^{(i)}) p(X_t^{(i)} | X_{t-1}^{(i)})}{q(X_t^{(i)} | X_{t-1}^{(i)}, Y_t)}$ Normaliser les poids selon  $\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^N w_t^{(j)}}$ **fin pour**Rééchantillonner en tirant  $N$  nouvelles particules  $\tilde{X}_t^{(j)}$  parmi les  $X_t^{(i)}$  selon une probabilité  $w_t^{(i)}$ , et leur affecter les poids  $\tilde{w}_t^{(j)} = \frac{1}{N}$ Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t | Y_{0:t}) = \sum_{j=1}^N \delta_{X_t^{(j)}}(X_t) \tilde{w}_t^{(j)}$ **fin pour****Fin**

Algorithme 2: SIR.

---

 Algorithme : Filtre particulaire générique
 

---

**Initialisation**,  $t = 0$  :Échantillonner  $N$  particules selon une distribution  $q(X_0)$  et poser  $\forall i w_0^{(i)} = \frac{1}{N}$ **pour**  $t \geq 1$  **faire****pour**  $i = 1$  à  $N$  **faire**Échantillonner les particules selon  $X_t^{(i)} \sim q(X_t^{(i)} | X_{t-1}^{(i)}, Y_t)$ Mettre à jour les poids selon  $w_t^{(i)} \propto w_{t-1}^{(i)} \frac{p(Y_t | X_t^{(i)}) p(X_t^{(i)} | X_{t-1}^{(i)})}{q(X_t^{(i)} | X_{t-1}^{(i)}, Y_t)}$ Normaliser les poids selon  $\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^N w_t^{(j)}}$ **fin pour****si**  $\hat{N}_{eff} > \text{seuil}$  **alors**Rééchantillonner en tirant  $N$  nouvelles particules  $\tilde{X}_t^{(j)}$  parmi les  $X_t^{(i)}$  selon une probabilité  $\tilde{w}_t^{(i)}$ , et leur affecter les poids  $\tilde{w}_t^{(j)} = \frac{1}{N}$ Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t | Y_{0:t}) = \sum_{j=1}^N \delta_{\tilde{X}_t^{(j)}}(X_t) \frac{1}{N}$ **sinon**Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t | Y_{0:t}) = \sum_{i=1}^N \delta_{X_t^{(i)}}(X_t) \tilde{w}_t^{(i)}$ **fin si****fin pour****Fin**

Algorithme 3: Filtre particulaire générique.

Plusieurs techniques spécifiques de rééchantillonnage ont été mises au point par nombre d'auteurs. Tout d'abord la méthode la plus directe, le rééchantillonnage multinomial. En tirant avec remise  $N$  nouvelles particules avec une probabilité  $w_t^{(i)}$ , on obtient  $N$  nouvelles particules indépendamment et identiquement distribuées selon  $p(X_t | Y_{0:t})$ , à qui un poids de  $1/N$  est donc conféré. Les nombres des descendants des particules suivent conjointement une loi multinomiale

de paramètres  $N$  et  $(w_t^{(i)})_{i=1}^N$ , d'où le nom donné à la méthode. Dans la pratique, l'algorithme 4 est utilisé.

---

Algorithme : Rééchantillonnage multinomial

---

**Initialisation : calcul des poids cumulés**,  $c_1 = 0$  :

**pour**  $i = 2$  à  $N$  **faire**

Faire  $c_i = c_{i-1} + w_t^{(i)}$

**fin pour**

Prendre un point de départ  $u_1$  selon la loi uniforme  $\mathcal{U}[0, 1/N]$

**pour**  $j = 1$  à  $N$  **faire**

Faire  $u_j = u_1 + 1/N(j-1)$

**tantque**  $c_i < u_j$  **faire**

Faire  $i = i + 1$

**fin tantque**

Mettre à jour les particules  $X_t^{(j)new} = X_t^{(i)old}$  et les poids selon  $w_t^{(j)new} = 1/N$

**fin pour**

**Fin**

---

Algorithme 4: Rééchantillonnage multinomial.

Une autre méthode éprouvée est celle de l'échantillonnage résiduel [Liu 98]. Elle consiste dans un premier temps à créer pour chaque particule mère  $X_t^{(i)}$   $\lfloor Nw_t^{(i)} \rfloor$  descendants. Une fois cette étape terminée,  $\bar{N} = \sum_{i=1}^N \lfloor Nw_t^{(i)} \rfloor < N$  nouvelles particules sont obtenues ( $\lfloor x \rfloor$  désigne la partie entière de  $x$ ). Il faut donc encore en générer  $R = N - \bar{N}$  pour garder un nombre constant de particules au fur et à mesure des itérations. A cet effet, la distribution résiduelle  $\frac{1}{R} \sum_{i=1}^N \lfloor Nw_t^{(i)} \rfloor \delta_{X_t^{(i)}}(X)$  est utilisée ( $\lceil x \rceil$  désigne la partie fractionnaire de  $x$ ).

Une dernière méthode existe : la méthode de branchement de Bernouilli [Crisan 99]. Elle est quasiment similaire à la précédente, à ceci près que les  $R$  particules manquantes ne sont pas toutes systématiquement obtenues. Le nombre de particules n'est donc pas constant. Un seul descendant supplémentaire est attribué à la particule  $X_t^{(i)}$  avec la probabilité résiduelle  $\lceil Nw_t^{(i)} \rceil$ .

Des trois méthodes citées, la première est simple à mettre en œuvre et possède la plus petite variance, elle est donc la plus couramment employée. En outre, quelle que soit la méthode employée, le rééchantillonnage comporte un certain nombre de défauts. Tout d'abord, les trajectoires suivies par les particules au cours du temps ne sont plus statistiquement indépendantes, or cette indépendance est nécessaire pour la preuve de convergence simple du filtre.

De plus, un phénomène d'appauvrissement de la diversité des particules est inévitable. Ce phénomène sera d'autant plus flagrant pour des systèmes à faible bruit. En conséquence, l'algorithme 3 est quasiment toujours préféré aux autres car il permet un compromis entre rééchantillonnage systématique et non-rééchantillonnage.

Il est possible de lutter contre l'appauvrissement des particules principalement par l'ajout de deux méthodes : une redistribution guidée par noyaux et une étape MCMC. Se reporter à [Arumpalam 02] pour la mise en œuvre de ces méthodes.

### 3.1.5 Autres éléments sur les filtres particuliers

Les paragraphes précédents présentent les éléments-clés constitutifs du filtre particulaire sous sa forme générique. Il existe toutefois de nombreuses autres formes de filtres particuliers. Nous évoquons déjà certaines implémentations de ces évolutions du filtre particulaire dans le cadre du suivi d'objets en 2.4.1. Elles sont reprises en partie dans cette sous-section. Les évolutions principales du filtre particulaire sont :

- le filtre particulaire rao-blackwellisé [Doucet 98, Doucet 00, Legland 03],
- le filtre particulaire sans parfum (unscented particle filter) [Van der Merwe 00],
- le filtre particulaire auxiliaire (FPA) [Pitt 99],
- le filtre particulaire guidé par les observations [Torma 04].

Les filtres rao-blackwellisés s'appliquent si une partie de l'état peut être déterminée par des méthodes optimales (filtre de Kalman par exemple). Il convient de réécrire les équations de transition et d'observation en séparant l'état en deux parties. Selon les différents systèmes alors obtenus, et en utilisant des marginales et des densités conditionnelles, il est possible de remonter à l'expression de la densité de filtrage. Le théorème de Rao-Blackwell montre que la variance en moyenne est réduite. Néanmoins la rao-blackwellisation n'est pas applicable dans le cas général.

Le filtre particulaire sans parfum approxime la densité d'importance optimale par une gaussienne. Pour chaque particule, un filtre de Kalman sans parfum (*cf.* annexe B) permet d'obtenir les paramètres de la gaussienne. Le même procédé est valable avec un filtre de Kalman étendu. L'estimation des paramètres utilise les observations, qui sont donc prises en compte avant la propagation.

Globalement, c'est la même idée qui est présente dans les FPAs et les filtres particuliers guidés par les observations. Le FPA fait appel à une variable aléatoire dite auxiliaire  $\tilde{X}_t^{(i)}$  qui permet un échantillonnage en deux temps. Si la fonction d'importance est optimale, l'équation de mise à jour est équivalente à :

$$w_t^{(i)} \propto w_{t-1}^{(i)} p(Y_t | X_{t-1}) \quad (3.19)$$

La mise à jour des poids peut alors être effectuée avant l'échantillonnage, le problème étant que la densité  $p(Y_t | X_{t-1})$  est difficile à évaluer. Les auteurs proposent alors de l'approcher par une densité  $p(Y_t | \mu_t^{(i)})$  où  $\mu_t^{(i)}$  caractérise la loi conditionnelle  $p(X_t | X_{t-1}^{(i)})$ . En général  $\mu_t^{(i)}$  est la moyenne de la variable  $X_t | X_{t-1}^{(i)}$ . Des poids auxiliaires sont alors définis :  $\lambda_t^{(i)} = w_{t-1}^{(i)} p(Y_t | \mu_t^{(i)})$ . Une première étape d'échantillonnage permet de rééchantillonner les particules issues de l'itération précédente et associées aux poids auxiliaires :  $\{X_{t-1}^{(i)}, \lambda_t^{(i)}\}$  pour obtenir les échantillons auxiliaires équirépartis :  $\{\tilde{X}_{t-1}^{(i)}, \frac{1}{N}\}$ . C'est à ce moment là que le FPA prend tout son intérêt car les particules sont redistribuées d'emblée dans des zones de l'image à plus forte vraisemblance.

La seconde partie de l'échantillonnage est plus classique. Elle permet l'obtention des particules à l'état présent via l'équation de transition :  $X_t^{(i)} \sim p(X_t | \tilde{X}_{t-1}^{(i)})$ . Au final il ne reste plus qu'à mettre les poids à jour :  $w_t^{(i)} \propto \frac{p(Y_t | X_{t-1})}{p(Y_t | \tilde{X}_{t-1})}$ .

Torma et Szepesvari [Torma 04] ont proposé un échantillonnage de la vraisemblance locale. Leur algorithme suit lui la marche d'un SIR classique avec densité d'importance *a priori*, mais

ils y incorporent une étape supplémentaire juste après la propagation des particules. D'autres particules temporaires sont alors tirées selon :  $Z_t^{(k)} \sim \frac{1}{\alpha_t^{(k)}} p(Y_t|X_t) g(X_t^{(k)} - X_t)$ , où  $\alpha_t^{(k)}$  est une normalisation et  $g(\cdot)$  une fonction à choisir. Les poids sont ensuite mis à jour selon  $w_t^{(k)} = \alpha_t^{(k)} p(Z_t^{(k)}|X_{t-1}^{(k)}) / p(X_t^{(k)}|X_{t-1}^{(k)})$ . La fonction  $g(\cdot)$  permet de déplacer les particules vers les modes de la vraisemblance  $p(Y_t|X_t)$ . De plus, si  $g(\cdot)$  est définie sur un support compact, positive, et que  $\int g(x) dx = 1$ , alors la méthode proposée n'introduit pas de biais supplémentaire par rapport au SIR.

Les trois dernières méthodes évoquées sont applicables dans le cas général et d'une certaine manière reposent toutes sur le même principe : donner plus d'importance à l'attache aux données image plutôt qu'au modèle de propagation. Nous n'avons pas pour but de détailler plus en avant dans ce mémoire ces divers filtres. L'état de l'art des travaux de thèse d'Arnaud [Arnaud 04a] contient des explications supplémentaires à ce sujet.

S'il est prouvé que ces filtres apportent un gain en robustesse pour des applications de suivi [Brêthes 06], il n'est pas indispensable dans notre cadre de travail d'y avoir recours. En effet nous n'avons pas pour ambition de contribuer à mettre au point une meilleure technique de filtrage particulaire mais de trouver un moyen habile pour faire de la fusion de données dans un tel filtre. A la limite, il est même préférable de garder un filtre relativement simple pour que la qualité de nos résultats dépende principalement de la qualité de la fusion de données. Nous optons donc pour un filtre particulaire générique avec rééchantillonnage multinomial pour nos expériences et nos développements futurs.

### 3.2 Filtre particulaire adapté au problème du suivi

Maintenant que nous avons choisi quel filtre utiliser, il faut expliciter chacune des variables, des relations et des densités nécessaires au bon fonctionnement du filtre. L'implémentation présentée dans cette partie s'inspire largement de celle de Perez *et al.* [Perez 02]. Nous commençons par rappeler la définition de la variable  $X_t$  qui correspond aux paramètres nécessaires au positionnement de l'objet dans l'image  $I_t$  d'après la représentation d'objet que nous avons choisie (boîte englobante).

$$X_t = \begin{pmatrix} i_t \\ j_t \\ height_t \\ width_t \end{pmatrix} \quad (3.20)$$

où le centre de l'objet-cible se trouve en  $(i_t, j_t)$ , et  $(height_t, width_t)$  sont les dimensions de la boîte englobante de l'objet.

L'équation de transition précise le mode de transition de l'état précédent vers l'état présent. Dans notre cadre applicatif cette transition est régie par le mouvement de l'objet entre deux images consécutives. Ce mouvement peut être connu *a priori* dans un cadre de travail spécifique où l'objet et son contexte sont connus. Dans les séquences routières, il existe des conditions sur le déplacement d'un véhicule, mais il est impossible de le prévoir totalement (accélérations, freinage, changements de voie intempestifs). L'autre solution consiste à estimer ce mouvement à

l'aide des méthodes présentées en 1.5. Cela pose alors le problème de la fiabilité de la méthode d'estimation de mouvement choisie car, en cas d'erreur, la prédiction favorisera à tort certaines zones de l'image. Pour ne pas prendre de risque, l'attitude souvent adoptée par les auteurs consiste à donner suffisamment de liberté à la prédiction pour que seules les zones où il est très improbable de retrouver l'objet soit pénalisées. Ces zones sont constituées des régions de l'image trop éloignées de la position précédente  $X_t$ . Le modèle le moins contraignant est la marche aléatoire :

$$X_t = X_{t-1} + \begin{pmatrix} \mathcal{N}(0, \sigma_1) \\ \mathcal{N}(0, \sigma_2) \\ \mathcal{N}(0, \sigma_3) \\ \mathcal{N}(0, \sigma_4) \end{pmatrix} \quad (3.21)$$

La position actuelle est supposée proche de la précédente, aucune direction n'est privilégiée, et le rayon d'action de la prédiction est limité par les écarts-types des gaussiennes. Nous utiliserons un modèle légèrement plus complexe en incorporant une information de vitesse, garantissant une certaine inertie de l'objet, ou cohérence temporelle du mouvement :

$$X_t = a_1 X_{t-1} + a_2 (X_{t-1} - X_{t-2}) + \begin{pmatrix} \mathcal{N}(0, \sigma_1) \\ \mathcal{N}(0, \sigma_2) \\ \mathcal{N}(0, \sigma_3) \\ \mathcal{N}(0, \sigma_4) \end{pmatrix} \quad (3.22)$$

avec  $a_1$  et  $a_2$  deux paramètres classiques d'un modèle auto-régressif. Ces paramètres peuvent être évolutifs ; pour plus de simplicité, ils seront constants dans nos expériences. A  $X_{t-1}$  fixée, l'équation détermine la densité *a priori* ou densité de transition  $p(X_t|X_{t-1})$  qui est donc gaussienne.

Concernant les observations  $Y_t$ , nous les définissons comme l'ensemble des caractéristiques extraites de l'image  $I_t$ . Les méthodes d'extraction utilisées sont celles définies par le modèle d'objet (partie caractérisation). Toutes ces caractéristiques sont extraites de la même "expérience-image". Dans l'absolu, il n'est pas nécessaire de décrire complètement le lien entre l'état  $X_t$  et les observations  $Y_t$ . Dans notre cadre de travail, ce lien s'avère complexe et non-linéaire. En revanche, il est nécessaire de définir le bruit aléatoire reliant ces deux grandeurs, c'est-à-dire la vraisemblance  $p(Y_t|X_t)$ . Cette dernière est indispensable à la mise à jour des poids du filtre particulaire. Quand les auteurs optent pour une caractérisation sans fusion, ils supposent en général la vraisemblance gaussienne. La variable sur laquelle est définie cette gaussienne est une grandeur scalaire issue de la mise en correspondance entre les caractéristiques connues *a priori*  $Y_0$  et les caractéristiques observées  $Y_t$  conditionnellement à une valeur  $X_t$ .

$$p(Y_t|X_t) \propto \exp\left(-\frac{d(Y_0, Y_t|X_t)^2}{\sigma^2}\right) \quad (3.23)$$

avec  $\sigma$  l'écart-type de la gaussienne, et  $d()$  une méthode de mise en correspondance, le plus souvent une distance. Nous précisons que la variable  $Y_t|X_t$  correspond aux caractéristiques observées localement à la position  $X_t$ . En d'autres termes, il s'agit des observations extraites dans l'imagette définie par  $X_t$ .

Il ne reste plus qu'une densité à définir : la densité d'importance. Comme évoqué dans la section précédente, ce choix est important quant à la qualité de l'estimation du filtre particulaire.

De plus, cette densité est celle qui propage les particules et les oriente dans l'image. En cas de propagation erronée, il sera impossible dans les étapes suivantes du filtre de retrouver la position réelle de l'objet. Dans un souci de simplicité par rapport à l'évaluation de l'apport de la fusion de données, nous choisissons la densité *a priori* en tant que densité d'importance. Comme cette densité a été définie prudemment, la propagation ne risque pas d'induire en erreur le filtre.

Au final, l'algorithme dans son contexte applicatif devient concrètement l'algorithme 5.

---

Algorithme : Filtre particulière simple appliqué au suivi

---

**Initialisation**,  $t = 0$  :

Échantillonner  $N$  particules selon une distribution  $p(X_0)$  et poser  $\forall i w_0^{(i)} = \frac{1}{N}$

**pour**  $t \geq 1$  **faire**

**pour**  $i = 1$  à  $N$  **faire**

    Échantillonner les particules selon  $X_t^{(i)} \sim p(X_t|X_{t-1})$

    Estimer la vraisemblance  $p(Y_t|X_t^{(i)})$

    Mettre à jour les poids selon  $w_t^{(i)} \propto w_{t-1}^{(i)} p(Y_t|X_t^{(i)})$

    Normaliser les poids selon  $\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^N w_t^{(j)}}$

**fin pour**

**si**  $\tilde{N}_{eff} > \text{seuil}$  **alors**

  Rééchantillonner en tirant  $N$  nouvelles particules  $\tilde{X}_t^{(j)}$  parmi les  $X_t^{(i)}$  selon une probabilité  $\tilde{w}_t^{(i)}$ , et leur affecter les poids  $\tilde{w}_t^{(j)} = \frac{1}{N}$

  Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t|Y_{0:t}) = \sum_{j=1}^N \delta_{\tilde{X}_t^{(j)}}(X_t) \frac{1}{N}$

**sinon**

  Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t|Y_{0:t}) = \sum_{i=1}^N \delta_{X_t^{(i)}}(X_t) \tilde{w}_t^{(i)}$

**fin si**

**fin pour**

**Fin**

---

Algorithme 5: Filtre particulière simple.

### 3.3 Processus et stratégies de fusion dans les filtres particulières

La fusion d'informations est un domaine de recherche en sciences de l'information en pleine expansion du fait de la multiplication des données transmises entre systèmes intelligents (systèmes à capteurs multiples, systèmes en réseau). Elle est aussi une réponse naturelle aux problèmes de robustesse des algorithmes de traitement de l'information car quand une approche est mise en difficulté par une situation particulière, une autre approche peut venir en soutien.

De manière générale, la fusion d'informations fonctionne selon quelques grands principes. Quand plusieurs sources d'information répondent à un même problème, c'est l'avis majoritaire qui l'emporte, toute la question étant de décider comment se définit une majorité. Théoriquement, plus le nombre de sources augmente, plus la décision fournie par le système de fusion doit être fiable. Néanmoins, si une source peu pertinente est intégrée au système, il est à craindre qu'elle ne fasse que bruyier le traitement de l'information. La prise en compte du contexte joue également un rôle important car elle permet de pondérer la confiance accordée aux sources et

d'affiner la décision finale.

Une fusion d'informations peut être qualifiée :

- d'implicite, si plusieurs sources interviennent dans le processus de fusion mais qu'elles ne sont pas combinées par une opération ou un procédé donné,
- d'explicite, si la fusion est réalisée dans le cadre d'une théorie où l'opération de fusion est énoncée à l'aide d'une règle donnée.

De nombreux auteurs ont déjà proposé de combiner différents extracteurs au sein d'un filtre particulaire. La plupart de ces approches utilisent une fusion implicite. Quand la fusion est explicite, elle se fait dans le cadre de la théorie des probabilités. Le filtre particulaire étant probabiliste, la méthode de fusion doit être compatible avec les probabilités. Cette section présente dans un premiers temps les approches existantes. Nous examinerons par la suite d'autres méthodes de fusion classiques et verrons si elles sont applicables au filtrage particulaire d'une part, et si elles peuvent contribuer à améliorer l'existant d'autre part.

### 3.3.1 Approches existantes

Nous présentons ici comment une fusion est opérée dans certains filtres particuliers.

#### Stratégies de fusion

Nous désignons par stratégies de fusion<sup>1</sup>, les approches dans lesquelles plusieurs sources sont utilisées à différents étages de l'exécution ou de la modélisation du filtre. La fusion est alors bel et bien effective, mais n'est pas contrôlée par une technique dédiée et l'étape de fusion n'apparaît pas explicitement dans le processus algorithmique. De nombreux auteurs ont recours à une stratégie de fusion plutôt qu'à une technique de fusion car une stratégie est beaucoup plus simple à mettre en œuvre et ne nécessite pas de justifications théoriques supplémentaires. Une première stratégie de fusion consiste à faire appel à une source d'information au niveau de la densité d'importance, tandis qu'une autre source est utilisée au niveau de la vraisemblance, c'est-à-dire pour l'estimation des poids. Bullock [Bullock 04] propose une telle approche, dans laquelle un extracteur de mouvement est utilisé au niveau de l'échantillonnage et un extracteur couleur au niveau de la mise à jour des poids. A  $t$  fixé, le filtre particulaire se résume presque entièrement aux deux étapes d'échantillonnage et de mise à jour ; il est donc logique de répartir les sources sur ces deux étapes.

L'autre stratégie possible consiste à utiliser plusieurs sources pour une de ces deux étapes. Que ce soit pour la densité d'importance ou pour la mise à jour, il est possible d'utiliser une source pour un groupe de particules et une autre pour le reste des particules. ICONDENSATION [Isard 98b] utilise un tirage aléatoire uniforme pour affecter un modèle de propagation aux particules. Selon ces modèles, diverses informations de forme ou de couleur sont utilisées.

Plutôt que de séparer la puissance de calcul pour divers modèles, il est également possible de rassembler toutes les données en sortie des extracteurs et de former un grand ensemble

---

<sup>1</sup>Nous précisons que, dans la littérature, la terminologie de stratégie de fusion peut faire référence au choix d'un ajencement opportun d'une étape de fusion au sein d'une architecture plus globale : étant donnée une certaine technique de fusion sélectionnée, à quel moment effectuer la fusion, comment exploiter l'information de fusion, etc.

d'attributs. Définir des densités conditionnellement à l'ensemble de tous les attributs est souvent délicat car il faut être capable de mettre en correspondance les attributs observés et le modèle d'objet même si certains attributs, à un moment donné, n'apportent aucune information discriminante. Les techniques de sélection de caractéristiques permettent de mettre en correspondance uniquement les attributs pertinents. Collins et Liu [Collins 03] utilisent des sommes pondérées des composantes des trois plans couleur pour former un ensemble de caractéristiques. À l'aide d'histogrammes et de rapports de vraisemblances, les auteurs déterminent les caractéristiques discriminant le mieux l'objet du fond de l'image. Toutefois, leur approche reste basée sur une seule catégorie d'information : la couleur. Leung et Gong [Leung 05] utilisent un critère d'information mutuelle entre attributs afin d'opérer la sélection.

### Fusion probabiliste

Les méthodes de fusion évoquées jusqu'à présent sont des méthodes implicites où la fusion est effectuée sans avoir recours à un processus spécifique. D'autres auteurs préfèrent une fusion explicite via la théorie des probabilités qui régit les filtres particulaires.

Au regard des probabilités, la solution de fusion la plus simple consiste à utiliser une hypothèse d'indépendance de variables aléatoires représentant des composantes de la variable  $Y_t$ . Soit  $y_t^{S_i}$  la variable aléatoire représentant les observations fournies par la source  $S_i$ . L'ensemble de ces variables correspond à la variable d'observation globale :  $Y_t = \{y_t^{S_i}\}_{i=1}^M$  si  $M$  extracteurs sont utilisés en parallèle. En supposant donc les variables  $y_t^{S_i}$  indépendantes, la vraisemblance s'écrit :

$$p(Y_t|X_t) = \prod_{i=1}^M p(y_t^{S_i}|X_t) \quad (3.24)$$

L'hypothèse d'indépendance est fautive dans l'absolu car les extracteurs sont appliqués sur la même image. L'hypothèse est plus crédible quand les sources d'information sont issues de capteurs différents. Toutefois, la pratique montre que cette estimation de la vraisemblance est meilleure qu'une vraisemblance définie sur une seule source. Brasnett *et al.* [Brasnett 06] utilisent cette technique tout en pondérant chaque source en fonction de sa distance avec le modèle d'objet. Dans la suite du mémoire, nous commettons un léger abus de langage en faisant référence à cette technique via le terme "fusion bayésienne".

Une autre solution probabiliste assez répandue est l'utilisation de mélanges. Les mélanges sont plus commodes à utiliser au niveau de la densité de filtrage elle-même :

$$p(X_t|Y_{0:t}) = \sum_{i=1}^M \pi_t^i p(X_t|y_t^{S_i}) \quad (3.25)$$

avec  $\pi_t^i$  les poids de chaque composante du mélange. Ces poids sont généralement déterminés par des méthodes d'apprentissage. Vermaak *et al.* [Vermaak 03] ont montré qu'il est possible d'évaluer récursivement les poids des composantes d'un mélange à partir des poids du filtre particulaire  $w_t^{(i)}$ . Les particules sont regroupées en différents ensembles attribués à chaque composante du mélange. L'approche n'est toutefois pas utilisée dans un cadre de fusion de sources.

Han *et al.* [Han 07] appliquent le mélange (3.25) à la fusion de données où chaque composante du mélange dépend d'une source. Les sources sont constituées d'extracteurs, de capteurs de nature différentes ou de capteurs similaires mais placés à différents endroits de la scène. Chaque composante de la vraisemblance est également représentée par un mélange. Il s'agit cette fois d'un mélange de gaussiennes dont les poids  $\kappa_{t,j}^i$  permettent d'évaluer directement les poids  $\pi_t^i$  du mélange principal. Nous précisons que l'approche utilisée ne correspond pas exactement à un filtre particulaire mais à un filtre bayésien à base de noyau [Han 05]. Ce filtre conserve néanmoins les deux étapes fondamentales d'échantillonnage et de d'évaluation de la vraisemblance. De très nombreux autres procédés et résultats sont utilisés au fil de l'algorithme proposé. Les résultats obtenus sont excellents en terme de suivi d'objets mais plusieurs critiques peuvent être émises. L'approche apparait au final complexe et lourde à mettre en œuvre. Les auteurs ne font pas mention du temps de calcul. De plus, cet algorithme utilise une très grande quantité d'informations (capteurs et extracteurs multiples) : le cadre applicatif est donc loin du notre.

Brethes *et al.* [Brèthes 06] ont réalisé une étude de différentes configurations de fusion dans les filtres particulaires. Ils utilisent trois modèles d'extracteurs : couleur, forme et mouvement. Trois vraisemblances et trois densités d'importance sont alors définies. Les techniques de fusion ne sont pas explicitées, néanmoins la présence de trois densités d'importance nous amène à penser que l'échantillonnage est réparti en groupes de particules ou qu'un mélange simple est utilisé. De même, la présence de trois vraisemblances semble indiquer que l'équation 3.24 est également utilisée. Leurs expériences combinent donc stratégies et techniques de fusion. Sans surprise, les auteurs concluent que les performances en robustesse augmentent avec le nombre de sources utilisées. Notons qu'ils précisent également que la fusion semble plus importante au niveau de la mise à jour des poids qu'à l'échantillonnage, car elle permet alors de rétablir le suivi après occultation.

Perez *et al.* [Perez 04] utilisent l'équation 3.24 mais l'exploitent différemment pour obtenir une fusion. Les auteurs proposent d'effectuer à chaque itération  $M$  échantillonnages et  $M$  mises à jour.  $M - 1$  états auxiliaires  $X_{t,1} \dots X_{t,M-1}$  sont nécessaires pour transiter de  $X_{t-1}$  à  $X_t$  :

$$p(X_t|X_{t-1}) = \int p_M(X_t|X_{t,M-1}) \dots p_1(X_t|X_{t,1}) dX_{t,M-1} dX_{t,1} \quad (3.26)$$

Il est alors possible de faire l'approximation selon laquelle la composante  $p(y_t^{S_j}|X_t)$  de la vraisemblance peut être incorporée après la  $j^{\text{ème}}$  propagation  $p_j(X_{t,i}|X_{t,j-1})$ . L'expression des poids  $w_t^{(i)}$  peut être obtenue en itérant sur  $j$  de la sorte. L'algorithme est appelé filtre particulaire avec échantillonnage en couches. Une critique possible de l'approche est le fait que l'ordre dans lequel sont pris en compte les observations intervient dans la qualité des résultats. Plusieurs scénarii sont examinés par les auteurs avec des modalités de couleur, de mouvement et de son pour le suivi du locuteur dans une conversation.

Wu et Huang [Wu 04] proposent ce qu'ils appellent une co-inférence. Deux filtres particulaires, l'un s'appuyant sur une information de couleur, l'autre sur une information de forme infèrent l'un dans l'autre. La figure 3.1 montre comment ces inférences se traduisent au niveau des dépendances entre variables. La fusion apporte un gain en robustesse significatif, en revanche il semble peu évident d'étendre la méthode à un nombre élevé de sources, les dépendances devenant de plus en plus complexes. Zhong *et al.* [Zhong 06] développent plus avant la dépendance

**Figure 3.1** : Exemples de deux chaînes d'état pour l'approche de Wu et Huang, figure tirée de [Wu 04]. L'exposant  $s$  fait référence au filtre utilisant l'information de forme. L'exposant  $c$  fait référence au filtre utilisant l'information de couleur.

graphique entre différentes parties de l'état et utilisent un algorithme de propagation de croyance pour gérer les dépendances.

Caron *et al.* [Caron 07a] utilisent l'équation 3.24 et introduisent des variables aléatoires supplémentaires décrivant la fiabilité des sources. Ainsi grâce à la prise en compte du contexte, les sources défaillantes sont éliminées du processus de fusion. Les auteurs proposent également de choisir entre plusieurs équations d'observation conditionnellement au contexte.

Dans la suite de cette partie, nous présentons d'autres approches de fusion, avant de voir si elles peuvent être adaptées au filtrage particulière. Si cette adaptation est possible, la fusion sera mise en œuvre au niveau d'une étape bien précise du filtre : l'estimation de la vraisemblance. Cibler explicitement la fusion à un endroit particulier nous permet de garder le contrôle sur la fusion de données et donc de dresser une évaluation objective de l'influence de la fusion sur le processus global. L'estimation de la vraisemblance est, comme nous l'avons vu dans les derniers paragraphes, une étape-clé dans laquelle il est naturel d'effectuer la fusion.

### 3.3.2 Méthodes à apprentissage

A partir du moment où des données étiquetées sont à disposition concernant la pertinence de telle ou telle combinaison de source, des techniques d'apprentissage peuvent être envisagées. Les méthodes d'apprentissage les plus usitées sont :

- les arbres de décision et forêts aléatoires,
- les séparateurs à vastes marges (SVM),
- les réseaux de neurones,
- les  $k$  plus proches voisins ( $k$ -ppv),
- les méthodes de boosting,
- les algorithmes génétiques,...

Cette liste est très loin d'être exhaustive. Comme nous l'avons vu dans les paragraphes précédents, dans un filtre particulière, la fusion peut intervenir principalement au niveau de la densité

d'importance ou au niveau de l'estimation de la vraisemblance. Il est possible d'imaginer que les méthodes d'apprentissage permettent de pondérer telle ou telle composante d'un mélange ou d'un terme multiplicatif de la vraisemblance. Cela revient alors à mesurer la fiabilité des sources, mais les techniques de combinaison de sources ne sont pas intrinsèquement améliorées. Les méthodes d'apprentissage viennent donc en complément mais ne peuvent se substituer aux règles probabilistes.

La densité  $p(y_t^{S_j} | X_t)$  peut être vue comme le score fourni par le classifieur  $j$  à la question : l'objet est-il dans l'imagette définie par  $X_t$ ? Pour aller plus loin dans le parallèle, des techniques de combinaisons de classifieurs peuvent être utilisées afin d'évaluer la vraisemblance globale. Kittler *et al.* [Kittler 98] ou Parikh *et al.* [Parikh 04] ont présenté des règles de décision pour la combinaison de classifieurs. Ces règles sont souvent, entre autres, la règle du max, du min, la règle multiplicative, la règle additive ou moyennée ou encore le vote à la majorité. Mis à part la règle multiplicative qui revient à la fusion bayésienne, toutes les autres ne trouvent aucune justification probabiliste et enfreignent même les règles de calcul probabiliste. L'analogie évoquée s'arrête donc ici et n'est pas applicable, faute de quoi la convergence du filtre particulaire ne serait plus assurée.

Les auteurs qui mêlent apprentissage et filtres particuliers, utilisent l'apprentissage pour améliorer les méthodes d'extraction de caractéristiques et non au sein du filtrage. Les travaux de Okuma *et al.* [Okuma 04] constituent une exception : les auteurs combinent l'approche de suivi de Vermaak *et al.* [Vermaak 03] et l'Adaboost de Viola et Jones [Viola 04]. Grâce à un apprentissage, l'Adaboost permet une meilleure détection des objets à suivre. Une densité liée à ce boosting est incorporée à la densité d'importance, ce qui fiabilise le processus global. Nous rappelons néanmoins que dans notre contexte de travail, l'apprentissage n'est pas possible.

### 3.3.3 Théorie des possibilités

#### Concepts des possibilités

La théorie des possibilités [Zadeh 78, Dubois 88] est une théorie mathématique de l'incertain née de la théorie des ensembles flous [Zadeh 65]. Elle permet de représenter à la fois l'incertitude et l'imprécision par l'intermédiaire de distributions de possibilités, souvent notées  $\pi(\cdot)$ , et définies sur un ensemble  $\Omega$ . La distribution de possibilités est en réalité une fonction d'appartenance à un ensemble flou. Contrairement au cas classique, l'appartenance à un ensemble n'est pas une valeur binaire (vrai ou faux) mais une grandeur continue qui traduit le degré d'appartenance. Cette alternative permet de traduire des jugements plus nuancés sur la validité des propositions examinées.

Deux mesures sont définies dans la théories des possibilités : la mesure de possibilité  $\Pi(\cdot)$  et la mesure de nécessité  $N(\cdot)$ . La première caractérise le fait que l'événement étudié est possible, tandis que la seconde quantifie le degré de certitude lié à l'événement. Nous énonçons quelques

propriétés sur ces mesures :

$$\Pi : 2^\Omega \longrightarrow [0, 1] \quad (3.27)$$

$$\Pi(\emptyset) = 0 \quad (3.28)$$

$$\Pi(\Omega) = 1 \quad (3.29)$$

$$\Pi(\cup_{i=1\dots n} A_i) = \max_{i=1\dots n} (A_i) \quad (3.30)$$

$$N : 2^\Omega \longrightarrow [0, 1] \quad (3.31)$$

$$N(\emptyset) = 0 \quad (3.32)$$

$$N(\Omega) = 1 \quad (3.33)$$

$$N(\cap_{i=1\dots n} A_i) = \min_{i=1\dots n} (A_i) \quad (3.34)$$

On a de plus :

$$N(A) = 1 - \Pi(\bar{A}) \quad (3.35)$$

$$N(A) \leq \Pi(A) \quad (3.36)$$

Si dans le cadre des probabilités, la mesure de probabilité  $P(A)$  représente le degré de confiance sur l'événement  $A$ , dans le cadre des possibilités c'est l'intervalle  $[N(A), \Pi(A)]$  qui représente cette confiance. La distribution de possibilité est définie comme une représentation atomique de la mesure  $\Pi()$  :

$$\forall A \subset \Omega, \Pi(A) = \max_{u \in A} \pi(u) \quad (3.37)$$

$$\forall A \subset \Omega, N(A) = \min_{u \in A} 1 - \pi(u) \quad (3.38)$$

Comme pour les probabilités, une distribution de possibilité peut être associée à une variable  $x$  dont  $\Omega$  constitue l'ensemble des valeurs possibles de  $x$ . D'après Dubois et Prade [Dubois 04], les interprétations suivantes d'une distribution de possibilité  $\pi_x$  sont valables  $\forall u \in \Omega$  :

- $\pi_x(u) = 0 \Leftrightarrow "x = u"$  est impossible.
- $\pi_x(u) = 1 \Leftrightarrow "x = u"$  est normal, sans surprise.

Outre ces aspects théoriques, la théorie des possibilités possède l'avantage de définir des opérations de fusion de sources. Ces opérations sont en pratique ce que nous recherchons pour notre problématique. De très nombreuses règles de combinaison sont définies sur les distributions de possibilités :

- les T-normes,
- les T-conormes,
- les moyennes,
- les sommes symétriques,
- des opérateurs de conflit ou de fiabilité.

Bloch décrit la grande majorité de ces règles de combinaisons dans [Bloch 96]. La recherche d'un opérateur de fusion optimal reste un thème de recherche actuel [Oussalah 03], et certains nouveaux opérateurs sont donc absents de l'étude de Bloch datant de 1996. Néanmoins, en pratique, les T-normes et T-conormes restent plus fréquemment utilisées. Les T-normes s'attachent à extraire de la concordance entre sources. Les T-conormes peuvent combiner des sources discordantes.

La théorie des possibilités offre donc tout un arsenal d'opérateurs de fusion intéressants. Cette diversité laisse espérer un gain de qualité du traitement de l'information comparativement aux probabilités plus rigides. Néanmoins, pour être applicable dans le filtre particulaire, il faut d'une part modéliser le problème de suivi en termes de possibilités et, d'autre part, la fusion possibiliste doit être transcrite en probabilités pour que le filtre particulaire puisse fonctionner. Nous proposons ci-après une méthode permettant la réalisation de ces deux points.

### Construction des possibilités

Pour construire des distributions de possibilité, il faut dans un premier temps définir un ensemble  $\Omega$  représentant l'ensemble des solutions au problème posé. Étant donné que nous cherchons à injecter la fusion de données au niveau de l'estimation de la vraisemblance  $p\left(Y_t|X_t^{(i)}\right)$ , il faudrait donc, pour chaque source donnée  $S_j$ , évaluer une possibilité pour chaque  $y_t^j$  et ce conditionnellement à  $X_t^{(i)}$ . L'espace sur lequel est définie  $y_t^j$  est certes discret mais beaucoup trop vaste pour être modélisé intégralement sous peine d'une complexité calculatoire prohibitive. De plus, cet espace dépend de la méthode d'extraction choisie pour la source  $S_j$ , il serait alors très délicat de fusionner des sources définies sur des ensembles hétérogènes.

Pour simplifier ce problème, nous proposons de définir un ensemble  $\Omega$  où l'hypothèse  $\omega_1 \in \Omega$  traduit le fait que l'objet suivi est observé dans l'imagette définie par  $X_t^{(i)}$ . Cette définition reste valable quelle que soit la source  $S_j$ . En d'autres termes, l'hypothèse  $\omega_1$  est intrinsèquement liée à la variable  $y_t^j|X_t^{(i)}$ . Deux autres hypothèses antagonistes à  $\omega_1$  constituent  $\Omega$ . La définition de  $\Omega = \{\omega_1, \omega_2, \omega_3\}$  est alors identique pour chaque source. Le problème revient en quelque sorte à un problème de classification de l'imagette analysée. Dans le cadre des fonctions de croyance, nous présenterons une définition de  $\Omega$  identique, plus de détails sont alors fournis sur le sens que prennent les autres hypothèses formant  $\Omega$  (cf. 3.5.1).

Le cardinal de  $\Omega$  étant faible, la définition des possibilités est simple. La condition d'existence d'une distribution de possibilité  $\pi(\cdot)$  est la suivante :

$$\sup \{\pi(x), x \in \Omega\} = 1 \quad (3.39)$$

Il nous faut définir une distribution  $\pi_j(\cdot)$  pour chaque source  $S_j$ . La construction des distributions se fait sur la base d'une mise en correspondance des observations relatives à la particule  $X_t^{(i)}$  et à la source  $S_j$  avec un modèle. Cette mise en correspondance s'effectue à l'aide d'une distance de Bhattacharyya  $d_{Bhattacharyya}(\cdot, \cdot)$ . Cependant, cette mise en correspondance dépend également de l'hypothèse  $\omega_k$  dont le degré d'incertitude est en cours d'évaluation. Il faut donc également que la distribution dépende de  $k$  :  $\pi_{j,k}(\cdot)$ . Nous proposons alors le modèle suivant :

$$\pi_{j,k}(\omega_k) = 1 \quad (3.40)$$

$$\pi_{j,k}(\omega_{k'}) = 1 - \alpha_{j,k} \exp(-\xi_{j,k} d^2) \text{ si } k' \neq k \quad (3.41)$$

$$\text{avec } d = d_{Bhattacharyya}\left(y_0^{S_j,k}|X_0, y_t^{S_j,k}|X_t^{(i)}\right) \quad (3.42)$$

$\alpha_{j,k}$  et  $\xi_{j,k}$  sont des paramètres du modèle ajustés par l'utilisateur selon un principe de type erreur/correction<sup>2</sup>. Nous notons  $y_t^{S_j,k}|X_t^{(i)}$  le vecteur d'observations effectuées à  $t$  sur l'imagette

<sup>2</sup>Dans l'absolu, il est préférable d'avoir recours à une méthode d'optimisation de type algorithme génétique pour déterminer cette valeur.

définie selon  $X_t^{(i)}$  par la source  $S_j$  en fonction du modèle relatif à l'hypothèse  $\omega_k \in \Omega$ .

Le modèle proposé est simple et nous avons conscience qu'il est possible de définir des modèles de possibilités plus évolués. Toutefois, dans nos expériences, le modèle présente montre un pouvoir discriminant suffisant pour établir une comparaison entre méthodes de l'incertain. Toutes les distributions ainsi générées sont ensuite agrégées à l'aide d'une règle de combinaison possibiliste, produisant ainsi, une unique distribution  $\pi(\cdot)$  contenant l'information de fusion.

### Filtre particulière avec fusion possibiliste

Maintenant qu'une distribution de possibilité a été obtenue par fusion, il faut ramener cette distribution dans le cadre probabiliste afin d'apporter l'information de fusion au filtre particulière. De même que pour les opérateurs de fusion, de très nombreux auteurs ont mis au point des transformations permettant de passer des possibilités vers les probabilités. Zadeh posa dès ses premiers travaux [Zadeh 78] un principe concernant de telles transformations :

- Une forte probabilité doit impliquer une forte possibilité, alors qu'une forte possibilité n'implique pas forcément une forte probabilité.
- Une faible possibilité doit impliquer une faible probabilité, alors qu'une faible probabilité n'implique pas forcément une faible possibilité.

Ce principe est connu sous le nom de principe de consistance possibilité-probabilité. En termes mathématiques, le principe revient à :

$$\forall u, u' \in \Omega, \pi(u) = 0 \Rightarrow P(u) = 0 \quad (3.43)$$

$$\text{et } \pi(u) > \pi(u') \Rightarrow P(u) > P(u') \quad (3.44)$$

Dubois et Prade [Dubois 83] vont plus loin dans les exigences et imposent :

$$\forall A \subseteq \Omega, \Pi(A) \geq P(A) \quad (3.45)$$

$$\text{et } \forall u, u' \in \Omega, \pi(u) > \pi(u') \Leftrightarrow P(u) > P(u') \quad (3.46)$$

Les deux auteurs parlent respectivement de consistance probabilité-possibilité et de préservation de préférence. Ils sont alors en mesure de proposer un processus de transformation respectant ces deux contraintes. La transformation suppose qu'il soit possible de ranger les possibilités de la plus grande à la plus faible sur  $\Omega : \pi(\omega_1) \geq \pi(\omega_2) \geq \dots \geq \pi(\omega_M)$ . La transformation est alors définie par :

$$P(\omega_i) = \sum_{j=i}^M \frac{1}{j} (\pi(\omega_j) - \pi(\omega_{j+1})) \quad (3.47)$$

avec une possibilité fictive  $\pi(\omega_{M+1})$  valant 0. La transformation inverse est possible mais ne présente pas d'intérêt pour notre cas.

Une étude de He et Qu [He 07] présente la grande majorité des autres transformations existantes. L'étude ne conclut pas à la supériorité d'une transformation par rapport aux autres, mais les auteurs indiquent que certaines sont difficiles à interpréter d'un point de vue de l'information. La transformation de Dubois et Prade étant rapide et clairement énoncée, elle a été retenue pour mettre en place un filtre particulière avec fusion possibiliste.

En fusionnant les sources, puis en transformant le résultat de cette fusion, la probabilité  $P(\omega_1)$  analyse l'observation globale  $Y_t$  pour un  $X_t = X_t^{(i)}$  donné. Nous proposons donc l'approximation suivante :

$$p\left(Y_t|X_t^{(i)}\right) \approx P(\omega_1) \quad (3.48)$$

Nous disposons à présent de tous les éléments nécessaires à la conception d'un algorithme de filtrage, et nous proposons la procédure 6.

---

Algorithme : Filtre particulaire simple avec fusion possibiliste

---

**Initialisation**,  $t = 0$  :

Échantillonner  $N$  particules selon une distribution  $p(X_0)$  et poser  $\forall i w_0^{(i)} = \frac{1}{N}$

**pour**  $t \geq 1$  **faire**

**pour**  $i = 1$  à  $N$  **faire**

    Échantillonner les particules selon  $X_t^{(i)} \sim p(X_t|X_{t-1})$

**pour**  $j = 1$  à  $M$  **faire**

      Évaluer la distribution de possibilité pour la source  $S_j$  sur un  $\Omega$

**fin pour**

    Fusionner les distributions de possibilité selon une règle donnée

    Appliquer la transformation 3.47 à la distribution issue de la fusion pour obtenir une probabilité  $P()$  sur  $\Omega$

    Accepter  $P(\omega_1)$  en tant qu'estimation de la vraisemblance  $p\left(Y_t|X_t^{(i)}\right)$

    Mettre à jour les poids selon  $w_t^{(i)} \propto w_{t-1}^{(i)} p\left(Y_t|X_t^{(i)}\right)$

    Normaliser les poids selon  $\tilde{w}_t^{(i)} = \frac{w_t^{(i)}}{\sum_{j=1}^N w_t^{(j)}}$

**fin pour**

**si**  $\hat{N}_{eff} > \text{seuil}$  **alors**

    Rééchantillonner en tirant  $N$  nouvelles particules  $\tilde{X}_t^{(j)}$  parmi les  $X_t^{(i)}$  selon une probabilité  $\tilde{w}_t^{(i)}$ , et leur affecter les poids  $\tilde{w}_t^{(j)} = \frac{1}{N}$

    Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t|Y_{0:t}) = \sum_{j=1}^N \delta_{\tilde{X}_t^{(j)}}(X_t) \frac{1}{N}$

**sinon**

    Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t|Y_{0:t}) = \sum_{i=1}^N \delta_{X_t^{(i)}}(X_t) \tilde{w}_t^{(i)}$

**fin si**

**fin pour**

**Fin**

---

Algorithme 6: Filtre particulaire simple avec fusion possibiliste.

### 3.3.4 Théories des fonctions de croyance

Développée conjointement par Dempster [Dempster 67], puis par son élève Shafer [Shafer 76], la théorie des fonctions de croyance est un formalisme mathématique destiné à répartir une quantité de croyance attribuée sur différentes hypothèses relatives à un même problème. Ces hypothèses sont rassemblées dans un ensemble  $\Omega$  appelé cadre de discernement. La répartition est symbolisée par une fonction  $m(\cdot)$ , appelée masse de croyance élémentaire, et définie sur l'ensemble des parties de  $\Omega$ . La théorie des fonctions de croyance est donc une troisième théorie de l'incertain. Il est surtout à noter, que cette théorie englobe les deux autres précédemment évoquées (probabilités et possibilités).

Dans la littérature, elle est souvent appelée théorie de Dempster-Shafer ou encore théorie de l'évidence<sup>3</sup>. Elle a déjà montré à travers de nombreuses publications, qu'elle peut donner de meilleures performances que les méthodes classiques dans des domaines tels que la classification [Denoeux 08], l'imagerie médicale [Zhang 07], la détection de frontière entre régions d'images [Debain 00], l'appariement des observations aux hypothèses dans le cadre d'un suivi par MHT [Mourllion 05], la fusion d'images [Bloch 08] ou le traitement du signal [Appriou 97].

Notons que quelques auteurs ont mêlé filtrage particulière et fonctions de croyance mais avec des objectifs différents des nôtres. Caron *et al.* [Caron 07a] utilisent un filtre particulière pour le suivi d'appareils en vols, tandis que les fonctions de croyance sont utilisées pour classifier les appareils en diverses catégories (gros-porteurs militaires, civils, avions de chasse). Faux *et al.* [Faux 06] utilisent les fonctions de croyance afin de concevoir un modèle de teinte de la peau par fusion de sources couleur. Ce modèle intègre un *a priori* contextuel, puis en s'appuyant sur ce modèle, un suivi du visage par filtrage particulière est réalisé mais il n'y a pas de communication entre le formalisme des fonctions de croyance et celui du filtrage.

De manière analogue aux possibilités, les fonctions de croyance peuvent :

- représenter des connaissances imprécises,
- être fusionnées à l'aide de règles dédiées,
- être transformées en probabilités à l'aide la transformation pignistique (*cf.* 3.4.7),
- permettre d'évaluer au final la vraisemblance  $p\left(Y_t|X_t^{(i)}\right)$  en utilisant une définition de  $\Omega$  similaire à celle présentée pour les possibilités.

Il est donc tout à fait possible de mettre au point un filtre particulière avec fusion crédibiliste<sup>4</sup> de données. Les fonctions de croyance constituent le cadre de fusion de données que nous avons retenu pour notre approche. Nous préférons d'abord justifier ce choix avant de détailler le formalisme des fonctions de croyance (*cf.* 3.4) et d'explicitier notre approche (*cf.* 3.5.3).

### 3.3.5 Justification du choix des fonctions de croyance comme méthode de fusion

Parmi les méthodes de fusion, seules les théories de l'incertain semblent présenter un réel intérêt pour effectuer une fusion au sein d'un filtre particulière. Parmi ces théories, celle dite des fonctions de croyance se distingue ; nous expliquons ici pourquoi.

#### Défauts des probabilités

Bien que les probabilités traduisent un degré de confiance sur la réalisation d'un événement, le jugement émis par une mesure de probabilité est relativement strict au sens où si l'événement  $A$  possède une probabilité  $\mu$ , alors automatiquement le non-événement  $\bar{A}$  possède la probabilité  $1 - \mu$ . Cette propriété est la conséquence directe d'un axiome du calcul des probabilités : *L'image de la réunion d'un nombre fini ou infini dénombrable de parties de  $\Omega$  disjointes est égale à la somme des images.*

$$P(A_1 \cup A_2 \cup \dots \cup A_n \cup \dots) = \sum_{i=1}^{\infty} P(A_i) \quad (3.49)$$

<sup>3</sup>Le terme évidence correspond à l'anglais *evidence*, qui signifie plutôt "preuve".

<sup>4</sup>Le terme crédibiliste fait référence à la théorie des fonctions de croyance.

Cet axiome n'a rien d'illogique mais, dans un processus de décision complexe, il induit un double jugement (sur  $A$  et  $\bar{A}$ ) qui peut être handicapant. Comme nous le verrons dans la partie suivante, la théorie des fonctions de croyance relâche grandement cette contrainte. Une des forces des fonctions de croyance est donc de n'affecter du crédit à un événement que si cette affectation est justifiée par la prise en compte de données.

Dans le cadre d'un processus multi-sources, il est relativement fréquent qu'une source favorise l'événement  $A$ , tandis qu'une autre favorise un événement  $B$  disjoint de  $A$ . On parle alors de situation de conflit. Contrairement aux probabilités, les fonctions de croyance disposent d'outils spécifiques formalisant la notion de conflit. Il est ainsi naturellement plus facile de le prendre en compte.

De même, le fait qu'une source n'ait aucune information à procurer à un moment donné doit être géré habilement afin que cette source ne bruite pas le processus de décision. Le phénomène d'absence d'information est appelé ignorance dans le cadre des fonctions de croyance. En probabilité, il se traduit par une équiprobabilité sur les hypothèses envisagées. Même si ces valeurs ne biaisent pas en général le traitement, elles interviennent sans que "leur ignorance" soit explicitement prise en compte. Les fonctions de croyance ont à disposition les outils nécessaires pour décrire et gérer l'ignorance.

Enfin et surtout, les fonctions de croyance disposent de règles de combinaison conçues pour la fusion de données alors que les probabilités, à la base, n'ont pas vraiment été développées à cet effet.

### Défauts des possibilités

La théorie des possibilités peut également combler les défauts des probabilités. Néanmoins, les possibilités, de même que les probabilités, sont une forme particulière de fonctions de croyance (*cf.* 3.4.1). Les fonctions de croyance offrent donc une plus grande liberté de modélisation que les possibilités. Quitte à choisir un formalisme généralisant les concepts des probabilités, autant choisir le plus vaste des deux.

### Défauts d'autres théories de l'incertain

Il existe une généralisation de la théorie des fonctions de croyance au cas continu. Toutefois, la complexité des calculs et des concepts est grandement accentuée par le recours aux fonctions de croyances continues. Bien que les premiers travaux sur ce sujet furent introduit par Dempster conjointement à la théorie relative au cas discret, en pratique, l'utilisation des fonctions de croyance continues est encore rare de nos jours. Quelques articles mettant en œuvre ces outils sont à noter : [Smets 05, Smets 07, Caron 07b].

Récemment, Dezert et Smarandache [Dezert 02, Smarandache 04b] ont proposé une théorie issue directement de celle des fonctions de croyance en généralisant la structure algébrique sur laquelle sont définies des fonctions de croyance. La théorie de Dezert-Smarandache (DSmT) tente de répondre à un manque de modélisation des fonctions de croyance vis-à-vis d'hypothèses conflictuelles. Par exemple, pour un problème où l'on voudrait déterminer la température

d'une pièce, l'hypothèse "chaud" et l'hypothèse "froid" sont modélisables par les fonctions de croyance. Avec la théorie de Dezert-Smarandache, l'hypothèse paradoxale "chaud et froid" sera prise en compte en tant que telle dans l'ensemble des hypothèses. Dans la théorie des fonctions de croyance, cette hypothèse est intégrée à la mesure de conflit qui absorbe tous les paradoxes. Au moment où nous avons émis le besoin d'opérer une fusion dans un filtre particulière, il nous est apparu que nous manquions de recul par rapport aux bénéfices apportés par la théorie de Dezert-Smarandache, qui n'est utilisée que depuis peu.

En vertu de l'ensemble de ces remarques, nous exploiterons les fonctions de croyance pour effectuer une fusion de données dans un filtre particulière. L'approche possibiliste servira de comparatif.

### 3.4 Formalisme des fonctions de croyance

La présentation qui est faite du formalisme des fonctions de croyance s'inscrit dans celle proposée par Smets [Smets 94] et appelée **modèle de croyances transférables** (Transferable Belief Model, TBM). D'autres modèles existent dans la littérature : modèle de Dempster [Dempster 68], modèle de Shafer [Shafer 76] ou encore le modèle des Hints de Kholas et Monney [Kholas 95].

#### 3.4.1 Masse de croyance élémentaire

Nous donnons ici une succession de définitions de base accompagnées de quelques remarques.

Soit un problème donné pour lequel un agent doit choisir une solution parmi un ensemble fini de solutions. Ces solutions, ou hypothèses  $\omega_i$ , sont supposées mutuellement exclusives<sup>5</sup> et exhaustives. L'ensemble des hypothèses constituant le cadre de discernement est noté  $\Omega = \{\omega_1, \dots, \omega_K\}$ . L'ensemble des parties de  $\Omega$  est noté  $2^\Omega$ . Toute l'information qui permet à un agent de construire ses croyances vis-à-vis du problème est appelé **élément d'évidence**  $Ev$ . Un élément d'évidence correspond au fait *qu'une source  $S$  déclare une proposition  $E$  vraie et que l'agent construisant ses croyances reconnaît la validité des informations fournies par  $S$* . Dans la suite du document, nous ferons souvent l'amalgame entre une source d'information et l'élément d'évidence d'un agent.

L'agent peut affecter une quantité de croyance à chaque élément  $A$  de  $2^\Omega$ . On appelle **masse de croyance** ou **fonction de masse** (FDM) la fonction représentant cette assignation. Elle est notée  $m^\Omega [Ev](\cdot)$ , ainsi,  $m^\Omega [Ev](A)$  symbolise le crédit attribué à l'événement  $A \subseteq \Omega$ . Pour plus de simplicité, la notation  $[Ev]$  sera généralement omise, de même que celle précisant le cadre sur lequel est définie la fonction.

Une fonction de masse respecte par définition les contraintes suivantes :

$$m(A) : 2^\Omega \rightarrow [0, 1] \quad (3.50)$$

$$\sum_{A \subseteq \Omega} m(A) = 1 \quad (3.51)$$

<sup>5</sup>L'intersection de ces hypothèses correspond à l'ensemble vide.

L'équation 3.51 exprime le fait que la somme totale des croyances attribuées vaut 1. Cette propriété remplace l'axiome probabiliste énoncé en 3.3.5. Cet axiome imposerait  $\sum_{\omega_i \in \Omega} P(\omega_i) = 1$ . Il est donc clair que la propriété 3.51 des fonctions de croyance est beaucoup moins contraignante. Elle est en revanche nécessaire pour que les rapports de taille entre masses prennent leur sens.

Nous noterons  $\mathfrak{B}^\Omega$  l'ensemble des fonctions de masse définies sur  $\Omega$ .

### Éléments particuliers

Parmi les éléments de  $2^\Omega$ , certains jouent un rôle particulier :

- Un ensemble  $A$  tel que  $m(A) > 0$  est appelé **élément focal**. Nous noterons  $\mathcal{F}_m^\Omega$  l'ensemble des éléments focaux d'une FDM  $m(\cdot)$ .
- L'ensemble vide  $\emptyset$  : Il représente la possibilité que la solution au problème ne soit pas contenue dans  $\Omega$ . Certains auteurs considèrent qu'aucune masse ne doit être attribuée à cette hypothèse en considérant que, le cas échéant, le problème a été mal posé. On parle alors de **monde fermé**. A l'inverse si cette hypothèse est prise en compte, ce qui revient à autoriser  $m(\emptyset) > 0$ , on parle de **monde ouvert**. Le TBM préconise l'usage d'un monde ouvert afin de rester le plus générique possible. De plus, quand deux sources sont conflictuelles et combinées de façon conjonctive (cf. 4.3.2 pour une définition de la conjonction), du crédit est attribué à l'hypothèse  $\emptyset$ . Certains auteurs appellent donc cette hypothèse "**conflit**", par léger abus de langage. La gestion du conflit fait partie des justifications de l'utilisation des fonctions de croyance, d'où un second intérêt à travailler en monde ouvert. Les avantages des mondes fermés ou ouverts ont été largement discutés par Smets [Smets 90].
- L'ensemble  $\Omega$  : Il représente le fait que la solution au problème est contenue dans  $\Omega$  mais qu'il est impossible de déterminer quelle hypothèse est la bonne. Ainsi  $m(\Omega)$  permet de quantifier l'ignorance face au problème. Par la suite, nous ferons référence par abus de langage à  $\Omega$  en tant qu'"**ignorance**" quand  $\Omega$  est vu en tant qu'hypothèse et non en tant qu'ensemble.

### Fonction de masse simple

Nous noterons  $A^w(\cdot)$  la fonction de masse telle que cette dernière ait seulement deux éléments focaux :  $A$  et  $\Omega$ , et que :

$$A^w(A) = 1 - w \quad (3.52)$$

$$A^w(\Omega) = w \quad (3.53)$$

Par exemple, la fonction de masse qui représente une ignorance totale face au problème est  $\Omega^0(\cdot)$ . Cette fonction est aussi appelée masse de croyance vide. La fonction de masse représentant le conflit total est  $\emptyset^0(\cdot)$ . Ces fonctions sont dites fonctions de masse simples. L'abréviation FDMS sera utilisée pour faire référence aux fonctions de masse simples.

### Fonction de masse dogmatique

Une fonction de masse est dogmatique si et seulement si :  $m(\Omega) = 0$

### Fonction de masse normale

Une fonction de masse est normale si et seulement si :  $m(\emptyset) = 0$ .

### Fonction de masse consonante

Une fonction de masse est dite consonante si et seulement si l'inclusion est une relation d'ordre totale sur ses éléments focaux. Si ses éléments focaux sont  $\{A_i\}_{i=1}^n$ , il existe une permutation  $perm(\cdot)$  sur les indices telle que :  $A_{perm(1)} \subseteq A_{perm(2)} \subseteq \dots \subseteq A_{perm(n)}$ . Les fonctions de croyance consonantes sont équivalentes à des possibilités (cf. [Dubois 86a]).

### Fonction de masse bayésienne

Une fonction de masse  $m(\cdot)$  est bayésienne si et seulement si tous ces éléments focaux sont des singletons  $\{\omega_i\}$ . En d'autres termes,  $\forall A \subseteq \Omega$ , tel que  $\#(A) > 1$ ,  $m(A) = 0$ . Les fonctions de croyance bayésiennes sont des distributions de probabilités.

#### 3.4.2 Le principe du minimum d'information

Le principe du minimum d'information du TBM n'est pas une propriété mais un précepte auquel dans l'absolu des opérations crédibilistes doivent se conformer. Il s'énonce de la manière suivante : *si un processus crédibiliste amène à choisir une fonction de masse parmi un ensemble de fonctions proposées, il convient de choisir la moins engagée des fonctions*. Le terme "**engagé**" peut se comprendre dans le sens "informatif" ou "risqué". La fonction de masse la moins engagée est naturellement  $\Omega^0(\cdot)$ , puisqu'elle ne prend pas position par rapport au problème. A l'inverse une fonction de masse de type  $\omega_i^0(\cdot)$  est totalement engagée pour l'hypothèse  $\omega_i$ .

Il est possible de quantifier l'engagement d'une fonction de masse. Diverses mesures ont été proposées à cet effet [Harmanec 96, Abellan 99, Klir 99], mais la plus commune reste celle dite de **non-spécificité** proposée par Dubois et Prade [Dubois 85] :

$$\eta(m) = \sum_{\emptyset \neq A \subseteq \Omega} m(A) \ln(\#(A)) \quad (3.54)$$

Cette formulation est facilement interprétable car les valeurs prises par la fonction sont pondérées par le cardinal des ensembles ainsi, par exemple, les singletons sont favorisés.

En outre, la notion de minimum d'engagement est à l'origine de la définition de plusieurs relations d'ordre partiel entre les fonctions de masses. Le paragraphe 2.3 de Denœux [Denœux 08] donne les définitions des différentes relations d'ordre existant ainsi que leur niveau de contrainte.

#### 3.4.3 Règles de combinaison

Le formalisme des fonctions de croyance offre la possibilité de fusionner différentes sources d'information, représentées par différentes fonctions de masses à l'aide d'opérateurs appelés règles de combinaison. Si l'on note une règle générique  $\odot$ , elle doit être telle que :

$$\odot : \mathfrak{B}^\Omega \times \mathfrak{B}^\Omega \longrightarrow \mathfrak{B}^\Omega \quad (3.55)$$

$$(m_1, m_2) \longrightarrow m_{1 \odot 2} \quad (3.56)$$

La première règle de combinaison fut introduite par Dempster dans le cadre d'un monde fermé. La règle de Dempster est notée  $\oplus$  et se définit ainsi :

$$\forall A \neq \emptyset, m_1 \oplus m_2 (A) = \frac{1}{1 - \kappa} \sum_{B, C / B \cap C = A} m_1 (B) m_2 (C) \quad (3.57)$$

$$\kappa = \sum_{B, C / B \cap C = \emptyset} m_1 (B) m_2 (C) \quad (3.58)$$

$$m_1 \oplus m_2 (\emptyset) = 0 \quad (3.59)$$

Le terme  $\kappa$  est en réalité l'expression du conflit. Dans le monde fermé sa valeur est rejetée, et comme il n'y a aucune raison pour attribuer cette valeur à une hypothèse plutôt qu'à une autre, elle est répartie identiquement sur l'ensemble des hypothèses par normalisation. La règle a été adaptée au monde ouvert par Smets. Elle est alors appelée règle conjonctive et notée  $\odot$ . Sa définition se simplifie et devient :

$$m_1 \odot m_2 (A) = \sum_{B, C / B \cap C = A} m_1 (B) m_2 (C) \quad (3.60)$$

Si l'on examine le terme  $m_1 \odot m_2 (\emptyset)$ , il apparaît que sa valeur est augmentée si la première source attribue du crédit à un ensemble  $A$  et que la seconde attribue du crédit à un autre ensemble  $B$ , disjoint de  $A$  ( $A \cap B = \emptyset$ ). Il s'agit donc bien d'une situation de conflit. Cela ne serait pas forcément le cas avec une règle différente. Comme nous le précisons, ce constat n'est valable que pour une règle de type conjonctif (*cf.* 4.3.2). Le terme conjonctif correspond à une propriété des règles de combinaison que nous énoncerons dans le chapitre 4.

Dans le présent chapitre, nous cherchons à mettre en place la physiologie globale de notre approche de suivi. Toutefois, l'étape de combinaison des fonctions de masse joue un rôle primordial dans l'obtention de bons résultats et dans la contribution que nous souhaitons apporter au problème de suivi. C'est là que tout se joue au niveau de la fusion. Les règles de combinaison nécessitent donc une étude approfondie qui sera présentée dans le chapitre 4. Nous présenterons les principales règles de combinaison rencontrées dans la littérature. Nous étudierons les mécanismes et les propriétés de ces règles, afin de comprendre comment y est traitée l'information. En fonction de nos besoins en terme de suivi d'objets, nous proposerons alors un opérateur de combinaison capable de satisfaire nos exigences.

Pour les tests qui seront menés dans la suite du chapitre, la règle utilisée sera la règle conjonctive qui reste la plus communément appliquée.

### 3.4.4 Autres représentations des croyances

Il existe d'autres manières de représenter les croyances que la fonction de masse :

$$\text{– fonction de crédibilité } bel(X) = \sum_{A \subseteq X, A \neq \emptyset} m(A) \quad (3.61)$$

$$\text{– fonction d'implicabilité } b(X) = \sum_{A \subseteq X} m(A) \quad (3.62)$$

$$\text{– fonction de plausibilité } pl(X) = \sum_{A \cap X \neq \emptyset} m(A) \quad (3.63)$$

$$\text{– fonction de communalité } q(X) = \sum_{X \subseteq A} m(A) \quad (3.64)$$

Dans la majorité des applications utilisant les fonctions de croyance, ces fonctions sont très utiles car elles permettent de prendre en compte différents degrés de certitude ou de risque dans le processus de prise de décision. Cependant, elles ne seront pas utilisées dans nos travaux.

Nous ferons néanmoins référence à deux autres représentations plus rarement évoquées : les fonctions de poids conjonctifs et les fonctions de poids disjonctifs. Quelques notions doivent être introduites avant de présenter ces deux types de fonctions. Smets [Smets 95] a montré qu'une fonction de masse non-dogmatique peut se décomposer en combinaison conjonctive de fonctions de masse simple généralisées (FDMSG). Une fonction de masse simple généralisée  $\mu(\cdot)$  est telle que :

$$\mu : 2^\Omega \longrightarrow \mathbb{R} \quad (3.65)$$

$$\mu(A) = 1 - w \quad (3.66)$$

$$\mu(\Omega) = w \quad (3.67)$$

avec  $A \neq \Omega$  et  $w \in [0, +\infty)$ . La définition est très proche de celle d'une fonction de masse simple, puisqu'il n'y a que deux éléments focaux, mais la différence vient du fait qu'elle est à valeur dans  $\mathbb{R}$ . Deux cas se présentent alors :

- $w \leq 1$  : La FDMSG est une FDMS.
- $w > 1$  : La FDMSG n'est pas une FDM. Smets parle alors de fonction de masse simple inverse (FDMSI).

Smets interprète les FDMSI comme une représentation de la croyance selon laquelle il existe des raisons de ne *pas* croire en  $A$ . En effet, pour une FDMSI donnée  $A^w(\cdot)$ , une combinaison conjonctive avec la FDMS  $A^{1/w}$  donne la masse de croyance vide  $\Omega^0(\cdot)$ . En d'autres termes, la FDMSI  $A^w(\cdot)$  correspond à une "dette de croyance" vis-à-vis de  $A$ , cette dette devant être effacée avant de commencer à attribuer du crédit à  $A$ . Smets montre alors que pour toute FDM non-dogmatique  $m(\cdot)$ , il existe une fonction  $w : 2^\Omega \longrightarrow [0, +\infty)$  telle que :

$$m = \bigoplus_{A \subseteq \Omega} A^{w(A)} \quad (3.68)$$

$$w(A) = \prod_{B \supseteq A} q(B)^{-1^{\#(B)} - \#(A) + 1} \quad (3.69)$$

L'équation 3.68 est appelée décomposition canonique de  $m(\cdot)$ . Au final la fonction  $w(\cdot)$  est appelée fonction de poids conjonctifs (FDPC) et représente la même quantité d'information

que  $m(\cdot)$ . La transformation inverse est d'ailleurs possible par définition de la décomposition canonique. Notons qu'il est plus commode de réécrire l'équation 3.69 sous la forme :

$$\ln w(A) = - \sum_{B \supseteq A} (-1)^{\#(B) - \#(A)} \ln q(B) \quad (3.70)$$

Enfin, en utilisant la règle disjonctive (cf. 4.2) et sur un raisonnement équivalent, une autre décomposition canonique pour des FDM non-normées est possible. Une fonction, appelée fonction de poids disjonctifs (FDPD), notée  $v(\cdot)$  est ainsi définie :

$$v : 2^\Omega \longrightarrow [0, \infty) \quad (3.71)$$

$$\ln v(A) = - \sum_{B \supseteq A} (-1)^{\#(A) - \#(B)} \ln b(B) \quad (3.72)$$

Nous soulignons que ces deux dernières représentations ne sont pas possibles pour l'ensemble des FDM  $\mathfrak{B}^\Omega$ . De nombreux compléments d'information sur ces deux fonctions sont disponibles également dans Denœux [Denœux 08].

### 3.4.5 Notion de distinction

Shafer [Shafer 76], qui a introduit la notion de distinction entre des éléments d'évidence, n'en a pas donné une définition formelle. La distinction rejoint en quelque sorte la notion d'indépendance conditionnelle entre variables aléatoires. Les éléments d'évidence sont considérés distincts si la construction des croyances selon un élément d'évidence ne restreint pas la construction des croyances pour les autres éléments. Il n'est pas plus approximatif de considérer des sources distinctes que des variables aléatoires indépendantes. En particulier, des sources provenant d'une même expérience ne peuvent être considérées comme distinctes, ce qui est le cas pour nos extracteurs tous dépendant des données image.

Dans [Smets 06], Smets propose également une définition de la distinction entre variables crédales : soient deux variables  $Z_1$  et  $Z_2$ , dont deux FDM  $m^{Z_1}$  et  $m^{Z_2}$  représentent les croyances relatives à la valeur réelle des variables. Supposons ces croyances forgées à partir de deux éléments d'évidence différents pour chaque fonction. Ces deux éléments d'évidence sont distincts si et seulement si une FDM  $m^{Z_1 \times Z_2}$  peut être calculée à partir de  $m^{Z_1}$  et  $m^{Z_2}$  en appliquant la règle conjonctive :

$$m^{Z_1} \otimes m^{Z_2}(A) = \begin{cases} m^{Z_1}(B) m^{Z_2}(C) & \text{si } A = B \times C, B \subseteq Z_1, C \subseteq Z_2 \\ 0 & \text{sinon} \end{cases} \quad (3.73)$$

Dans la suite du mémoire, nous parlerons de sources distinctes.

### 3.4.6 Opération d'affaiblissement

Quand une connaissance *a priori* permet de savoir qu'une source  $S$  n'est pas fiable, il est possible de réduire l'influence de cette source dans le processus de fusion en opérant un affaiblissement sur la FDM  $m[S](\cdot)$  relative à cette source. Un coefficient d'affaiblissement  $\alpha \in [0, 1]$  est défini, et la masse affaiblie est alors :

$$m[\alpha, S](X) = \begin{cases} (1 - \alpha) m[S](X) & \text{si } X \neq \Omega \\ (1 - \alpha) m[S](X) + \alpha & \text{si } X = \Omega \end{cases} \quad (3.74)$$

Plus  $\alpha$  est grand, plus l'affaiblissement est important. Grâce à l'affaiblissement, la FDM d'une source non fiable est transformée en une autre FDM plus proche de  $\Omega^0(\cdot)$ , elle est moins engagée.

### 3.4.7 Transformation pignistique

Le TBM sépare le traitement de l'information en deux niveaux :

- le niveau crédal dans lequel les croyances sont modélisées, combinées et traitées,
- le niveau pignistique, postérieur à l'étape crédibiliste, et dans lequel une décision est prise par rapport aux croyances établies.

La figure 3.2 représente la chaîne de traitement globale où *bel* est la fonction de crédibilité et *betP* est la **transformation pignistique**<sup>6</sup> qui transforme les croyances en probabilités. Comme

**Figure 3.2 :** Processus décisionnel dans le TBM.

nous l'avons expliqué, il est intéressant d'utiliser les fonctions de croyance pour gérer des données imprécises et mettre au point des processus complexes de fusion. En revanche pour prendre une décision, les probabilités sont mieux adaptées car, à ce stade, l'imprécision ne fait que bruyier le traitement et doit être éliminée, d'où l'intérêt de se doter d'une transformation permettant le passage des croyances vers les probabilités.

Le traitement de l'information au niveau crédal doit aboutir à la création d'une fonction de masse  $m_{\odot}(\cdot)$  synthétisant l'ensemble des connaissances à l'instant où doit être prise la décision. Pour ramener  $m_{\odot}$  vers un espace probabiliste, Smets [Smets 90] se base alors sur le principe de la raison insuffisante. Ce principe stipule qu'en l'absence de raison de privilégier une hypothèse plutôt qu'une autre, les hypothèses sont toutes supposées équiprobables. Ainsi, pour tout élément focal  $A$  de  $m_{\odot}(\cdot)$  la masse sera redistribuée uniformément sur les éléments de  $A$ . Cette idée rejoint celle de Dempster pour la normalisation des fonctions de masses par le conflit. La transformation se définit alors :

$$BetP(\{\omega_1\}) = \sum_{A \subseteq \Omega} \frac{|\{\omega_1\} \cap A|}{|A|} \frac{m_{\odot}(A)}{1 - m_{\odot}(\emptyset)} \quad (3.75)$$

La formule est donnée pour  $\{\omega_1\}$  et  $m_{\odot}(\cdot)$  mais est valable pour tout ensemble et toute FDM. Notons qu'à travers ce procédé, une seule distribution de probabilité est associée à une fonction

<sup>6</sup>L'adjectif pignistique introduit par Smets vient du latin pignus, signifiant "pari".

de masse, tandis qu'une infinité de fonctions de masse est associée à une distribution de probabilité. Cela reflète d'une part que l'espace des fonctions de masse est bel et bien plus vaste que celui des distributions de probabilité, mais aussi d'autre part que la transformation pignistique occasionne une perte ou un filtrage d'information.

Cette transformation est indispensable à l'utilisation des fonctions de croyance dans un filtre particulière. A ce stade, seuls restent quelques points de modélisation des fonctions de masse et du cadre de discernement pour mettre au point le filtre particulière avec fusion crédibiliste.

### 3.4.8 Autres processus crédibilistes

Le formalisme des fonctions de croyance contient de nombreux autres outils mathématiques. Comme nous n'aurons pas recours à ces éléments dans notre approche, nous nous permettons de ne citer que quelques uns d'entre eux, sans énoncer leur définition afin de ne pas alourdir la lisibilité du mémoire :

- Le conditionnement, qui est le pendant crédibiliste du conditionnement probabiliste,
- Le théorème de Bayes généralisé (TBG). De même que la règle de Bayes le fait pour les probabilités, le TBG fonctionne à l'aide d'une hypothèse  $\theta$  appartenant au cadre de discernement  $\Theta$  et évaluée conditionnellement à une autre hypothèse  $\omega$  appartenant à  $\Omega$ . Le TBG permet alors de renverser le conditionnement et d'évaluer  $\omega$  sachant  $\theta$ ,
- Le raffinement et l'extension vide. Le raffinement permet de passer d'un cadre de discernement à un autre où les hypothèses du premier cadre sont divisées en sous-hypothèses dans le nouveau cadre. Si de nouvelles croyances sont affinées dans le nouveau cadre, l'extension vide permet de mettre à jour les croyances dans l'ancien cadre,
- Le grossissement et la marginalisation. Le grossissement est l'opération duale du raffinement, où les hypothèses d'un nouveau cadre sont des agrégats d'hypothèses de l'ancien cadre. La marginalisation est un cas particulier de grossissement où le cadre de départ est un espace produit  $\Omega \times \Theta$ . Afin d'exprimer les croyances sur  $\Omega$  seulement,  $\forall \theta$  les hypothèses  $\omega$  sont alors regroupées pour déterminer  $m(\omega)$  ( $m \in \mathfrak{B}^\Omega$ ).

Récemment, Ramasso [Ramasso 07] a soutenu une thèse appliquant les fonctions de croyance à un problème d'analyse de séquences d'images sportives. L'état de l'art présenté dans le cadre de ses travaux permet un tour d'horizon de nombreux outils crédibilistes.

## 3.5 Approche proposée

Dans cette partie, nous expliquons comment le formalisme des fonctions de croyance est concrètement appliqué au problème de suivi d'objets. Nous donnons également quelques résultats génériques concernant l'approche afin de valider l'intérêt de la fusion dans un filtre particulière et le recours aux fonctions de croyance.

### 3.5.1 Définition du cadre de discernement $\Omega$

Un des problèmes les plus importants dans la mise en place d'une approche crédibiliste est la définition du cadre de discernement dans lequel l'algorithme va évoluer. Le nombre d'hypothèses doit être en général assez limité car la taille de  $2^\Omega$  croît exponentiellement avec celle de  $\Omega$ . Cela dit, en général seuls quelques sous-ensembles de  $\Omega$  présentent un intérêt réel pour le problème traité, ce qui allège les traitements.

– **Définitions des hypothèses du cadre de discernement** Comme nous l'évoquions pour le filtre particulaire avec fusion possibiliste 3.3.3, définir un cadre de discernement à partir des variables  $Y_t$  et  $X_t$  n'est pas raisonnable car la cardinalité du cadre ainsi défini serait trop élevée. Or, en examinant le sens donné à la variable  $Y_t|X_t^{(i)}$ , il apparaît que le filtre particulaire a besoin d'évaluer l'adéquation de toutes les observations  $Y_t$  mesurées dans l'imagerie définie par  $X_t^{(i)}$  avec les données du modèle. Nous proposons donc de définir une hypothèse  $\omega_1$  représentant ce niveau d'adéquation.

L'idée est de construire des croyances de sorte que la transformée pignistique pour l'hypothèse  $\omega_1$  corresponde à la vraisemblance  $p(Y_t|X_t^{(i)})$ , c'est-à-dire :

$$p(Y_t|X_t^{(i)}) = \text{BetP}(\{\omega_1\}) \quad (3.76)$$

Pour parvenir à ce but, les autres hypothèses du cadre doivent être antagonistes pour respecter le principe d'exclusion mutuelle. Nous définissons donc le cadre  $\Omega = \{\omega_1, \omega_2, \omega_3\}$  composé de trois hypothèses :

- $\omega_1$  : l'objet suivi se trouve dans l'imagerie relative à  $X_t^{(i)}$ .
- $\omega_2$  : l'imagerie relative à  $X_t^{(i)}$  contient une partie du fond de l'image.
- $\omega_3$  : l'imagerie relative à  $X_t^{(i)}$  contient un objet autre que notre cible et indépendant du fond de la scène.

La figure 3.3 retranscrit visuellement ces différentes hypothèses. Pour chaque source d'informa-

**Figure 3.3 :** Exemple d'imageries correspondant à différentes valeurs de  $X_t^{(i)}$ . Dans le cadre du filtrage particulaire, les imageries sont tirées au hasard selon la densité d'importance et peuvent correspondre à ces différents cas. Pour le cadre vert, la situation correspond à celle décrite par l'hypothèse  $\omega_1$ , cette hypothèse doit donc se voir attribuer un crédit très important dans un tel cas de figure. Le raisonnement est analogue pour le cadre rouge vis-à-vis de  $\omega_2$  et pour le cadre bleu vis-à-vis de  $\omega_3$ .

tion, une FDM est définie sur le cadre  $\Omega$ . Après fusion des différentes FDM par une règle de combinaison, la probabilité pignistique peut être calculée pour chacune de ces hypothèses, mais seule celle concernant  $\omega_1$  revêt un intérêt, car c'est elle qui sémantiquement correspond à  $Y_t|X_t$ .

D'après le sens donné aux hypothèses de  $\Omega$ , toutes les situations possibles sont couvertes par  $\Omega$ , et deux hypothèses ne peuvent se réaliser conjointement :  $\forall i, j \{\omega_i\} \cap \{\omega_j\} = \emptyset$ . Néanmoins,

si une imagerie contient, par exemple, pour une moitié l'objet-cible et pour l'autre moitié du fond de la scène, alors  $\{\omega_1\}$  et  $\{\omega_2\}$  peuvent être deux éléments focaux de la fonction de masse<sup>7</sup>. Cette situation paradoxale ne pose pas de problème car elle est canalisée à deux niveaux. En premier lieu, si l'objet n'est qu'à moitié présent dans l'imagerie, il y aura au moins une source pour laquelle le crédit attribué à  $\{\omega_1\}$  sera significativement plus faible que pour une l'imagerie correspondant exactement à l'objet. En second lieu, dans cette situation la valeur de l'ignorance  $m(\Omega)$  est automatiquement plus faible (cf. equation (3.51)), et donc d'après la formule (3.75) la valeur de la probabilité pignistique sera également plus faible. En somme, l'adéquation observation/modèle est bien graduellement évaluée.

Nous aurions pu nous contenter de ne définir qu'une seule autre hypothèse antagoniste représentant  $\{\bar{\omega}_1\}$ , c'est-à-dire l'hypothèse "l'imagerie contient quelque chose d'autre que l'objet ciblé". Nous avons choisi de diviser cette hypothèse en deux sous-hypothèses  $\{\bar{\omega}_1\} = \{\omega_2, \omega_3\}$  à cause du caractère imprécis de certaines sources. Les paragraphes suivants justifient ce choix.

#### – Sélection des ensembles crédités d'après la mise en correspondance entre modèle et observations

En théorie, lors de la construction des fonctions de masse, chaque sous-ensemble de  $\Omega$  peut être crédité à hauteur d'une certaine valeur en fonction des données observées. En pratique, seul un nombre très limité d'entre eux sont effectivement crédités. En effet, cette limitation permet d'une part d'alléger les temps de calcul et, d'autre part, certains sous-ensembles ont une signification telle que, par rapport au problème posé, il est impossible de concevoir un modèle en rapport avec eux. C'est le cas, entre autres, de  $\{\omega_3\}$ . Comment se doter d'un modèle représentant tout autre objet que celui suivi ? Il faudrait se doter de modèles pour chaque autre type d'objet pouvant intervenir dans la scène, ce qui dans le contexte des séquences routières ne paraît pas raisonnable.

En conséquence, seuls deux sous-ensembles peuvent être crédités pour une source donnée  $S_j$  lors de la construction de la fonction de masse  $m[S_j]$  relative à cette source. Ces ensembles sont :

- le singleton  $\{\omega_1\}$  ou l'ensemble  $\{\omega_1, \omega_3\}$  :
  - le singleton  $\{\omega_1\}$  : il s'agit de l'ensemble le plus naturel à créditer, celui qui symbolise la présence de l'objet-cible.
  - l'ensemble  $\{\omega_1, \omega_3\}$  : cet ensemble symbolise l'ensemble des objets indépendants du fond de la scène. Il est crédité par des sources imprécises qui peuvent distinguer ces objets du fond, mais pas catégoriser ces objets. Dans les séquences routières, cela représente l'ensemble des véhicules observables.

Le choix du singleton  $\{\omega_1\}$  ou l'ensemble  $\{\omega_1, \omega_3\}$  est effectué en fonction de la précision du pouvoir discriminant des sources (cf. 3.5.2, 3.5.2 et 3.5.2). Ce sont ces différences de précision qui expliquent le besoin de définir l'hypothèse  $\omega_3$ .

- le singleton  $\{\omega_2\}$ , qui peut être crédité si une partie du fond de la scène revient de manière systématique. Dans les séquences routières, il peut s'agir du bitume qui possède des propriétés visuelles relativement stables.

L'attribution du crédit pour les ensembles cités ci-dessus se fait grâce à la mise en cor-

---

<sup>7</sup>Cette situation n'est pas une situation de conflit. Le conflit intervient entre deux sources et non entre deux hypothèses.

respondance entre observations et modèles. Il existe différentes manières d'effectuer cette mise en correspondance et également différentes manières d'exploiter cette mise en correspondance pour l'obtention de fonctions de masse. Dans la sous-section suivante, nous faisons un rappel des méthodes existantes dans ce domaine, avant de présenter la méthode adoptée dans ces travaux de thèse.

### 3.5.2 Calcul des masses de croyance pour chaque source

#### Approches existantes

Pour chaque source d'information  $S_j$ , une FDM  $m[S_j](\cdot)$  est définie d'après les informations fournies par l'extracteur associé à la source  $S_j$ . Plusieurs auteurs ont déjà abordé la question de la modélisation des croyances aboutissant à l'obtention de FDM. Une première famille de méthodes utilise des vraisemblances pour construire les masses de croyance. Deux méthodes de référence ont été développées indépendamment par Appriou [Appriou 91] et par Smets [Smets 93] via le TBG. Cette famille d'approches est en réalité surtout intéressante pour des problèmes de classification où les hypothèses  $\omega_i$  correspondent à des classes. Pour une observation  $Y$ , les vraisemblances  $L(\omega_i|Y) = p(Y|\omega_i)$  sont en général connues conditionnellement aux classes grâce à un apprentissage. Les masses peuvent être alors définies ainsi :

$$m[S_j](\{\omega_i\}) = 0 \quad (3.77)$$

$$m[S_j](\{\bar{\omega}_i\}) = \alpha_i(1 - R.L(\omega_i|Y)) \quad (3.78)$$

$$m[S_j](\Omega) = 1 - \alpha_i(1 - R.L(\omega_i|Y)) \quad (3.79)$$

avec  $\alpha_i$  un coefficient d'affaiblissement et  $R$  un paramètre à définir. Appriou propose un second modèle conjointement à celui présenté ci-dessus. Ce second modèle s'exprime de la manière suivante :

$$m[S_j](\{\omega_i\}) = \alpha_i(R.L(\omega_i|Y)) \quad (3.80)$$

$$m[S_j](\{\bar{\omega}_i\}) = \alpha_i(1 - R.L(\omega_i|Y)) \quad (3.81)$$

$$m[S_j](\Omega) = 1 - \alpha_i \quad (3.82)$$

D'après l'auteur, ces deux modèles possèdent en pratique des comportements similaires. Quoiqu'il en soit, ce schéma ne s'applique pas à notre cas car nous n'avons pas de telle correspondance entre les hypothèses du cadre de discernement et le problème de suivi.

L'autre famille de méthodes repose sur des distances entre observations et modèles. Denoeux [Denoeux 95] a proposé un modèle de construction inspiré des k-ppv. Toujours en partant d'un problème de classification, cette méthode suppose connu un ensemble de vecteurs  $\{Y^k\}_{k=1}^{N_{vec}}$  ainsi que les étiquettes des classes de chacun de ces vecteurs. Pour chaque vecteur de cet ensemble, il est possible de définir une FDMS de la manière suivante :

$$m^{k,i}(\{\omega_i\}) = \alpha_i \Phi_i(d(Y, Y^k)) \quad (3.83)$$

$$m^{k,i}(\Omega) = 1 - \alpha_i \Phi_i(d(Y, Y^k)) \quad (3.84)$$

avec  $\alpha_i$  un coefficient d'affaiblissement et  $\Phi_i$  une fonction telle que :

$$\Phi_i : \mathbb{R} \longrightarrow [0, 1] \quad (3.85)$$

$$\Phi_i(0) = 1 \quad (3.86)$$

$$\lim_{d \rightarrow \infty} \Phi_i(d) = 0 \quad (3.87)$$

La FDM ainsi obtenue est simple puisqu'elle possède comme éléments focaux :  $\{\omega_i\}$  et  $\Omega$ . Cette construction se justifie par le fait que la distance  $d(Y, Y^k)$  ne peut donner un avis que sur  $\{\omega_i\}$ , tout le reste de la croyance doit être affecté à l'ignorance. Une fois ces FDMS calculées, l'auteur propose des les agréger par la règle de Dempster d'abord par rapport à l'ensemble des vecteurs étiquetés (indice  $k$ ), puis par rapport aux hypothèses (indice  $i$ ). Le processus donne au final une FDM évaluée pour la source ayant produit l'observation  $Y$ .

### Approche proposée pour le problème de suivi

– **Mise en correspondance observations/modèles** La mise en correspondance observations/modèles s'effectue à l'aide de la distance de Bhattacharyya  $d_{Bhattacharyya}(\cdot, \cdot)$ . Chaque méthode d'extraction employées dans nos travaux, c'est à dire chaque source, fournit des observations pouvant être présentées sous forme vectorielle, ce qui rend l'utilisation de la distance de Bhattacharyya possible. De manière analogue aux possibilités, nous notons  $y_t^{S_j, k} | X_t^{(i)}$  le vecteur d'observations effectuées à  $t$  sur l'imagette définie selon  $X_t^{(i)}$  par la source  $S_j$  en fonction du modèle relatif à un élément focal  $A_k \subsetneq \Omega$ .

Dans le cas où il s'agit de construire les croyances pour  $A_k = \{\omega_1\}$  ou  $A_k = \{\omega_1, \omega_3\}$ ,  $y_t^{S_j, k} | X_t^{(i)}$  est mis en correspondance avec les caractéristiques extraites sur l'imagette correspondant à l'objet suivi prélevée à l'instant initial à l'aide de  $X_0$ . Nous notons ces caractéristiques  $y_0^{S_j, k} | X_0$ . En revanche, dans le cas où il s'agit de construire les croyances pour  $A_k = \{\omega_2\}$ , il faut effectuer une mise en correspondance avec un modèle du fond de la scène. Pour obtenir un tel modèle, nous étendons légèrement notre *a priori* à la connaissance d'un vecteur  $X_0^{\{\omega_2\}}$ . On note ces caractéristiques  $y_0^{S_j, k} | X_0^{\{\omega_2\}}$ .

D'après la définition de la distance de Bhattacharyya, la mise en correspondance pour la particule  $X_t^{(i)}$ , la source  $S_j$  et l'élément focal  $A_l$  est une valeur réelle appartenant à l'intervalle  $[0, 1]$ . Nous la notons  $d_{i,j,l}$ .

– **Construction des FDM** Concernant la construction des FDM, nous adoptons une stratégie similaire à celle de Denceux, mais plus directe, car nous ne disposons pas d'un ensemble de vecteurs comme requis. Des FDMS sont construites pour chaque élément focal  $A_l \neq \Omega$  de  $m[S_j](\cdot)$  au lieu d'être construites pour chaque classe.

Par rapport au modèle de Denceux, nos connaissances sur l'objet sont limitées à une seule imagette de l'objet, la plupart du temps issue de la première image de la séquence. La fusion sur l'indice  $k$  des vecteurs appris n'a donc pas lieu d'être. Nous obtenons donc directement des FDMS de la forme suivante :

$$m^l[S_j](A_l) = \alpha_{l,j} \Phi_{l,j}(d_{i,j,l}) \quad (3.88)$$

$$m^l[S_j](\Omega) = 1 - \alpha_{l,j} \Phi_{l,j}(d_{i,j,l}) \quad (3.89)$$

avec  $A_l$  un élément focal et  $d_{i,j,l}$  la distance traduisant la mise en correspondance des observations avec les données *a priori*. Un indice  $j$  apparaît sur le paramètre  $\alpha_{l,j}$  et la fonction  $\Phi_{l,j}$  afin de marquer la dépendance vis-à-vis d'une source donnée. De manière analogue à Kharbouche et Vannoorenberghe [Kharbouche 04], nous choisissons de plus :

$$\Phi_{l,j}(d) = \exp(-\xi_{l,j}d_{i,j,l}^2) \text{ (modèle gaussien)} \quad (3.90)$$

avec  $\xi_{l,j}$  est un paramètre fixé. La construction des FDM repose en tout sur deux paramètres :  $\alpha_{l,j}$  et  $\xi_{l,j}$ . Ces paramètres dépendent de l'ensemble crédité (l'indice  $l$ ) mais aussi de la source (indice  $j$ ). Notons que dans nos expériences nous prendrons  $\alpha_{l,j} = 1$  dans tous les cas de figure. Quant à la valeur de  $\xi_{l,j}$ , elle est réglée dans nos expériences au cas par cas par l'utilisateur selon un principe de type erreur/correction<sup>8</sup>.

Enfin, notre but est de définir une seule FDM par source. Il faut donc effectuer une pré-fusion entre les FDM relatives à une même source mais à différents éléments focaux  $A_l$ . Nous choisissons d'effectuer cette opération à l'aide de la règle conjonctive, choix classique à ce stade de la modélisation.

Les figures 3.4 et 3.5 résument le processus d'obtention d'une FDM pour une source donnée.

**Figure 3.4 :** Chaîne de traitement pour l'obtention d'une FDMS pour une source et pour un élément focal.

Notons qu'une fusion de FDMS est très rapide à calculer. Nous donnons à présent quelques détails relatifs aux différentes catégories de sources d'information pour le calcul des FDM.

### FDM texture-couleur

En suivi d'objets, il est difficile de dire, d'un point de vue général, quels sont les éléments focaux d'une source texture-couleur. Le choix le plus délicat est toujours entre le singleton  $\{\omega_1\}$  et l'ensemble  $\{\omega_1, \omega_3\}$ . Ces éléments sont crédités à partir de la distance entre les observations

<sup>8</sup>Dans l'absolu, il est préférable d'avoir recours à une méthode d'optimisation de type algorithme génétique pour déterminer cette valeur.

**Figure 3.5 :** Chaîne de traitement pour l'obtention d'une FDM pour une source.

et le modèle de l'objet-cible lui même. Si le crédit est attribué à  $\{\omega_1\}$  alors cela signifie que l'on considère que le modèle texture-couleur est suffisamment discriminant pour qu'aucun autre objet de la scène puisse s'y apparenter. Si le crédit est attribué à  $\{\omega_1, \omega_3\}$  alors cela signifie que la source n'est pas suffisamment précise pour bien distinguer deux objets distincts.

Dans les séquences routières, deux véhicules peuvent avoir la même couleur et des textures similaires. Néanmoins les mesures de texture-couleur sont dans la pratique assez fines pour distinguer deux véhicules différents de même couleur. Il est donc raisonnable de choisir  $\{\omega_1\}$  comme élément focal d'une source texture-couleur, à condition tout de même d'être assez sélectif sur le modèle, c'est-à-dire de choisir  $\xi_{l,j}$  de sorte que le mode de la gaussienne  $\Phi_{l,j}$  soit étroit. Le cas limite est bien sûr la situation de fouillis total où deux véhicules de même modèle, de même marque et de même couleur se trouvent côte à côte dans la scène filmée. Ces cas restent peu probables, et comme nous l'évoquions en 2.7.2, le fouillis total doit être géré à un autre niveau algorithmique.

Concernant l'autre élément focal restant  $\{\omega_2\}$ , là encore les raisonnements ne sont valables que sur une catégorie de séquences données. Dans notre cas, il est possible de caractériser le bitume grâce aux textures-couleurs. Dans cette situation, il convient donc de créditer  $\{\omega_2\}$ . D'après les extracteurs employés dans ces travaux de thèse, il n'apparaît d'ailleurs pas possible de caractériser le fond de la scène à l'aide de sources d'une autre nature que la texture-couleur. L'attribution de crédit à cet élément focal pour les autres types de FDM n'est donc pas abordé.

### FDM forme

Pour la forme, le même raisonnement que pour la texture-couleur est applicable mais les conclusions diffèrent. Il semble en effet difficile de créditer  $\{\omega_1\}$  car les voitures sont toutes de formes relativement similaires. Nous choisissons donc, via le modèle de l'objet, de créditer  $\{\omega_1, \omega_3\}$ . Il serait possible de définir un modèle pour des classes de véhicules très différentes des automobiles, comme un poids lourd si une imagerie d'un tel véhicule est donnée *a priori*. Ces modèles permettrait alors de créditer  $\{\omega_3\}$ . Cela reste relativement dangereux à cause du mouvement de changement d'échelle, en effet avec la perspective un poids lourd vu de loin a une

forme très similaire à celle d'une voiture vue de près. Pour des motocyclettes, ce problème ne se pose pas mais leur forme étant singulièrement différente de celle d'une voiture, il y a peu de risques de confusion. Définir un modèle spécifique à chaque type d'objet alourdirait le processus sans améliorer notablement son efficacité. Au final, un seul élément focal est retenu pour une source de forme :  $\{\omega_1, \omega_3\}$ .

### FDM mouvement

Les sources de type mouvement sont les seules pour lesquelles il est possible de généraliser un peu le choix des éléments focaux au delà de l'application au suivi de véhicules. En effet, il est extrêmement rare qu'un mouvement soit suffisamment discriminant pour distinguer un objet en particulier. Le rôle d'une source mouvement est de souligner la présence d'objets se détachant clairement du fond de la scène. En ce sens, si un objet en mouvement est détecté alors l'ensemble  $\{\omega_1, \omega_3\}$  doit être crédité. A l'inverse, une absence de mouvement ne prouve pas qu'il s'agisse du fond de la scène car l'objet suivi peut naturellement stopper son mouvement pour quelques instants. Dans nos séquences, cette situation est très fréquente en milieu urbain (feux rouges et autre signalisation) mais aussi sur autoroute (embouteillage). De même que pour une source de forme, seul  $\{\omega_1, \omega_3\}$  constitue un élément focal pour une source mouvement.

### Bilan sur l'obtention des FDM

Au final, les remarques précédentes simplifient grandement l'obtention des FDM. Seules les sources de type texture-couleur nécessitent une pré-fusion entre une FDMS créditant  $\{\omega_1\}$  et une seconde FDMS créditant  $\{\omega_2\}$ . Toutes les autres sources sont obtenues directement en appliquant le modèle gaussien de la fonction  $\Phi_{l,j}(\cdot)$ .

### 3.5.3 Fusion des FDM et intégration de la masse fusionnée au filtre particulière

Une fois que nous disposons d'une FDM par source, il suffit d'appliquer une règle  $\odot$  à ces FDM pour opérer la fusion de données :

$$m_{\odot} = \odot_{j=1}^M m[S_j] \quad (3.91)$$

Ainsi,  $m_{\odot}$  contient l'information de fusion. Il suffit ensuite d'appliquer la transformation pignistique à cette FDM pour revenir à des probabilités et faire le lien avec le filtrage particulière, comme évoqué en 3.5.1.

Nous résumons à présent toutes les démarches nécessaires à l'utilisation du formalisme des fonctions de croyance pour un suivi par filtrage particulière :

- définition d'un cadre de discernement adapté,
- choix des éléments focaux pour les sources,
- obtention des FDM pour chaque source,
- fusion des sources à l'aide d'une règle de combinaison,
- calcul de la transformation pignistique à partir des croyances issues de la fusion afin de les transcrire en probabilités,
- acceptation de cette probabilité en tant qu'estimation de la vraisemblance globale du fait de la définition donnée à  $\omega_1$ .

Il est à présent possible de donner la procédure algorithmique correspondant au filtre particulaire avec fusion crédibiliste (FPFC) (cf. algorithme 7).

---

Algorithme : Filtre particulaire simple avec fusion crédibiliste, FPFC

---

**Initialisation**,  $t = 0$  :

Échantillonner  $N$  particules selon une distribution  $p(X_0)$  et poser  $\forall i w_0^{(i)} = \frac{1}{N}$

**pour**  $t \geq 1$  **faire**

**pour**  $i = 1$  à  $N$  **faire**

    Échantillonner les particules selon  $X_t^{(i)} \sim p(X_t|X_{t-1})$

**pour**  $j = 1$  à  $M$  **faire**

      Évaluer la fonctions de masses de la source  $S_j$  sur un  $\Omega$

**fin pour**

    Fusionner les fonctions de masse selon une règle donnée  $\odot$

    Appliquer la transformation pignistique à la fonction de masse issue de la fusion :

$$BetP(\{\omega_1\}) = \sum_{A \subseteq \Omega} \frac{|\{\omega_1\} \cap A|}{|A|} m_{\odot}(A)$$

    Accepter  $BetP(\{\omega_1\}) \approx p(Y_t|X_t^{(i)})$

    Mettre à jour les poids selon  $w_t^{(i)} \propto w_{t-1}^{(i)} p(Y_t|X_t^{(i)})$

    Normaliser les poids selon  $\tilde{w}^{(i)} = \frac{w^{(i)}}{\sum_{j=1}^N w^{(j)}}$

**fin pour**

**si**  $\hat{N}_{eff} > \text{seuil}$  **alors**

    Rééchantillonner en tirant  $N$  nouvelles particules  $\tilde{X}_t^{(j)}$  parmi les  $X_t^{(i)}$  selon une probabilité  $\tilde{w}_t^{(i)}$ , et leur affecter les poids  $\tilde{w}_t^{(j)} = \frac{1}{N}$

    Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t|Y_{0:t}) = \sum_{j=1}^N \delta_{\tilde{X}_t^{(j)}}(X_t) \frac{1}{N}$

**sinon**

    Estimer l'état caché en utilisant l'estimation de la distribution  $\hat{p}(X_t|Y_{0:t}) = \sum_{i=1}^N \delta_{X_t^{(i)}}(X_t) \tilde{w}_t^{(i)}$

**fin si**

**fin pour**

**Fin**

---

Algorithme 7: Filtre particulaire simple avec fusion crédibiliste.

### 3.5.4 Tests de suivi génériques

Nous n'avons pas la prétention de fournir un algorithme de suivi visuel générique, mais la forme sous laquelle se présente l'algorithme 7 reste utilisable dans divers contextes de suivi. L'implémentation concrète de notre approche est de toute façon *ad hoc* puisque la modélisation des FDM est *ad hoc* et que les extracteurs de caractéristiques sont eux-mêmes *ad hoc*. Nous présentons ici quelques résultats dans un cadre plus large que notre application de suivi de véhicules. Ces tests ont pour but de mettre en évidence la pertinence de la fusion d'informations proposée.

Pour ces tests, trois sources de catégories différentes sont utilisées : une source texture-couleur, une source de forme et une source de mouvement. Les trois exemples de séquence comportent chacun un certain nombre d'événements perturbateurs : changement d'illumination, mouvement rotationnel, occultation et fouillis.

Nous comparons dans un premier temps notre approche FPFC à de simples filtres particu-

lares fondés uniquement sur une source. Cela nous permet de montrer qu'une approche mono-source est moins robuste qu'une approche multi-source, ce qui intuitivement paraît logique. Le FPFC est ensuite comparé aux deux grandes autres approches évoquées dans ce chapitre que sont le FP avec fusion bayésienne et le FP avec fusion possibiliste. La règle de combinaison choisie pour le FPFC est la règle conjonctive  $\odot$ .

*Nota bene* : Dans la suite du manuscrit, les résultats sont principalement présentés en fonction des méthodes de fusion. Nous désignons par algorithmes-coopérateurs, les extracteurs d'information utilisés pour chaque source, ainsi que le filtre particulière. Pour chaque vidéo expérimentée, les réglages de ces algorithmes-coopérateurs seront conservés pour chaque méthode de fusion afin de permettre un comparatif clair. Toutefois, la qualité des résultats dépend naturellement aussi de ces réglages. Si à un moment donné, le choix des paramètres des algorithmes-coopérateurs pénalise une approche, nous le soulignerons et donnerons une estimation qualitative ou quantitative des résultats de l'approche pour un jeu de paramètres différent. De manière générale, les paramètres sont choisis de sorte à ne pas favoriser implicitement une méthode plutôt qu'une autre. Par exemple, l'écart-type de la densité  $p(X_t|X_{t-1})$  est toujours choisi suffisamment grand pour que l'algorithme ait suffisamment de liberté et qu'une estimation correcte ne soit pas le fruit d'heureuses circonstances.

### Cas de sources défaillantes

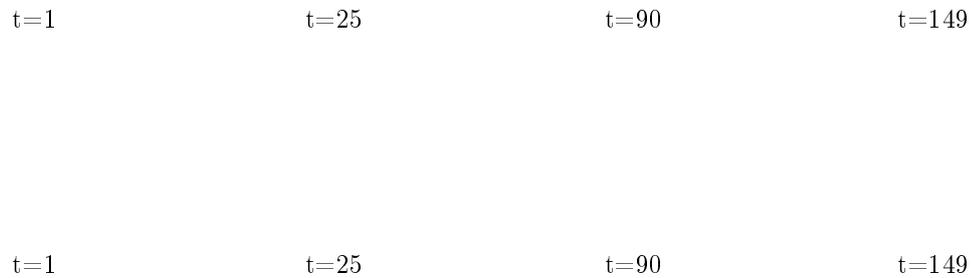
Nous présentons ici une séquence comportant un événement perturbateur : un changement brutal d'illumination de la scène. Ce changement a pour conséquence de rendre la source texture-couleur inopérante car les observations postérieures au changement sont trop éloignées du modèle texture-couleur. Notons que, comme évoqué en 2.7, à chaque type de source correspond un événement susceptible de la rendre inopérante. Nous définissons dans ce mémoire une **source défaillante** comme étant quasiment incapable de créditer un ensemble autre que  $\Omega$ . La FDM correspondant à cette source est donc proche de  $\Omega^0$ . La défaillance est la forme la plus sévère d'un défaut plus général d'une source d'information que nous appelons **indétermination**<sup>9</sup>. Plus une source est indécise, et moins elle cible en particulier une hypothèse du cadre de discernement. Nous reviendrons sur ces termes en 4.4.1.

La séquence présentée ici possède les caractéristiques suivantes : basse résolution et qualité moyenne (léger bruit). L'objet suivi est une balle de tennis, dont la couleur est facilement repérable par la source texture-couleur. Cependant, une fois le changement de luminosité intervenu, cette couleur est changée et n'est plus reconnue par la source. La forme de l'objet étant simple et la caméra immobile, les sources de forme et de mouvement fonctionnent relativement bien tout au long de la séquence.

La figure 3.6 montre la perte de suivi du filtre particulière texture-couleur. Le FPFC parvient à maintenir le suivi en s'appuyant sur les autres sources. Une analyse plus fine des résultats est disponible dans la figure 3.7 où les taux de suivi<sup>10</sup> sont représentés dans trois graphes. Chaque graphe compare un filtre mono-source au filtre multi-source. Le premier graphe compare le filtre texture-couleur au FPFC et au Camshift (version d'OpenCV). Le Camshift étant lui aussi essen-

<sup>9</sup>Cette notion est à opposer à celle d'engagement évoquée en 3.4.2.

<sup>10</sup>La définition du taux de suivi est disponible en 2.6.3, eq(2.31).



**Figure 3.6 :** En haut : suivi par filtre particulaire texture-couleur, en bas : suivi par FPFC (3 sources).

**Figure 3.7 :** Évaluation des performances dans le cas d'une source défaillante.

tiellement basé sur une seule source texture-couleur (densité de couleur) et étant une approche de référence, il nous est apparu opportun de l'ajouter au comparatif. Le taux du filtre particulaire texture-couleur baisse brutalement dès le changement d'illumination matérialisé par le trait magenta. Le Camshift subit aussi "de plein fouet" les effets du changement mais parvient

à maintenir un suivi partiel de la balle de tennis. En réalité, les pixels du cœur de la balle sont moins affectés que ceux en périphérie par le changement, et c'est ce cœur qui est suivi par Cam-shift.

Sur le deuxième graphe, le filtre particulaire basé sur la forme "décroche" pendant quelques images vers la fin de la séquence. Les cartes de symétrie sont utilisées comme extracteur, or par moments, les changements d'illumination affectent localement la balle de tennis et modifient l'apparence de sa symétrie.

Enfin sur le dernier graphe, le filtre particulaire basé sur le mouvement possède des performances fluctuant avec l'intensité du mouvement de la balle qui, par moments, est immobile. Pendant ces phases sans mouvement, la cohérence temporelle du filtre particulaire rend possible la continuité du suivi, même si le suivi est alors moins précis.

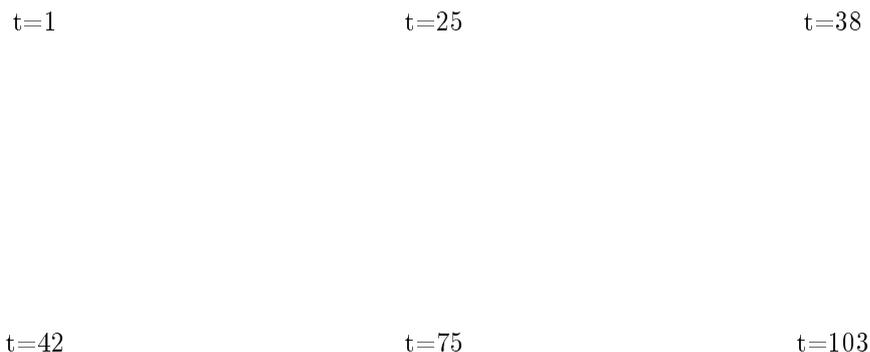
Le FPFC maintient le suivi tout au long de la séquence. Il est intéressant de noter que le FPFC fait mieux que chacune des approches mono-source. En ce sens, la procédure de fusion remplit parfaitement son office en prenant le meilleur des trois sources.

### Cas de sources leurrées

Quand une source peut fournir des informations au moins en partie erronées, cette source est qualifiée de **non fiable**. Comme nous le verrons en 4.4.1, de nombreux travaux ont été consacrés à la fusion d'informations en présence de telles sources. Dans ce mémoire, nous désignons par **source leurrée**, une source non fiable, dont la proportion d'information erronée est potentiellement très élevée. En termes de suivi d'objets, cela se traduit par une source croyant reconnaître l'objet suivi en des positions de l'image où il ne se trouve pas. Dans un tel cas, la source fournit illégitimement un crédit élevé à  $\{\omega_1\}$  ou  $\{\omega_1, \omega_3\}$ . Cela se produit notamment en cas de fouillis où un objet similaire à celui suivi sera crédité. Le problème peut également survenir quand le modèle d'objet n'est pas assez discriminant par rapport au reste de la scène filmée.

Pour discuter de ces cas plus complexes à gérer pour les algorithmes de fusion, nous présentons deux séquences d'images.

– **Premier exemple** La séquence présentée pour cette expérience possède les caractéristiques suivantes : basse résolution et qualité moyenne (léger bruit). L'objet à suivre est une face arrière de véhicule ; la séquence correspond donc à une séquence routière, c'est-à-dire la catégorie de séquence sur laquelle ces travaux de thèse sont destinés à être appliqués. Une situation de fouillis est créée par un autre véhicule de forme et de couleur relativement proches de l'objet-cible. La source texture-couleur et la source de forme fonctionnent convenablement tant que l'autre véhicule est suffisamment loin du véhicule suivi. La source de mouvement est peu efficace, car de nombreux objets en mouvement jalonnent la scène (véhicule constituant le fouillis, véhicules en contre-sens, piétons). La figure 3.8 montre les résultats obtenus par une approche mono-source (filtre particulaire texture-couleur) et l'approche multi-source (FPFC). Ce sont principalement les variations d'illumination qui apportent de la confusion aux sources et finalement créent la situation de fouillis. La séquence ayant été acquise par un temps ensoleillé, la réflexion de la lumière du soleil sur les carrosseries des véhicules fait varier leur rendu visuel. Le véhicule per-



**Figure 3.8** : suivi par FPFC (3 sources).

turbateur est de couleur blanche or, quand celui-ci se situe dans l'ombre portée d'autres objets de la scène, sa couleur devient très proche de celle du véhicule suivi qui est gris. Par moments, les reflets lumineux endommagent la symétrie du véhicule suivi, et la source de forme se rabat alors sur le véhicule perturbateur.

Le phénomène est encore plus clair en examinant les taux de suivi fournis par la figure 3.9. Quand un taux descend à environ 0,5 l'algorithme hésite entre les deux véhicules. La densité de filtrage possède deux modes centrés sur les deux véhicules, et comme l'estimateur effectue un moyennage ces modes (*cf.* équation 3.7), la boîte englobante se retrouve au milieu entre les deux véhicules. C'est le cas notamment pour le FPFC entre les images 40 et 65.

Les courbes relatives au FPFC et au FP texture-couleur sont quasiment superposées ce qui prouve que, sur cet exemple, le FPFC est très influencé par la source texture-couleur. Cela s'explique par le fait que cette source crédite directement le singleton  $\{\omega_1\}$  au lieu de l'ensemble  $\{\omega_1, \omega_3\}$  qui lui est crédité par les deux autres sources. Le poids de la source texture-couleur est donc plus fort au niveau de la transformée pignistique. Le FPFC ne suit les autres sources que si le crédit fourni à  $\{\omega_1, \omega_3\}$  est réellement significatif, ce qui n'est pas le cas dans cet exemple. La figure 3.9 montre également que la source de forme ne produit pas un suivi satisfaisant car le taux avoisine trop souvent des niveaux proches de 0,5. Toutefois, il est assez net que la source de forme produit un meilleur résultat que FPFC, la fusion d'informations est donc ici imparfaite.

– **Deuxième exemple** La séquence présentée ici possède les caractéristiques suivantes : haute résolution et fort bruit en cas de mouvement rapide. L'objet suivi est un ballon de basketball mis en mouvement par un chien.

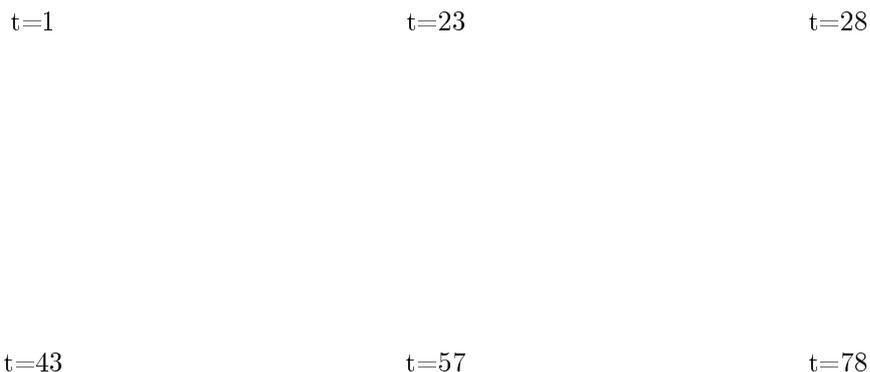
Le chien impulse un mouvement rotationnel au ballon ce qui perturbe la source de forme

**Figure 3.9 :** Évaluation des performances en présence de défaillances pour chaque source.

**Figure 3.10 :** Variabilité visuelle de l'objet suivi.

fondée sur la symétrie par rapport à un axe fixe. Cette source est donc défaillante. Toujours à cause du chien, le mouvement du ballon est imprévisible et peut devenir très rapide par moment. Par ailleurs, la caméra n'est pas fixe mais peu mobile dans cette séquence. La source de mouvement va suivre de façon alternative le ballon et le chien qui sont les deux objets dont les mouvements sont les plus intenses. Ce problème est dû à l'imprécision de la source mouvement qui est donc leurrée dans plusieurs parties de la séquence. La source texture-couleur semble la plus fiable des trois mais le bruit occasionné par les mouvements intenses endommage considérablement le rendu visuel du ballon comme en atteste la figure 3.10. La source texture-couleur est donc elle aussi leurrée par intermittence.

En plus de ces difficultés, le chien occulte le ballon l'espace de quelques images ce qui, par définition, affecte l'ensemble des sources. Étant donné toutes ces difficultés, le FPFC ne parvient pas à maintenir le suivi après occultation, comme le montre la figure 3.11. Sur cette même figure,



**Figure 3.11** : suivi avec FPFC (3 sources).

on observe que l'occultation débute à la deuxième image, qu'elle est totale à la troisième et qu'elle se termine à la quatrième. Comme dans l'exemple précédent, le FPFC n'est pas en mesure de juguler les sources leurrées. Les taux de suivi de la figure 3.12 donne le détail des performances du FPFC.

Pour représenter les taux de suivi, nous avons éprouvé le besoin de matérialiser l'influence de l'occultation. Cette dernière est représentée par le taux d'occultation, qui représente la proportion visible de l'objet par rapport à sa taille totale. Ce taux est estimé par expertise. La définition du taux de suivi est de plus légèrement modifiée, au sens où ce sont les surfaces où l'objet est visible qui sont prises en compte. Ainsi quand l'occultation est totale, le taux de suivi est nécessairement nul. Le FPFC ne produit pas non plus cette fois une fusion d'informations adéquate, car le FP texture-couleur donne de meilleurs résultats. Bien que meilleurs, ces résultats restent tout de même insuffisants car la qualité du suivi après occultation est trop faible. Dans cet exemple, la source de mouvement est prépondérante par rapport aux autres or, comme le chien se déplace plus intensément que le ballon, la source de mouvement oriente le FPFC vers le chien au lieu de l'orienter vers le ballon. A cause du bruit évoqué en figure 3.10, la source texture-couleur n'est pas en mesure de contre-balancer cet effet indésirable.

Au vu des trois expériences présentées, le FPFC propose une solution adaptée au problème de suivi d'objets multi-sources pour des sources normales ou indélicates. En revanche, la présence de sources leurrées constitue une limite de cette approche.

Nous présentons dans les paragraphes qui suivent un comparatif des trois théories sélec-

**Figure 3.12 :** Évaluation des performances en présence de défaillances pour chaque source.

tionnées suite à l'étude bibliographique menée dans les sections précédentes du chapitre. Elles permettent d'opérer une fusion d'informations explicite dans un filtre particulaire. Pour plus d'équité dans la comparaison, nous rappelons que la fusion est effectuée lors de la même étape du FP pour les trois théories : l'estimation de la vraisemblance.

### Comparaison avec l'approche probabiliste

Dans le cadre de la théorie des probabilités, la fusion au stade de la vraisemblance s'effectue sous hypothèse d'indépendance des sources. Elle pourrait être effectuée autrement si nous disposions d'un modèle global de la densité  $p(y_t^{S_1}, \dots, y_t^{S_M} | X_t)$ , ce qui n'est pas le cas. Nous comparons donc le FPFC avec un filtre particulaire utilisant la fusion bayésienne, que nous baptisons filtre particulaire avec fusion bayésienne (FPFB). La figure 3.13 représente les taux de suivi pour ces deux approches sur chacune des séquences évoquées dans cette partie. A cause de l'effet multiplicatif de la fusion bayésienne, le FPFB est très sensible aux sources défaillantes comme le montre les graphes 1 et 3. En effet il suffit qu'une seule des trois sources se retrouve avec une vraisemblance marginale  $p(y_t^{S_j} | X_t)$  proche de 0 pour que la vraisemblance globale soit proche de 0 elle aussi. Les poids sont alors trop équirépartis pour attirer le filtre vers l'objet suivi, d'où la perte de la cible.

A l'inverse, le graphe 2 montre que le FPFB résiste bien aux sources leurrées car, cette fois, l'effet multiplicatif de la fusion bayésienne revêt un caractère sélectif. Le FPFB arrive à filtrer une source leurrée et à avantager une autre qui ne l'est pas. Il accentue la correspondance obser-

**Figure 3.13 :** Comparaison entre l'approche probabiliste FPFB et l'approche crédibiliste FPFC.

vation/modèle quand les sources sont leurrées simultanément. Il fait donc mieux que le FPFC, à l'exception du laps de temps délimité par les images 40 et 48, où les vraisemblances locales sont là aussi significativement faibles pour le FPFB.

Dans le cas de sources défaillantes, l'utilisation brutale de la fusion bayésienne est naïve. De simples procédés peuvent être mis en place pour limiter les désagréments de marginales trop faibles. Pour plus de justesse dans le comparatif, nous éliminons les marginales à faibles valeurs par seuillage avant d'utiliser la fusion bayésienne afin d'écartier les sources clairement défaillantes. La mise en œuvre d'un tel seuillage est explicitée dans le chapitre suivant (*cf.* 4.4.2). En théorie, le seuillage induit une dépendance des observations à une nouvelle variable aléatoire symbolisant l'état d'une source (défaillante ou non, à l'image de l'approche proposée dans [Caron 07a]). En pratique, le procédé paraît assez naturel à mettre en œuvre. La figure 3.14 montre les taux de suivi obtenus avec cette modification. Une nette amélioration est visible pour la séquence du premier et du troisième graphe. En examinant le premier graphe, on constate que le suivi est prolongé jusqu'aux alentours de l'image 100, au lieu de l'image 50 sans seuillage. Tout le problème consiste à trouver un bon seuil, car avec un seuil trop faible des marginales aux valeurs très faibles peuvent fortement influencer le filtre, tandis que un seuil trop haut engendre une perte d'information. D'autres valeurs du seuil permettent d'augmenter encore la durée du suivi et de parvenir à des performance équivalente au FPFC.

Sur le troisième graphe, il apparaît que la fusion probabiliste offre ici des résultats net-

**Figure 3.14 :** Comparaison entre l'approche probabiliste FPFB modifiée et l'approche crédibiliste FPFC.

tement supérieurs en étant capable de reprendre le suivi après occultation. Sur cet exemple, l'opération multiplicative effectuée par la fusion bayésienne permet d'attribuer des poids élevés à des zones où la source texture-couleur et la source mouvement estiment toutes deux que les observations sont proches du modèle de l'objet. Voici un exemple permettant de comprendre ce phénomène :

**Exemple 1.** Pour simplifier nous prenons un filtre à seulement deux particules  $i_1$  et  $i_2$ .

- admettons que la particule  $i_1$  cible le bon objet et que les FDM des sources pour cette particule soient les suivantes :

FDM	$\{\omega_1\}$	$\{\omega_1, \omega_3\}$	$\Omega$	autres ensembles
$m[S_1]$	0,3	0	0,7	0
$m[S_2]$	0	0,5	0,5	0
$m[S_3]$	0	0,1	0,9	0

Les résultats de la fusion<sup>11</sup> et de la transformation pignistique donnent :

FDM	$\{\omega_1\}$	$\{\omega_1, \omega_3\}$	$\Omega$	autres ensembles
$m_{\odot}$	0,3	0,385	0,315	0
$BetP$	0,2642	×	×	×

Après normalisation et en supposant les poids antérieurs équirépartis, le poids  $w_t^{(1)}$  vaut

<sup>11</sup>La valeur de la transformée pignistique est légèrement modifiée par rapport à la formule donnée en 3.4.7 (cf. 5.2.3 pour plus de détails sur cette modification).

0,5028.

En transposant au cas probabiliste, cela revient à avoir les différentes marginales comme suit :

densité	
$p\left(y_t^{S_1} X_t\right)$	0,3
$p\left(y_t^{S_2} X_t\right)$	0,5
$p\left(y_t^{S_3} X_t\right)$	0,1
$p\left(Y_t X_t\right)$	0,015

Après normalisation et en supposant les poids antérieurs équirépartis, le poids  $w_t^{(1)}$  vaut 0,6666.

- admettons que la particule  $i_2$  cible un mauvais objet et que les FDM des sources pour cette particule soient les suivantes :

FDM	$\{\omega_1\}$	$\{\omega_1, \omega_3\}$	$\Omega$	autres ensembles
$m[S_1]$	0,3	0	0,7	0
$m[S_2]$	0	0,5	0,5	0
$m[S_3]$	0	0,05	0,95	0

Les résultats de la fusion et de la transformation pignistique donnent :

FDM	$\{\omega_1\}$	$\{\omega_1, \omega_3\}$	$\Omega$	autres ensembles
$m_{\odot}$	0,3	0,3675	0,3325	0
$BetP$	0,2613	×	×	×

Après normalisation et en supposant les poids antérieurs équirépartis, le poids  $w_t^{(2)}$  vaut 0,4972.

En transposant au cas probabiliste, cela revient à avoir les différentes marginales comme suit :

densité	
$p\left(y_t^{S_1} X_t\right)$	0,3
$p\left(y_t^{S_2} X_t\right)$	0,5
$p\left(y_t^{S_3} X_t\right)$	0,05
$p\left(Y_t X_t\right)$	0,0075

Après normalisation et en supposant les poids antérieurs équirépartis, le poids  $w_t^{(2)}$  vaut 0,3333.

La différence est minimisée par la règle conjonctive. La multiplication effectuée par la fusion bayésienne permet une sélection plus étroite des particules. Grâce à la source  $S_3$ , le FPFb accorde un poids deux fois plus élevé à  $i_2$  qu'à  $i_1$ , ce qui est ici judicieux. En ce sens, le FPFb résiste mieux aux sources leurrées.

En revanche pour la séquence du second graphe la modification reste insuffisante, et le FPFb modifié se focalise sur le mauvais objet sur les images 40 et 48. Cette fois-ci, le changement de valeur du seuil ne change rien au problème. En effet, même une valeur faible d'une marginale possède un pouvoir discriminant qu'il est regrettable de ne pouvoir exploiter. En augmentant le seuil, la perte d'information fait baisser la précision du suivi. En abaissant le seuil, de faibles valeurs de vraisemblance viennent perturber le processus de fusion. Voici un exemple permettant de comprendre ce phénomène :

**Exemple 2.** Pour simplifier nous prenons un filtre à seulement deux particules  $i_1$  et  $i_2$ . Comme dans l'exemple précédent, seule la source  $S_3$  fournit des valeurs différentes pour les deux particules.

- admettons que la particule  $i_1$  cible le bon objet et que les FDM des sources pour cette particule soient les suivantes :

FDM	$\{\omega_1\}$	$\{\omega_1, \omega_3\}$	$\Omega$	autres ensembles
$m[S_1]$	0,01	0	0,99	0
$m[S_2]$	0	0,005	0,995	0
$m[S_3]$	0	0,01	0,99	0

Les résultats de la fusion et de la transformation pignistique donnent :

FDM	$\{\omega_1\}$	$\{\omega_1, \omega_3\}$	$\Omega$	autres ensembles
$m_{\odot}$	0,01	0,0148	0,9752	0
$BetP$	0,0092	×	×	×

Après normalisation et en supposant les poids antérieurs équirépartis, le poids  $w_t^{(1)}$  vaut 0,4600.

En transposant au cas probabiliste, cela revient à avoir les différentes marginales comme suit :

densité	
$p(y_t^{S_1}   X_t)$	0,01
$p(y_t^{S_2}   X_t)$	0,005
$p(y_t^{S_3}   X_t)$	0,01
$p(Y_t   X_t)$	0,000005

Après normalisation et en supposant les poids antérieurs équirépartis, le poids  $w_t^{(1)}$  vaut 0,3333.

- admettons que la particule  $i_2$  cible un mauvais objet et que les FDM des sources pour cette particule soient les suivantes :

FDM	$\{\omega_1\}$	$\{\omega_1, \omega_3\}$	$\Omega$	autres ensembles
$m[S_1]$	0,01	0	0,99	0
$m[S_2]$	0	0,005	0,995	0
$m[S_3]$	0	0,02	0,98	0

Les résultats de la fusion et de la transformation pignistique donnent :

FDM	$\{\omega_1\}$	$\{\omega_1, \omega_3\}$	$\Omega$	autres ensembles
$m_{\odot}$	0,01	0,0247	0,9653	0
$BetP$	0,0108	×	×	×

Après normalisation et en supposant les poids antérieurs équirépartis, le poids  $w_t^{(2)}$  vaut 0,5400.

En transposant au cas probabiliste, cela revient à avoir les différentes marginales comme suit :

densité	
$p(y_t^{S_1}   X_t)$	0,01
$p(y_t^{S_2}   X_t)$	0,005
$p(y_t^{S_3}   X_t)$	0,02
$p(Y_t   X_t)$	0,00001

Après normalisation et en supposant les poids antérieurs équirépartis, le poids  $w_t^{(2)}$  vaut 0,6666.

Si, comme dans le cas de la séquence n°2 ("deux voitures"), toutes les sources doivent être prises en compte pour assurer le suivi, la fusion bayésienne attribue à la particule  $i_2$  une probabilité légèrement supérieure à celle de  $i_1$ , mais surtout un poids deux fois plus important. Après normalisation des poids du filtre particulaire, ce rapport de facteur 2 fera pencher fortement le filtre vers la particule  $i_2$ . Ainsi, la décision repose uniquement sur la source  $S_3$  pour des valeurs de vraisemblance pourtant très faibles, ce qui est ici une mauvaise opération. Si dans l'exemple précédent la fusion bayésienne aidait à converger plus rapidement vers le bon objet, dans le cas présent elle amplifie l'erreur d'appréciation commise par les extracteurs. Pour autant, il n'est pas possible de relever le seuil pour filtrer  $S_3$  car alors toutes les autres seraient également éliminées.

A l'inverse, la règle conjonctive prend en compte la source  $S_3$  sans pour autant lui conférer une importance majeure dans le processus de fusion. Ce comportement prudent est ici judicieux. Comme dans l'exemple précédent, la différence entre les masses après fusion sera faible, mais pas inexistante comme ce serait le cas si la source  $S_3$  était éliminée du processus par seuillage. Même s'il n'est pas bon dans l'absolu de favoriser  $i_2$ , car cette particule cible le mauvais objet, il est inévitable de lui conférer un poids supérieur à celui de  $i_1$  étant donné les informations fournies par les extracteurs. Quand le nombre de particules ne se limite pas à deux, le cumul des poids du filtre particulaire tend à limiter encore plus les conséquences des informations erronées. En effet, le bon objet se trouve généralement plus près de l'ancienne position du filtre et donc un nombre plus élevé de particules est généré autour du bon objet. Il faut alors espérer que dans les images suivantes, les extracteurs reviennent dans le droit chemin et cessent de fournir des informations erronées, sinon n'importe quel algorithme finira par suivre le mauvais objet. Voilà pourquoi dans l'exemple relatif au deuxième graphe, le FPPC est la seule approche capable de maintenir un suivi stable tout au long de la séquence bien que, comme nous l'expliquions précédemment, le résultat du FPPC sur cette vidéo n'atteint pas un score de suivi moyen suffisant pour le moment.

L'analyse des résultats est assez décevante car il apparaît qu'aucune des approches proposées ne parvient à produire la fusion adéquate en présence de sources normales, défaillantes ou leurrées. Néanmoins, deux conclusions très intéressantes sont à souligner :

- les approches crédibiliste et probabiliste apportent des réponses adéquates dans des situations complémentaires,
- les comportements des approches crédibiliste et probabiliste diffèrent de par la nature même des opérations de calcul effectuées.

En travaillant sur cette notion de nature de fusion, il paraît donc possible de proposer une solution alternative qui puisse répondre favorablement à nos exigences. L'élaboration de cette solution constitue tout l'enjeu du chapitre 4.

### Comparaison avec l'approche possibiliste

Nous ajoutons à présent le FPPF (algorithme 6) au comparatif. Plusieurs versions du FPPF sont testées pour différentes règles de combinaison possibilistes. Le nombre de ces règles

étant élevé, nous nous contentons de tester trois des plus classiques d'entre elles. Pour faciliter le comparatif, les résultats sont regroupés dans le tableau 3.1 donnant pour chaque approche la moyenne temporelle du taux de suivi ainsi que l'écart-type correspondant. Les résultats des approches probabilistes et crédibilistes y sont également présentés. Un taux de suivi moyen élevé

Approche utilisée	séquence de la balle de tennis		séquence des deux voitures		séquence du chien et du ballon	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
FPFB	0,3470	0,4205	<b>0,7197</b>	<b>0,1319</b>	0,2396	0,3509
FPFB modifié	0,5514	0,3933	<b>0,7390</b>	<b>0,1180</b>	<b>0,5954</b>	<b>0,2818</b>
FPFP <i>min</i> T-norm	0,0758	0,1792	0,0955	0,2227	0,3367	0,3173
FPFP moyenne arithmétique	0,0722	0,1745	0,0827	0,2557	0,3579	0,3279
FPFP <i>max</i> T-conorm	0,0716	0,1734	0,0805	0,2101	0,2811	0,2812
FPFC	<b>0,7673</b>	<b>0,1471</b>	0,6390	0,1330	0,3195	0,3481

**Tableau 3.1** : Comparatif des différentes approches de fusion dans un filtre particulière par taux de suivi moyen ( $\mu$ ) et l'écart-type du taux ( $\sigma$ ). Les résultats les plus significatifs sont en gras.

permet de savoir si le suivi a été satisfaisant pour la séquence étudiée. Un écart-type élevé traduit un manque de robustesse et une instabilité de l'approche. D'après le niveau de qualité de suivi constaté, le taux moyen de suivi atteint par FPFC pour la vidéo "balle de tennis" est par exemple satisfaisant. Un taux moyen en dessous de 0,7 indique une qualité de suivi insuffisante (sauf cas particulier d'une occultation biaisant la moyenne vers des valeurs plus faibles).

Un écart type moyen supérieur à 0,2 indique une instabilité du suivi tendant à montrer que l'approche échouerait avec une paramétrisation légèrement différente. Notons qu'il n'est pas pertinent de comparer les écarts types de deux approches si les taux moyen ne sont pas suffisamment proches. Enfin, ces statistiques ne reflètent pas tous les aspects qualitatifs d'un suivi. Par exemple, l'effet de moyennage a tendance à gommer des erreurs sporadiques, comme celles du FPFB modifié pour la vidéo "deux voitures". Ces erreurs temporaires constituent pourtant un défaut grave en terme de suivi.

Les résultats des approches possibilistes sont singulièrement mauvais. D'après les investigations menées lors de nos expérimentations, le paramétrage des algorithmes-coopérateurs (extracteurs et filtre particulière) influe plus sur les approches possibilistes que sur les deux autres types d'approches. Ainsi, de bien meilleurs résultats peuvent être obtenus avec un jeu de paramètres différent notamment pour les deux premières séquences, comme le montre la table 3.2. Seule l'approche FPFP par T-conorme max semble vraiment inadaptée. Pour ce même jeu de paramètres, les performances du FPFC ou du FPFB sont légèrement altérées, mais restent un ton au-dessus de celles du FPFP. Étant donné cette constatation, et le fait que le FPFP n'offre pas d'alternative claire en terme de qualité de fusion par rapport au FPFB et au FPFC, ce type d'approche est écarté dans le reste du manuscrit au profit des deux autres.

Approche utilisée	séquence de la balle de tennis		séquence des deux voitures		séquence du chien et du ballon	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
FPFP <i>min</i> T-norm	0,6319	0,1374	0,5596	0,3531	0,3786	0,3639
FPFP moyenne arithmétique	0,6581	0,1659	0,4669	0,3949	0,3375	0,3053
FPFP <i>max</i> T-conorm	0,0468	0,1535	0,1715	0,2797	0,2811	0,2812

**Tableau 3.2** : Comparatif des différentes approches de fusion dans un filtre particulaire par taux de suivi moyen ( $\mu$ ) et l'écart-type du taux ( $\sigma$ ).

### Remarques et bilan des résultats

Nous avons conscience qu'il est difficile de généraliser ces résultats sans poursuivre les tests sur d'autres séquences. La réalisation d'un test au complet nécessite une phase d'acquisition, de pré-traitement (compression non destructrice), de paramétrisation des extracteurs-sources, et d'évaluation de la vérité terrain, qui au final rendent difficile l'exécution de tests à grande échelle. Toutefois, nous estimons que les séquences utilisées présentent des caractéristiques différentes et adaptées à l'évaluation des méthodes de fusion. Les résultats étant assez nets et argumentés, il nous paraît probant de conclure que l'utilisation des fonctions de croyance dans un filtre particulaire permet d'augmenter la robustesse d'un suivi d'objets dans le cas où les sources sont défaillantes. Quand les sources sont leurrées, la fusion bayésienne conserve l'avantage, même si la réponse apportée n'est pas entièrement satisfaisante. Des tests supplémentaires viendront confirmer ces affirmations dans les chapitres suivants.

La conclusion principale à retenir de ces premiers tests est la complémentarité des atouts du FPFC et du FPFB. Étant donné que la théorie des fonctions de croyance est le formalisme de modélisation de l'incertain le plus vaste, ce formalisme se prête mieux à l'étude des opérations de fusion et du sens qu'elles revêtent. Il existe en effet différentes façon de combiner des FDM, c'est-à-dire différentes règles de combinaison. Il convient d'étudier plus en détails les mécanismes des règles de combinaison afin de déterminer quelle règle peut s'apparenter dans son fonctionnement à celle de Bayes, puis de concevoir un processus pouvant allier cette règle à la règle conjonctive pour combiner leurs avantages.

## 3.6 Conclusion

Le présent chapitre s'est articulé de la manière suivante : présentation des filtres particuliers et de leurs mécanismes, présentation des approches existantes mêlant fusion et filtrage particulaire, présentation d'autres méthodes possibles, justification théorique de l'intérêt des fonctions de croyance par rapport aux autres méthodes, et enfin justification pratique de l'apport des fonctions de croyances en terme de suivi par filtre particulaire.

Nous nous sommes maintenant dotés d'un modèle et d'une procédure multi-source de suivi. Il nous paraît à présent naturel de nous pencher sur la meilleure façon de réaliser la

fusion à l'intérieur du cadre de travail que nous avons défini. Dans les fonctions de croyance, tout le processus de fusion se résume à la règle de combinaison choisie, d'où l'idée de nous pencher plus longuement sur ces règles dans le chapitre suivant. Les expériences ont montré que l'approche proposée comble un défaut de l'approche probabiliste classique mais qu'en revanche l'approche probabiliste fonctionne mieux dans le cas spécifique de plusieurs sources leurrées. Ces différences proviennent de propriétés disjointes permettant de répondre favorablement à un cas particulier. Nous nous intéresserons donc particulièrement aux propriétés des règles de combinaison afin de trouver un processus permettant de conjuguer les propriétés adéquates et donc de produire un suivi satisfaisant dans l'ensemble des exemples cités dans les expériences. A plusieurs reprises dans ce chapitre, l'importance du traitement sous-jacent de l'information relatif à la combinaison de sources, ainsi que la prise en compte du contexte ont déjà suscité des remarques sur leur importance dans la qualité finale des résultats. C'est notamment en analysant ces notions que nous pourrions proposer un mode de combinaison plus robuste. L'optimisation désirée du processus combinatoire sera justifiée directement par nos besoins en suivi d'objets.

# Chapitre 4

## Contribution à l'obtention d'une stratégie de fusion contextuel

### Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>121</b>
<b>4.2</b>	<b>Les principales règles de combinaisons</b>	<b>122</b>
4.2.1	Règles particulières	122
4.2.2	Familles de règles	126
4.2.3	Représentation matricielle	128
4.2.4	Remarques supplémentaires sur les règles	129
<b>4.3</b>	<b>Mécanismes et propriétés des règles de combinaisons</b>	<b>130</b>
4.3.1	Propriétés algébriques	130
4.3.2	Nature des combinaisons	132
4.3.3	Influence de l'élément d'évidence	133
4.3.4	Propriétés liées aux sous-ensembles	134
4.3.5	Autres propriétés	138
4.3.6	Comparatif des règles	139
<b>4.4</b>	<b>Stratégie de combinaison contextuel</b>	<b>139</b>
4.4.1	Combinaison de sources anormales	140
4.4.2	Utilisation du contexte dans une stratégie de combinaison de sources non fiables	145
<b>4.5</b>	<b>Expérimentations</b>	<b>149</b>
4.5.1	Cas de sources défaillantes	150
4.5.2	Cas de sources leurrées	151
4.5.3	Comparatif global des règles et de la stratégie de combinaison	156
<b>4.6</b>	<b>Conclusion</b>	<b>160</b>

---

### 4.1 Introduction

Comme évoqué dans le chapitre précédent, le formalisme des fonctions de croyance permet de mettre au point des outils de fusion appelés règles de combinaison. Bien que ces règles suivent des préceptes communs, elles peuvent être conçues de manières différentes pour répondre à tels ou tels enjeux de fusion. Selon ces enjeux, un certain nombre de propriétés sont fixées, et une règle possédant l'ensemble de ces propriétés est recherchée. Ce caractère *ad hoc* que peuvent revêtir les règles de combinaison est parfois critiqué par rapport aux probabilités où les calculs sont indépendants de ce genre de préoccupations. De notre point de vue, le fait de ne pas avoir à choisir une méthode de fusion est plus une façon de contourner le problème que d'y apporter

une solution.

L'étape de fusion d'informations dont nous avons besoin pour améliorer nos performances en suivi d'objets se concentre au niveau de la règle de combinaison. Nous proposons donc d'en étudier quelques unes des règles parmi les plus répandues en fonction des propriétés qu'elles peuvent remplir. Une fois ce comparatif fait, nous énoncerons nos besoins par rapport au suivi d'objets dans les séquences vidéos. Nous verrons que ces besoins sont liés aux difficultés évoquées en 2.7 et se traduisent d'un point de vue fusion par la présence de sources anormales à fusionner. Nous pourrions voir qu'aucune règle ne peut totalement satisfaire entièrement à nos besoins. Nous proposerons alors une solution alternative que nous appelons stratégie de combinaison contextuelle. Des résultats sont proposés en fin de chapitre afin de mettre en évidence l'intérêt de notre stratégie.

## 4.2 Les principales règles de combinaisons

Dubois et Prade [Dubois 86a] ont montré dans un article fondateur que pour combiner deux FDM, il est possible de proposer plusieurs règles mais que seul le produit des masses peut constituer l'élément de base pour transférer les croyances vers la FDM-résultat. En d'autres termes, en partant de la masse  $m_1(A)$  pour une première source et de la masse  $m_2(B)$  pour une seconde source, le transfert de masse se fait par l'affectation de  $m_1(A)m_2(B)$  à un élément  $C$  sur lequel est définie  $m_{1\odot 2}$ .

La valeur de  $m_{1\odot 2}(C)$  est alors égale à la somme des produits de masses qui lui sont affectés. Les règles de combinaison varient en fonction de la méthode permettant de sélectionner à partir d'un couple d'ensembles  $(A, B)$  quel sera l'ensemble  $C$  qui recevra le produit  $m_1(A)m_2(B)$ .

Nous présentons ici brièvement quelques unes de ces règles de combinaison sans trop donner de détails sur leurs différents usages dans un premier temps. Nous préférons ensuite présenter l'ensemble des propriétés que peuvent avoir les règles, et discuter de l'intérêt de ces propriétés. Au final, un comparatif des règles suivant les propriétés est présenté.

### 4.2.1 Règles particulières

Dans le reste du mémoire, la notation  $\odot_{xx}$  représente le symbole correspondant à la règle de combinaison  $xx$ . Nous notons de plus la FDM résultant de la combinaison suivant cette règle  $m_{xx}$ . Exceptionnellement, certains auteurs définissent un symbole spécifique à leur règle, nous reprenons alors ce symbole *symb* et la masse après fusion est notée  $m_{symb}$ .

Dans le chapitre précédent en 3.4.3, nous avons déjà présenté les règles de Dempster  $\oplus$  et conjonctive  $\odot$  dont les définitions sont très proches. Il existe en réalité une infinité de règles de combinaison possiblement utilisables dans le cadre des fonctions de croyance. Nous présentons ici quelque unes des plus citées d'entre elles.

La **règle disjonctive** vient en complément de la règle conjonctive ; elle est d'une certaine manière l'opération duale. La règle disjonctive, notée  $\odot$ , redistribue le crédit en provenance des deux FDM fusionnées en s'appuyant sur l'union des éléments focaux (au lieu de l'intersection

pour la règle conjonctive) :

$$m_{\odot}(A) = \sum_{B,C|B \cup C = A} m_1(B) m_2(C) \quad (4.1)$$

Cette règle fut introduite par Dubois et Prade [Dubois 86b], puis discutée par Smets dans le cadre du GBT [Smets 93]. La règle disjonctive interprète de manière différente la nature de l'information fournie par les sources (plus d'éléments sur cette question sont donnés en 4.3.2). De ce point de vue, les règles conjonctive et disjonctive sont deux visions extrêmes du problème, d'où l'idée de proposer des règles adoptant un comportement intermédiaire. Dubois et Prade [Dubois 92] ont donc introduit une autre règle, cette fois appelée **règle de Dubois et Prade**. Nous noterons cette règle  $\odot_{dp}$  qui s'écrit :

$$m_{dp}(A) = m_{\odot}(A) + \sum_{\substack{B \cap C = \emptyset \\ B \cup C = A}} m_1(B) m_2(C) \quad (4.2)$$

$$m_{dp}(\emptyset) = 0 \quad (4.3)$$

Cette règle est en quelque sorte un compromis entre la règle conjonctive et la règle disjonctive via un mécanisme de redistribution du conflit. Le mécanisme de redistribution du conflit part du constat que si deux sous-ensembles  $A$  et  $B$  sont disjoints, alors il faut appliquer le principe du minimum de spécificité, et donc transférer la croyance au plus petit ensemble contenant  $A$  et  $B$ , c'est-à-dire  $A \cup B$ .

Une des voies principale d'amélioration du processus de fusion est la gestion de la redistribution du conflit  $\kappa$ . Zadeh [Zadeh 86] souleva un cas paradoxal pour lequel la règle de Dempster produit un résultat contraire à la logique. Le contre-exemple de Zadeh se présente sous la manière suivante :

- le cadre de discernement se compose de trois éléments  $\Omega = \{\omega_1, \omega_2, \omega_3\}$ .
- deux FDM  $m_1$  et  $m_2$  sont fusionnées. Elles possèdent chacune deux éléments focaux dont un est commun aux deux  $\{\omega_2\}$  et dont les autres diffèrent.
- $m_1$  est telle que :  $m_1(\{\omega_1\}) = 0,99$ ,  $m_1(\{\omega_2\}) = 0,01$  et  $m_1(\{\omega_3\}) = 0$ .
- $m_2$  est telle que :  $m_2(\{\omega_1\}) = 0$ ,  $m_2(\{\omega_2\}) = 0,01$  et  $m_2(\{\omega_3\}) = 0,99$ .
- et par combinaison de Dempster :  $m_{\oplus}(\{\omega_2\}) = 1$ , certitude pour  $\{\omega_2\}$ .

Les deux FDM attribuent avant fusion une très faible croyance pour  $\{\omega_2\}$ , il est donc choquant d'obtenir la fonction totalement engagée pour  $\{\omega_2\}$  en sortie. Pour de nombreux auteurs, la clé du problème de combinaison est liée à la gestion du conflit  $m_{\odot}(\emptyset)$ . Yager [Yager 87a, Yager 87b] part de la règle de Dempster et propose de changer la répartition du conflit afin d'obtenir un résultat plus logique sur le contre-exemple de Zadeh :

$$m_y(A) = m_1 \odot m_2(A), \forall A \neq \emptyset, \neq \Omega, \quad (4.4)$$

$$m_y(\emptyset) = 0 \quad (4.5)$$

$$m_y(\Omega) = m_1 \odot m_2(\emptyset) + m_1 \odot m_2(\Omega) \quad (4.6)$$

Nous notons la **règle de Yager**  $\odot_y$ . Cette règle redistribue l'ensemble du conflit sur l'ignorance. Il n'est pas illogique en effet de considérer que le conflit est source d'ignorance. Avec cette règle, la réponse au contre-exemple de Zadeh est :  $m_y(\{\omega_2\}) = 0,0001$  et  $m_y(\Omega) = 0,9999$ , ce qui est

un résultat plus logique. La règle disjonctive produit le même résultat.

Notons que sur le contre-exemple de Zadeh, la règle conjonctive donne en sortie :  $m_{\odot}(\{\omega_2\}) = 0,0001$  et  $m_{\odot}(\emptyset) = 0,9999$ , ce qui reste une situation non-dommageable pour l'utilisateur puisque la masse attribuée à  $\{\omega_2\}$  reste faible. La situation décrite par Zadeh est typiquement une situation de fort conflit entre sources, le comportement de la règle conjonctive semble donc logique. Cet exemple illustre l'intérêt d'un monde ouvert, qui permet ainsi de détecter simplement des situations conflictuelles. La règle de Yager fait l'amalgame entre le conflit et l'ignorance, il y a donc une perte d'information par rapport à la règle conjonctive.

Toujours dans une optique de gestion du conflit, Daniel [Daniel 03] a présenté la **règle minC** (pour minimum de conflit). L'auteur propose que pour un produit de masses donné  $m_1(A)m_2(B)$ , résultant d'un conflit, c'est-à-dire  $A \cap B = \emptyset$ , cette masse ne soit réallouée qu'à des ensembles non disjoints de  $A \cup B$ . Deux schémas de redistribution sont proposés mais pas de formule générique pour les règles correspondantes. Le point de vue de Daniel s'oppose donc à celui de Dempster, qui estimait que de la masse conflictuelle n'appartenait pas à un ensemble plus qu'à un autre.

La redistribution du conflit a été également étudiée par Smarandache et Dezert [Smarandache 05b] dans le cadre de la théorie DSMT. Les auteurs ont proposé plusieurs règles déclinées sous l'appellation de règles à redistribution proportionnelle du conflit (Proportional Conflict Redistribution, PCR). La plus significative de ces règles a été adaptée à la théorie des fonctions de croyance par Florea *et al.* [Florea 06]. La **règle PCR** est notée  $\odot_{pcr}$ , et sa formule pour deux sources est la suivante :

$$m_{pcr}(A) = m_{\odot}(A) + \sum_{B \in 2^{\Omega} | A, B \cap A = \emptyset} \left[ \frac{m_1(A)^2 m_2(B)}{m_1(A) + m_2(B)} + \frac{m_2(A)^2 m_1(B)}{m_2(A) + m_1(B)} \right] \quad (4.7)$$

Les règles PCR ont été conçues spécialement pour redistribuer une part du conflit aux ensembles à l'origine de ce même conflit afin d'éviter que trop de conflit soit accumulé par  $\emptyset$ . Notons que cette règle est la seule à faire intervenir des non-linéarités.

La règle PCR a néanmoins été étendue par Martin et Osswald [Martin 07] qui lui ont adjoint un terme redistribuant une part du conflit à l'union des ensembles (à l'image de la règle de Dubois et Prade). Ce terme est pondéré par un coefficient d'affaiblissement<sup>1</sup>, la règle est donc dite à redistribution proportionnelle du conflit affaiblie (Discounted PCR, DPCR). La **règle DPCR** est notée  $\odot_{dpcr}$  et sa formule est la suivante :

$$m_{dpcr}(A) = m_{\odot}(A) + \sum_{B \cap A = \emptyset} \epsilon \left[ \frac{m_1(A)^2 m_2(B)}{m_1(A) + m_2(B)} + \frac{m_2(A)^2 m_1(B)}{m_2(A) + m_1(B)} \right] + \sum_{B \cup C = A, B \cap C = \emptyset} (1 - \epsilon) m_1(B) m_2(C) \quad (4.8)$$

avec  $\epsilon$  le coefficient d'affaiblissement relatif à la règle DPCR.

<sup>1</sup>Cet affaiblissement ne correspond pas à la définition formelle fournie en 3.4.6. La masse est réaffectée à l'union des ensembles au lieu de  $\Omega$ , mais le principe de "désengagement" de la FDM est similaire.

Zhang [Zhang 94] est parti sur une voie différente et propose de pondérer les produits de masse à l'aide des cardinaux des ensembles entrant en jeu. Pour le reste, la définition de cette règle est identique à celle de Dempster, nous la notons donc  $\odot_z$  :

$$\forall A \neq \emptyset, m_z(A) = \frac{1}{Z} \sum_{B, C / B \cap C = A} \frac{|A|}{|B||C|} m_1(B) m_2(C) \quad (4.9)$$

$$m_z(\emptyset) = 0 \quad (4.10)$$

avec  $Z$  un facteur de normalisation. Cette règle permet de prendre en compte à quel point deux ensembles s'intersectent avant de leur affecter le produit de leur croyance. La pondération proposée est donc une mesure d'intersection entre deux ensembles et, à ce titre, une autre mesure sémantiquement équivalente pourrait lui être substituée. De même, des mesures définissant un degré d'inclusion ou d'union existent et pourraient être insérées dans une règle. En revanche, ce genre de règle impose une renormalisation afin de respecter l'équation 3.51 définissant les FDM. Elles ne répondent pas non plus spécifiquement au contre-exemple de Zadeh.

Murphy [Murphy 00] a introduit la règle dite de la moyenne statistique. Dans ce mémoire, nous la référençons sous l'appellation de **règle de Murphy**, et elle est notée  $\odot_m$ . Elle se définit par :

$$m_m(A) = \frac{1}{2} (m_1(A) + m_2(A)) \quad (4.11)$$

Le principe de moyennage peut être généralisé à l'aide de pondérations différentes de  $\frac{1}{2}$ . Cette règle souffre d'un défaut majeur au sens où elle n'est pas clairement justifiée d'un point de vue du traitement de l'information. Il n'y a pas plus de raisons de moyennner des probabilités que des fonctions de croyance.

La moyenne intervient de nouveau dans Josang *et al.* [Josang 03] mais pour définir une répartition du conflit. La règle est appelée opérateur pondéré par la moyenne (Weighted Average Operator, WAO) et nous la notons  $\odot_{wao}$ . La **règle WAO** se définit par :

$$m_{wao}(A) = m_{\odot}(A) + m_{\odot}(\emptyset) [m_1(A) + m_2(A)] \text{ pour } A \neq \emptyset \quad (4.12)$$

$$m_{wao}(\emptyset) = m_{\odot}(\emptyset) [m_1(\emptyset) + m_2(\emptyset)] \quad (4.13)$$

Le cœur de la règle reste une combinaison conjonctive.

Delmotte *et al.* [Delmotte 95] ont étudié un aspect différent en vue d'élaborer une règle plus efficace : la fiabilité des sources. La fiabilité de chaque source est évaluée par un coefficient  $R_i$ , avec  $i$  l'indice relatif à la source  $S_i$ . La première règle proposée par les auteurs repose sur un principe de moyennage, nous l'appellerons donc **règle de la moyenne de Delmotte**. Nous la notons  $\odot_{del-moy}$  et sa formule est la suivante :

$$m_{del-moy}(A) = R_1 m_1(A) + R_2 m_2(A) \quad (4.14)$$

Une deuxième règle prend en compte ces mêmes coefficients de fiabilité en les répartissant sur un premier terme lié à la règle conjonctive et un second terme lié à la règle disjonctive. C'est un cas particulier de la famille des règles mixtes, qui seront présentées dans la sous-section suivante.

Nous appelons donc cette règle, la **règle mixte de Delmotte**. Nous la notons  $\odot_{del-mix}$  et sa formule est la suivante :

$$m_{del-mix}(A) = \Omega^0(A) \text{ si } \forall i, R_i = 0 \quad (4.15)$$

$$m_{del}(\emptyset) = 0 \quad (4.16)$$

$$m_{del}(A) = \frac{1}{Z} [And(R_1..R_M) m_{\odot}(A) + Xor(R_1..R_M) m_{\odot}(A)] \text{ si } A \neq \emptyset \quad (4.17)$$

$$And(R_1, \dots, R_M) = \prod_{i=1}^M R_i \quad (4.18)$$

$$Xor(R_1, \dots, R_M) = \left[ 1 - \prod_{i=1}^M (1 - R_i) \right] \left( 1 - \prod_{i=1}^M R_i \right) \quad (4.19)$$

$$Z = Xor(R_1, \dots, R_M) + \kappa And(R_1, \dots, R_M) \quad (4.20)$$

La formule est présentée pour  $M$  sources afin de mieux voir la répartition des coefficients de fiabilité. Les règles conjonctive et disjonctive intervenant dans cette formule, cette règle joue donc aussi sur l'aspect de la nature de la combinaison.

#### 4.2.2 Familles de règles

Certains auteurs ont proposé des formalismes unificateur des règles de combinaison sous la forme de familles de règles. Lefevre *et al.* [Lefevre 02] par exemple présentent une famille de règles fondées sur la règle de Dempster et qui englobe toutes les redistributions possibles du conflit sur les éléments de  $2^\Omega$ . Les termes intervenant dans la redistribution sont appelés des poids dont la somme vaut  $m(\emptyset)$ . Nous notons  $\beta(A)$  le poids de l'ensemble  $A$  et  $\odot_l$  une **règle de Lefevre** :

$$m_l(A) = m_{1\odot 2}(A) + \beta(A) m_{1\odot 2}(\emptyset) \text{ pour } A \neq \emptyset \quad (4.21)$$

$$m_l(\emptyset) = \beta(\emptyset) m_{1\odot 2}(\emptyset) \quad (4.22)$$

$$\text{et } \sum_{A \subseteq \Omega} \beta(A) = 1 \quad (4.23)$$

La meilleure redistribution du conflit est apprise par minimisation d'un critère d'erreur. Cette formulation du problème regroupe la règle de Dempster, de Yager et de Dubois & Prade dans un même ensemble.

Une idée similaire avait été développée auparavant par Inagaki [Inagaki 91] mais un seul paramètre  $k$  code la répartition du conflit :

$$m_i(A) = (1 + km_{1\odot 2}(\emptyset)) m_{1\odot 2}(A) \text{ pour } \forall A \in 2^\Omega / \{\emptyset, \Omega\} \quad (4.24)$$

$$m_i(\Omega) = (1 + km_{1\odot 2}(\emptyset)) m_{1\odot 2}(\Omega) + (1 + km_{1\odot 2}(\emptyset) - k) m_{1\odot 2}(\emptyset) \quad (4.25)$$

$$m_i(\emptyset) = 0 \quad (4.26)$$

avec  $\odot_i$  une **règle d'Inagaki** de paramètre  $k$ . Cette méthode, appelée règle de combinaison unifiée d'Inagaki est donc moins globale que la famille de règles décrite par Lefevre.

Récemment, Denœux [Denœux 08] a introduit une famille de règles dont nous ne présentons que les deux plus significatives, appelées respectivement règle conjonctive prudente et règle

disjonctive hardie. La **règle conjonctive prudente**, notée  $\odot$ , se définit à partir de la fonction de poids conjonctifs  $w(\cdot)$ .

$$m_{\odot}(A) = \odot_{AC\Omega} A^{w_1(A) \wedge w_2(A)} \quad (4.27)$$

où  $\wedge$  est l'opérateur de minimum. Elle est dite conjonctive, puisque comme le montre la formule 4.27, elle résulte de combinaisons conjonctives. De manière duale, la **règle disjonctive hardie**, notée  $\oslash$  se définit à partir de la fonction de poids disjonctifs  $v(\cdot)$  :

$$m_{\oslash}(A) = \oslash_{AC\Omega} A^{v_1(A) \vee v_2(A)} \quad (4.28)$$

Là encore, le terme disjonctif se justifie par l'équivalence à un ensemble de combinaisons disjonctives. Le principal intérêt de ces règles est leur capacité à pouvoir fusionner des sources non-distinctes. A cause de la définition des fonctions de poids conjonctifs et disjonctifs, ces règles ne sont pas applicables sur n'importe quelles FDM. La règle conjonctive prudente s'utilise sur des FDM non-dogmatiques et la règle disjonctive hardie sur des FDM non-normées.

Par ailleurs, Smets [Smets 97] a introduit une famille de règles généralisant les règles conjonctives et disjonctives. Un paramètre  $\alpha \in [0, 1]$  permet de définir des opérateurs avec un degré de conjonction ou de disjonction modulable, d'où le nom de **règles  $\alpha$ -jonctives** donné à cette famille. Une règle appartenant à cette famille mérite en particulier d'être citée pour sa singularité : la **règle disjonctive exclusive**. Nous la notons  $\underline{\oslash}$  et sa définition est la suivante :

$$m_{\underline{\oslash}} = \sum_{(B \cap \bar{C}) \cup (\bar{B} \cap C) = A} m[S_1](B) m[S_2](C) \quad (4.29)$$

Nous verrons que cette règle possède un intérêt surtout théorique car elle correspond à une situation de fusion particulière (cf. 4.3.2).

Plus récemment, Florea *et al.* [Florea 06] ont introduit la famille des **règles adaptatives**. Cette famille prône également une solution intermédiaire à la conjonction ou à la disjonction. Nous la notons  $\odot_f$ . Elle se définit de la manière suivante :

$$m_f(A) = \gamma_1(\kappa) m_{\odot}(A) + \gamma_2(\kappa) m_{\oslash}(A) \text{ if } A \neq \emptyset \quad (4.30)$$

$$m_f(\emptyset) = 0 \quad (4.31)$$

avec  $\gamma_1(\cdot)$  et  $\gamma_2(\cdot)$  deux fonctions du conflit  $\kappa$ . Les auteurs proposent une définition dite symétrique de ces fonctions :

$$\gamma_1(\kappa) = \frac{\kappa}{1 - \kappa + \kappa^2} \quad (4.32)$$

$$\gamma_2(\kappa) = \frac{1 - \kappa}{1 - \kappa + \kappa^2} \quad (4.33)$$

Par la suite, nous désignerons par **règle adaptative de Florea**, la règle de Florea dotée des fonctions définies ci-dessus. La famille de Florea a été étendue par Martin *et al.* [Martin 07]. Cette extension est appelée famille des **règles mixtes**, que nous notons  $\odot_{mix}$ , et qui se définissent de la manière suivante :

$$m_{mix}(A) = \sum_{B \cup C = A} \delta_1(B, C) m_1(B) m_2(C) + \sum_{B \cap C = A} \delta_2(B, C) m_1(B) m_2(C) \quad (4.34)$$

avec  $\delta_1(.,.)$  et  $\delta_2(.,.)$  deux fonctions associant un couple de sous-ensembles à un coefficient réel. Si  $\delta_1 = \beta_1$  et  $\delta_2 = \beta_2$ , alors la règle mixte de Martin revient à la règle adaptative. Les auteurs préconisent néanmoins l'usage des fonctions suivantes :

$$\delta_1(B, C) = 1 - \frac{\#(B \cap C)}{\#(B \cup C)} \quad (4.35)$$

$$\delta_2(B, C) = 1 - \delta_1(B, C) \quad (4.36)$$

Par la suite, nous désignerons par **règle mixte de Martin**, la règle mixte dotée des fonctions définies ci-dessus. Comme ces fonctions préconisées prennent en compte le cardinal des sous-ensembles, la règle mixte de Martin prend en compte l'indétermination des sources de manière similaire à la règle de Zhang.

Martin et Osswald ont de plus combiné la règle DPCR et la règle mixte dans la règle logiquement dite à redistribution proportionnelle du conflit affaiblie mixte (Mixed DPCR, MDPCR). La combinaison des deux règles s'effectue de manière simple : les masses non-conflictuelles sont gérées par la règle mixte, tandis que les masses conflictuelles sont gérées par la règle DPCR. La **règle MDPCR** est notée  $\odot_{mdpcr}$ , et sa formule donne donc :

$$m_{mdpcr}(A) = \sum_{\substack{B \cup C = A \\ B \cap C \neq \emptyset}} \delta_1(B, C) m_1(B) m_2(C) \quad (4.37)$$

$$+ \sum_{\substack{B \cap C = A \\ B \cap C \neq \emptyset}} \delta_2(B, C) m_1(B) m_2(C) \quad (4.38)$$

$$+ \sum_{B \cap A = \emptyset} \epsilon \left[ \frac{m_1(A)^2 m_2(B)}{m_1(A) + m_2(B)} + \frac{m_2(A)^2 m_1(B)}{m_2(A) + m_1(B)} \right] \quad (4.39)$$

$$+ \sum_{\substack{B \cup C = A \\ B \cap C = \emptyset}} (1 - \epsilon) m_1(B) m_2(C) \quad (4.40)$$

### 4.2.3 Représentation matricielle

Certaines règles de combinaison peuvent être représentées à l'aide de matrices [Smets 00]. Il convient alors de représenter une FDM  $m_i(.)$  sous forme vectorielle  $\vec{m}_i$  :

$$\vec{m}_i = \begin{pmatrix} m_i(\emptyset) \\ \vdots \\ m_i(A) \\ \vdots \\ m_i(\Omega) \end{pmatrix} \quad (4.41)$$

Cette représentation suppose un ordonnancement des sous-ensembles de  $\Omega$ . Par convention, il est commode de choisir cet ordonnancement par analogie avec le codage binaire. Voici une illustration

pour un cadre de discernement à trois éléments  $\Omega = \{a, b, c\}$  :

$$\begin{array}{ccc}
 a, b, c & \text{Ensemble} & \text{vecteur } \vec{m} \\
 0, 0, 0 & \emptyset & \left( \begin{array}{c} m_i(\emptyset) \\ m_i(\{c\}) \\ m_i(\{b\}) \\ m_i(\{b, c\}) \\ m_i(\{a\}) \\ m_i(\{a, c\}) \\ m_i(\{a, b\}) \\ m_i(\{\Omega\}) \end{array} \right) \\
 0, 0, 1 & \{c\} & \\
 0, 1, 0 & \{b\} & \\
 0, 1, 1 & \{b, c\} & \longrightarrow \\
 1, 0, 0 & \{a\} & \\
 1, 0, 1 & \{a, c\} & \\
 1, 1, 0 & \{a, b\} & \\
 1, 1, 1 & \{\Omega\} & 
 \end{array} \quad (4.42)$$

Une règle de combinaison  $\odot$  se réécrit alors :

$$\vec{m}_{1\odot 2} = M_{\odot}^{m_1} \vec{m}_2 \quad (4.43)$$

$M_{\odot}^{m_1}$  est une matrice de taille  $|2^{\Omega}| \times |2^{\Omega}|$  et  $\vec{m}_{1\odot 2}$  le vecteur masse de croyance issu de la fusion. La valeur de  $M_{\odot}^{m_1}(A, B)$  représente la fraction de la masse  $m_2(B)$  qui sera attribuée à  $A$  par le processus de combinaison avec  $m_1$ . Il est important de noter que  $M_{\odot}^{m_1}$  caractérise la règle  $\odot$  mais dépend d'une FDM donnée  $m_1$ . Si la règle  $\odot$  est représentable matriciellement, alors  $\forall m \in \mathfrak{B}^{\Omega}$  il existe  $M_{\odot}^m$  représentant  $m \odot \dots$

La règle conjonctive accepte la représentation matricielle. L'exemple suivant montre cette règle pour une FDM  $m_1$  quelconque et pour un cadre de discernement simple  $\Omega = \{a, b\}$  :

$$M_{\odot}^{m_1} = \begin{array}{c} \emptyset \\ \{a\} \\ \{b\} \\ \Omega \end{array} \left( \begin{array}{cccc} 1 & m_1(\emptyset) + m_1(\{b\}) & m_1(\emptyset) + m_1(\{a\}) & m_1(\emptyset) \\ 0 & m_1(\{a\}) + m_1(\Omega) & 0 & m_1(\{a\}) \\ 0 & 0 & m_1(\{b\}) + m_1(\Omega) & m_1(\{b\}) \\ 0 & 0 & 0 & m_1(\Omega) \end{array} \right) \quad (4.44)$$

Le processus de combinaison peut être itéré de sorte que :  $\vec{m}_1 \odot \vec{m}_2 \odot \dots \odot \vec{m}_n = M_{\odot}^{m_1} M_{\odot}^{m_2} \dots M_{\odot}^{m_{n-1}} \vec{m}_n$ . La représentation matricielle aide à mieux visualiser certaines propriétés des règles de combinaison pour des cadres de discernement à faibles cardinaux. Les résultats observables sur des cadres de petite taille sont généralisables à des dimensions plus élevées. La représentation matricielle permet entre autres de voir quels sous-ensembles sont prioritairement bénéficiaires ou déficitaires vis-à-vis de la fusion.

#### 4.2.4 Remarques supplémentaires sur les règles

Certains commentaires de cette partie sont tirés de Smarandache [Smarandache 04a], Sentz et Ferson [Sentz 02] et Smets [Smets 06]. Nous recommandons la lecture de ces articles pour plus de détails sur les différentes règles citées. Ces articles contiennent également d'autres définitions de règles ou des références vers des articles présentant d'autres règles. La littérature regorge en effet de règles de combinaison. Les règles qui ont été écartées de ce mémoire sont jugées désuètes, d'usage anecdotique, s'éloignant du formalisme des fonctions de croyance ou encore inapplicables par rapport au cadre de discernement défini.

Par ailleurs, quelques auteurs intègrent au processus de combinaison une étape de conditionnement ou d'affaiblissement. S'il est vrai que de telles étapes ont une influence sur la qualité des résultats, nous ne considérons pas qu'elles puissent appartenir au processus de combinaison,

car que les sources soient affaiblies ou conditionnées, les choix fondamentaux qui contraignent le transfert des croyances restent les mêmes. De plus, le conditionnement est étroitement lié à la règle conjonctive, et l'affaiblissement à la règle disjonctive.

En outre, les formules qui sont données dans cette partie sont présentées dans le cas où seulement deux sources sont à fusionner mais la majorité des règles peuvent être généralisées à  $M > 2$  sources. La discussion sur l'associativité des règles (*cf.* 4.3.1) vient en complément de cette remarque.

Enfin, il est à noter que certains auteurs considèrent que seule la règle de Dempster est valable [Haenni 02] ou que le contre-exemple de Zadeh n'en est pas un [Haenni 05]. Quoi qu'il en soit, il n'y a pas de règle qui fasse l'unanimité dans la communauté. La profusion des règles est interprétée par les détracteurs des fonctions de croyance comme une preuve de la faiblesse de l'axiomatique des fonctions de croyance. De notre point de vue, chaque règle se justifie par les propriétés qu'elle remplit et par le sens donné à ces propriétés. La section suivante propose une étude de ces propriétés. Nous excluons de la suite des travaux les familles de règles pour lesquelles il est impossible de statuer dans le cas général sur leurs propriétés.

## 4.3 Mécanismes et propriétés des règles de combinaisons

### 4.3.1 Propriétés algébriques

Nous appelons propriétés algébriques, les propriétés des règles de combinaison s'apparentant aux propriétés des lois internes d'ensembles algébriques. Pour plus de clarté, les propriétés seront souvent énoncées pour le cas de deux ou trois sources à fusionner.

#### Commutativité

Une règle de combinaison  $\odot$  est commutative si et seulement si,  $\forall m_1$  et  $m_2 \in \mathfrak{B}^\Omega$  :

$$m_1 \odot m_2 = m_2 \odot m_1 \quad (4.45)$$

La commutativité peut s'interpréter comme le fait que la règle traite de façon égale deux opérands. Si une règle n'est pas commutative, cela implique qu'un des opérands est favorisé lors du transfert de masse par rapport à l'autre. Ce genre de comportement peut trouver une justification dans des contextes précis, mais conséquemment une telle règle perd énormément en polyvalence. De plus, un contexte est par définition amené à évoluer au cours des processus, il est donc dangereux de constamment opérer des combinaisons déséquilibrées. La commutativité apparaît donc comme une propriété indispensable pour une règle de combinaison, et toutes les règles évoquées dans ce mémoire sont commutatives.

#### Associativité

Une règle de combinaison  $\odot$  est associative si et seulement si,  $\forall m_1, m_2$  et  $m_3 \in \mathfrak{B}^\Omega$  :

$$(m_1 \odot m_2) \odot m_3 = m_1 \odot (m_2 \odot m_3) \quad (4.46)$$

L'associativité couplée à la commutativité peut s'interpréter comme le fait que l'ordre dans lequel les sources sont agrégées n'a pas d'importance, et que toutes les sources n'ont pas besoin d'être

agrégées en même temps. La première partie de cette interprétation est la plus importante d'un point de vue du traitement de l'information. Si toutes les sources sont disponibles en même temps, la règle doit pouvoir s'abstraire de l'ordre dans lequel sont réalisées les fusions en agrégeant les sources de manière globale :

$$\odot_{i=1}^M m_i = \Gamma(m_1, \dots, m_M) \quad (4.47)$$

où  $\Gamma(\cdot)$  associe  $M$  FDM à une seule FDM. Si quelle que soit la permutation  $perm(\cdot)$  sur l'indice de la source  $\Gamma(m_1, \dots, m_M) = \Gamma(m_{perm(1)}, \dots, m_{perm(M)})$ , on parle alors d'opérateur  $M$ -aire. De même que la commutativité, cette propriété est indispensable au bon fonctionnement d'une règle sauf contexte très particulier. Toutes les règles évoquées dans ce mémoire sont  $M$ -aires. En revanche, toutes ne sont pas associatives comme la règle de Yager, de Zhang ou de Dubois & Prade. En conséquence, si l'information en provenance d'une source parvient en retard, il faut soit attendre toutes les sources, soit recalculer en entier la combinaison via l'équation 4.47.

Notons que si dans l'absolu, une règle  $M$ -aire permet de fusionner équitablement plusieurs sources, il est beaucoup plus coûteux en temps de calcul d'utiliser l'équation de 4.47 plutôt que de faire des combinaisons successives entre deux termes comme le permet l'associativité. C'est pourquoi Haenni [Haenni 02] estime que toute règle devrait être associative. Yager [Yager 87b] n'est pas de cet avis, il souligne que ce qui importe réellement est de pouvoir par un moyen quelconque prendre en compte l'arrivée d'une nouvelle source quand la combinaison a déjà été entamée. S'il existe un tel moyen, Yager parle de règle **quasi-associative**. Pour illustrer son idée, Yager prend l'exemple de la moyenne arithmétique, qui peut être mise à jour quand une nouvelle donnée arrive sans reprendre tout le calcul, bien que l'opération de moyenne ne soit pas associative. Les règles de Yager et Dubois & Prade sont quasi-associatives.

Si les informations en provenance des sources arrivent à des instants différents et que les dates d'arrivée importent (données périssables), il peut être intéressant de moduler la combinaison selon ces dates. En rapport avec ce thème, Smets parle d'exigence markovienne dans Smets [Smets 06]. En général, le dernier élément d'évidence arrivé est jugé plus fiable que les autres. Cela dit, il peut être plus simple de séparer le problème en deux avec d'abord une phase d'atténuation des sources selon leur date d'arrivée et ensuite une phase de combinaison classique. Nous ne donnons pas plus de détails sur cette notion, car dans notre cadre applicatif, toutes les sources sont synchrones.

## Idempotence

Une règle de combinaison  $\odot$  est idempotente si et seulement si,  $\forall m \in \mathfrak{B}^\Omega$  :

$$m \odot m = m \quad (4.48)$$

A notre connaissance, seule la famille de règles introduite par Denœux et les règles de type moyenne possèdent la propriété d'idempotence. Son interprétation peut être vue sous deux angles différents. Si deux sources donnent exactement le même avis sur le problème posé, il semble logique de valider cet avis en sortie de la fusion, d'où l'intérêt de l'idempotence. En revanche, deux avis identiques peuvent également tendre à renforcer le crédit attribué aux éléments focaux de la FDM fusionnée, ce qui implique que l'ignorance  $m(\Omega)$  doive baisser suite à la fusion. En conclusion, selon l'évolution souhaitée des croyances, l'idempotence peut être intéressante mais n'est pas indispensable de façon générale.

### Existence d'un élément neutre ou absorbant

Une FDM  $m_e$  est un élément neutre de la règle de combinaison  $\odot$  si et seulement si  $\forall m \in \mathfrak{B}^\Omega$  :

$$m_e \odot m = m \quad (4.49)$$

Une FDM  $m_a$  est un élément absorbant de la règle de combinaison  $\odot$  si et seulement si  $\forall m \in \mathfrak{B}^\Omega$  :

$$m_a \odot m = m_a \quad (4.50)$$

De tels éléments sont par exemple :

- $\Omega^0(.)$  est neutre pour  $\oplus$
- $\Omega^0(.)$  est absorbant pour  $\oplus$
- $\emptyset^0(.)$  est neutre pour  $\ominus$
- $\emptyset^0(.)$  est absorbant pour  $\ominus$

Sur l'intérêt que telle ou telle FDM soit absorbante ou neutre, plusieurs points de vue sont adoptables. Par exemple si une des sources à fusionner est  $\Omega^0(.)$  et que la décision doit être prise avec la plus grande prudence, alors la FDM fusionnée doit être également  $\Omega^0(.)$ .  $\Omega^0(.)$  serait donc un élément absorbant. D'un autre côté,  $\Omega^0(.)$  peut être considérée comme vide d'information et écartable du processus de fusion, donc quelle que soit l'autre terme à fusionner, il doit être égal à la FDM fusionnée.  $\Omega^0(.)$  serait donc cette fois élément neutre.

Comme pour l'idempotence l'existence de tels éléments est potentiellement intéressante mais pas indispensable. Notons que l'unicité de ces éléments n'est absolument pas garantie.

#### 4.3.2 Nature des combinaisons

Soient deux éléments d'évidence  $Ev_1$  et  $Ev_2$  induits par deux sources, et représentés respectivement par deux FDM  $m[S_1]$  et  $m[S_2]$ . Soit  $A$  un élément focal de  $m[S_1]$  et  $B$  un élément focal de  $m[S_2]$ . Afin de ne négliger aucune hypothèse ayant reçu le moindre crédit et de prendre le moins de risque possible, la masse de croyance relative au couple  $(A, B)$  doit être transférée à  $A \cup B$ . Cette combinaison de FDM ne rejette aucune hypothèse, et ne génère pas de conflit. C'est la vue disjonctive d'une combinaison où il est supposé qu'au moins l'une des deux sources  $S_1$  ou  $S_2$  dit vrai mais qu'il est impossible de savoir laquelle.

A l'inverse, si les avis fournis par les deux sources sont considérés tous deux exacts, selon un certain degré de croyance, alors c'est que la "vérité" se trouve dans l'intersection  $A \cap B$ . C'est alors à cet ensemble que la croyance doit être transférée. Cette combinaison est sélective car elle élimine des hypothèses et peut générer du conflit. C'est la vue conjonctive d'une combinaison, où il est supposé que les deux sources disent vrai.

Les avantages de ces deux modes de pensées diffèrent et il est difficile de dire dans l'absolu quel est le meilleur. Le mode disjonctif est plus prudent que le mode conjonctif, mais le mode conjonctif amène plus rapidement à une solution au risque de tomber sur la fonction de conflit total. Comme le précise Smets dans [Smets 97] une FDM résultant d'une combinaison de nature conjonctive sera plus engagée que les FDM dont elle est issue. La notion d'engagement rejoint celle de sélection de la solution. Des éléments supplémentaires de discussion sur l'intérêt du conjonctif par rapport au disjonctif sont disponibles dans Dubois et Prade [Dubois 86b] et

dans Smets [Smets 97, Smets 06].

Il existe des règles qui possèdent à la fois une part de conjonction et de disjonction. Ce sont par exemple la famille des règles  $\alpha$ -jonctives et particulièrement la règle exclusive disjonctive. Cette règle est sensée répondre à la situation selon laquelle une seule des deux sources dit vrai mais il est impossible de savoir laquelle. Il est assez difficile d'imaginer un usage concret de cette règle, puisqu'il faut au moins savoir qu'une règle dit vrai et l'autre non. Si cette information est disponible, elle est très généralement nominative pour les sources et nul besoin d'utiliser la règle, il suffit de censurer la source incorrecte.

La règle de Dubois & Prade est à base conjonctive de par le premier terme de la formule 4.2 mais possède également une part de disjonction de par le second terme de la même formule. La table 4.2 présentée à la fin de cette section résume entre autres la nature sous-jacente des combinaisons des règles que nous étudions.

Notons par ailleurs que la règle disjonctive appliquée à des FDM bayésiennes revient quasiment à utiliser la fusion bayésienne sur des distributions de probabilités. La fusion bayésienne est donc de nature disjonctive, c'est pour cela qu'elle fonctionne dans des cas où le FPFC (de nature conjonctive) ne fonctionne pas et inversement. D'après les expériences présentées au chapitre 3, il serait intéressant d'avoir un opérateur à la fois disjonctif et conjonctif.

### 4.3.3 Influence de l'élément d'évidence

De même qu'il existe plusieurs natures de combinaison, il existe plusieurs natures d'éléments d'évidence. Cette nature peut avoir un impact significatif sur le résultat d'une fusion. La nature d'un élément d'évidence est conditionnée principalement par deux aspects :

- conjonctif ou disjonctif : selon la nature du problème traité, les éléments d'évidence peuvent également être conjonctif ou disjonctif, mais la définition de ces termes pour un élément d'évidence est singulièrement différente de celle pour la nature d'une combinaison. Dans le cas disjonctif, un seul élément d'un ensemble  $A$  peut être la solution au problème. Dans le cas conjonctif, tous les éléments d'un ensemble  $B$  sont des réponses exactes au problème.

Ce sont notamment Dubois et Prade [Dubois 86b] qui ont discuté de la nature d'un élément d'évidence. Les auteurs donnent l'exemple suivant : l'événement "John est grand" signifie que la valeur réelle de la taille de "John" est contenue dans un intervalle symbolisé par l'ensemble "Grand", mais seulement une de ces valeurs est possible, car "John" ne peut avoir deux tailles différentes. Il s'agit dans ce cas d'un élément d'évidence disjonctif.

Par ailleurs, la phrase "John a vécu à Paris de 1980 à 1984" signifie qu'il est vrai de dire que "John" était effectivement à Paris pour chacun des éléments suivants : 1980, 1981, 1982, 1983, 1984. Il s'agit dans ce cas d'un élément d'évidence conjonctif.

Il est assez rare que des auteurs s'intéressent à des éléments d'évidence conjonctifs. D'après le principe d'exclusion mutuelle des hypothèses d'un cadre de discernement, les éléments d'évidence disjonctifs semblent impératifs. D'après la façon dont nous avons

construit notre cadre de discernement (cf. 3.5.1), les éléments d'évidence à agréger sont tous disjonctifs.

- distinct ou non : il s'agit de savoir si les éléments d'évidence sont distincts entre eux. La définition de la distinction a été donnée en 3.4.5. Pour une règle donnée, la propriété souhaitable est de pouvoir fusionner des sources qu'elles soient distinctes ou non. En théorie, seule la famille de règles introduite par Denœux possède une telle propriété, mais en pratique, les autres règles sont appliquées même entre sources non-distinctes. Pour deux sources non-distinctes à fusionner, Smets recommande de construire les FDM jointes sur l'espace produit  $\Omega \times \Omega$  afin de séparer chaque hypothèse en couples dont chaque composante correspond à une source. Ce procédé alourdit néanmoins le processus de fusion.

#### 4.3.4 Propriétés liées aux sous-ensembles

Dans cette sous-section, nous introduisons les propriétés qui y sont présentées afin de traduire le fait que certains sous-ensembles de  $\Omega$  revêtent un rôle particulier au sein des combinaisons de croyances. Nous partons du constat que les sous-ensembles de  $\Omega$  ne sont pas traités de manière équitable par les règles de combinaison et que ce traitement caractérise les règles. Il est clair qu'une règle qui traiterait totalement équitablement les sous-ensembles ne serait pas forcément plus pertinente qu'une autre car le propre d'une combinaison est aussi de faire transiter les masses de croyance selon un schéma favorable à une prise de décision.

Néanmoins, le comportement de certaines règles vis-à-vis de certains sous-ensembles tend à biaiser les combinaisons, ce sont donc des limites que nous cherchons à tracer plutôt que des contraintes rigides. Nous introduirons notamment à cet effet la notion de règle impartiale. L'idée est que pour une telle règle, toute masse attribuée à un ensemble quelconque puisse augmenter ou diminuer après combinaison.

Des définitions viennent en préambule à ces propriétés proposées afin de faciliter leur écriture.

**Définition :**  $A \subseteq \Omega$  est un sous-ensemble absorbant pour la règle  $\odot$ , si et seulement si  $\forall m$  et  $m' \in \mathfrak{B}^\Omega$ ,  $m \odot m'(A) \geq m(A)$  et  $m' \odot m(A) \geq m(A)$ .

En d'autres termes, si une règle possède un sous-ensemble absorbant, la croyance associée à ce dernier ne peut qu'augmenter. Par exemple,  $\emptyset$  est absorbant pour la règle conjonctive, et  $\Omega$  est absorbant pour la règle disjonctive. Il existe clairement une corrélation entre sous-ensemble absorbant et élément absorbant. Nous estimons qu'il est plus important de déterminer les sous-ensembles absorbants d'une règle, que ses éléments absorbants. Par exemple,  $\emptyset^0(\cdot)$  est un élément absorbant pour la règle conjonctive, mais la notion d'absorption des croyances ne se résume pas à cette seule FDM mais à toutes celles pour lesquelles  $\emptyset$  est un élément focal.

**Définition :**  $A \subseteq \Omega$  est un sous-ensemble de rejet<sup>2</sup> pour la règle  $\odot$ , si et seulement si  $\forall m$  et  $m' \in \mathfrak{B}^\Omega$ ,  $m \odot m'(A) \leq m(A)$  et  $m' \odot m(A) \leq m(A)$ .

<sup>2</sup>La dénomination "rejet" ne fait pas référence au sens attribué à ce terme par la communauté de reconnaissance des formes. Ce nom indique que le sous-ensemble rejette les produits de masses de croyances qui pourraient lui être affectés lors d'une combinaison de masse selon une règle donnée.

En d'autres termes, si une règle possède un sous-ensemble de rejet, la croyance associée à ce dernier ne peut que diminuer. Par exemple,  $\emptyset$  est un sous-ensemble de rejet pour la règle disjonctive, et  $\Omega$  est sous-ensemble de rejet pour la règle conjonctive. La même remarque que pour les sous-ensembles absorbants est valable pour les sous-ensembles de rejet en comparaison aux éléments neutres d'une règle.

Nous donnons logiquement alors la définition suivante de l'**impartialité** :

**Définition :** *Une règle de combinaison est impartiale si et seulement si elle ne possède aucun sous-ensemble de rejet ou absorbant.*

En utilisant une règle impartiale, les croyances peuvent évoluer de manière plus libre. Malheureusement, très peu de règles de combinaison sont impartiales. Les règles conjonctive et disjonctive, comme nous l'avons vu sont clairement partiales. Les règles conjonctive prudente et disjonctive hardie étant respectivement des conjonctions et disjonctions multiples ne sont pas non plus impartiales. *Stricto sensu*, toutes les règles prévues pour un monde fermé sont partiales, car elles maintiennent toutes artificiellement l'hypothèse  $\emptyset$  à une croyance nulle. Il s'agit des règles de Dempster, Yager, Zhang ou de Dubois & Prade. Concernant les règles de type moyenne, leur impartialité se démontre facilement. En effet, pour tout ensemble  $A$ , pour toute FDM  $m_1$ , il est toujours possible de trouver une FDM  $m_2$  qui contredise l'absorption ou le rejet. Pour statuer plus facilement sur le cas d'autres règles, nous introduisons d'autres définitions.

**Définition :** *Une règle de combinaison est dite directe, si et seulement si  $\forall A \subseteq \Omega$  il existe une application  $\psi_A : \Omega \times \Omega \rightarrow [0, 1]$  telle que*

$$m_1 \odot m_2(A) = \sum_{B \subseteq \Omega, C \subseteq \Omega} \psi_A(B, C) m_1(B) m_2(C) \quad (4.51)$$

$$\sum_{A \subseteq \Omega} \psi_A(B, C) = 1 \quad (4.52)$$

Cette définition couvre une large gamme de règles, mais nous n'avons pas l'intention de définir une famille de règles. Le seul intérêt de cette définition est de faciliter l'écriture de preuves à venir. Les règles conjonctive prudente, disjonctive hardie ou de type moyenne sont des exemples de règles non-directes. L'équation 4.52 traduit le fait que tout le produit  $m_1(B) m_2(C)$  doit être transféré sans perte lors de la combinaison. Elle est une conséquence directe des équations (3.51) et (4.51). La fonction  $\psi_A(.,.)$  évaluée en  $B$  et  $C$  représente la proportion du produit  $m_1(B) m_2(C)$  qui est transférée vers  $A$ . Ainsi, l'ensemble des fonctions  $\psi_A(.,.)$  décrit intégralement la règle de combinaison à laquelle il est associé. Nous appelons ces fonctions  $\psi_A(.,.)$ , **fonctions caractéristiques d'une règle directe**.

Certaines propriétés classiques se traduisent de manière simple au niveau des fonctions caractéristiques :

**Proposition :** *Pour une règle de combinaison directe donnée, la commutativité est équivalente à  $\forall A, B, C \subseteq \Omega, \psi_A(B, C) = \psi_A(C, B)$*

**Preuve :** cf. annexe C.1

**Proposition :** Pour une règle de combinaison directe donnée, l'associativité est équivalente à  $\forall A, C, D, E \subseteq \Omega, \sum_{B \subseteq \Omega} \psi_A(B, C) \psi_B(D, E) = \sum_{B \subseteq \Omega} \psi_A(D, B) \psi_B(E, C)$

**Preuve :** cf. annexe C.2

Grâce entre autres à l'expression de la commutativité pour les règles directes, nous pouvons proposer les résultats suivants :

**Proposition :** un ensemble  $A \subseteq \Omega$  est absorbant pour une règle de combinaison directe commutative si et seulement si  $\forall B \subseteq \Omega, \psi_A(A, B) = 1$

**Preuve :** cf. annexe C.3

**Proposition :** Un ensemble  $A \subseteq \Omega$  est un sous-ensemble de rejet pour une règle de combinaison directe commutative si et seulement si  $\forall B, C \subseteq \Omega$  tels que  $(B, C) \neq (A, A), \psi_A(B, C) = 0$

**Preuve :** cf. annexe C.4

Suite à ces deux résultats, l'impartialité d'une règle de combinaison directe commutative écrite sous forme matricielle se résume en deux contraintes :

- la matrice ne doit pas contenir de 1 sur la diagonale.
- la ligne de la matrice correspondant à l'ensemble  $A$  ne peut être de la forme :  $m(A)$  pour l'élément diagonal et 0 ailleurs.

Il suffit à présent d'observer la forme matricielle des règles de combinaison directes commutatives, pour examiner leur impartialité. Nous examinons les règles pour le cadre de discernement simple  $\Omega = \{a, b\}$ . La famille de règles de Lefevre s'écrit sous forme matricielle :

$$M_{\odot l}^m = \begin{pmatrix} \beta(\emptyset) & \beta(\emptyset)(m(\emptyset)+m(b)) & \beta(\emptyset)(m(\emptyset)+m(a)) & \beta(\emptyset)m(\emptyset) \\ \beta(a) & m(a)+m(\Omega)+\beta(a)[m(\emptyset)+m(b)] & \beta(a)(m(\emptyset)+m(a)) & m(a)+\beta(a)m(\emptyset) \\ \beta(b) & \beta(\emptyset)(m(\emptyset)+m(b)) & m(b)+m(\Omega)+\beta(b)[m(\emptyset)+m(a)] & m(b)+\beta(b)m(\emptyset) \\ \beta(\Omega) & \beta(\Omega)(m(\emptyset)+m(b)) & \beta(\Omega)(m(\emptyset)+m(a)) & m(\Omega)+\beta(\Omega)m(\emptyset) \end{pmatrix} \quad (4.53)$$

Les règles de Lefevre sont commutatives et de plus directes, elles correspondent au cas où :

- si  $B \cap C = A$  et  $A \neq \emptyset, \psi_A(B, C) = 1$
- si  $B \cap C = \emptyset$  et  $A \subseteq \Omega, \psi_A(B, C) = \beta(A)$ .
- sinon  $\psi_A(B, C) = 0$ .

D'après leur forme matricielle, une condition suffisante d'impartialité est :  $\forall A \subseteq \Omega, \beta(A) \in ]0, 1[$ .

Pour les règles  $\alpha$ -jonctives, il est plus difficile de conclure car nous ne disposons pas de formule générale. En revanche, la règle disjonctive exclusive s'écrit matriciellement :

$$M_{\underline{\odot}}^m = \begin{pmatrix} m(\emptyset) & m(a) & m(b) & m(\Omega) \\ m(a) & m(\emptyset) & m(\Omega) & m(b) \\ m(b) & m(\Omega) & m(\emptyset) & m(a) \\ m(\Omega) & m(b) & m(a) & m(\emptyset) \end{pmatrix} \quad (4.54)$$

Elle est donc impartiale. Nous avons donné les éléments nécessaires pour statuer sur l'impartialité de l'ensemble des règles que nous avons présélectionnées pour notre étude. Nous introduisons<sup>3</sup>

<sup>3</sup>A notre connaissance, une telle règle n'a pas encore été introduite.

néanmoins une nouvelle règle directe, que nous appelons **règle intermédiaire** et que nous notons  $\odot_{inter}$  :

$$m_{inter}(A) = \frac{1}{2}m_{1\odot 2} + \frac{1}{2}m_{1\oplus 2} \quad (4.55)$$

Cette règle appartient à la famille définie par Martin. Elle présente comme intérêt majeur de représenter une croyance moyenne entre conjonction et disjonction. Matriciellement, elle s'écrit :

$$M_k^m = \begin{pmatrix} \frac{m(\emptyset)}{2} + \frac{1}{2} & \frac{m(\emptyset)}{2} + \frac{m(b)}{2} & \frac{m(\emptyset)}{2} + \frac{m(a)}{2} & \frac{m(\emptyset)}{2} \\ \frac{m(a)}{2} & m(a) + \frac{m(\emptyset)}{2} + \frac{m(\Omega)}{2} & 0 & \frac{m(a)}{2} \\ \frac{m(b)}{2} & 0 & m(b) + \frac{m(\emptyset)}{2} + \frac{m(\Omega)}{2} & \frac{m(b)}{2} \\ \frac{m(\Omega)}{2} & \frac{m(b)}{2} + \frac{m(\Omega)}{2} & \frac{m(a)}{2} + \frac{m(\Omega)}{2} & \frac{m(\Omega)}{2} + \frac{1}{2} \end{pmatrix} \quad (4.56)$$

Elle est donc elle aussi impartiale, car commutative.

Il est cependant intéressant de noter que cette règle n'est pas associative, de même que le sous-ensemble des règles de Lefevre dont nous avons établi l'impartialité. Il paraît donc opportun d'examiner la compatibilité de ces propriétés. A ce stade, la seule règle associative et impartiale est la règle disjonctive exclusive, dont le comportement en terme de fusion est pourtant insatisfaisant.

Afin d'écartier des règles au comportement contre-intuitif, une propriété dite de consistance faible est également introduite :

**Définition :** Une règle de combinaison directe est faiblement consistante, si et seulement si  $\forall A \neq \emptyset, \psi_A(A, A) = 1$  et  $\forall A, B, C, \psi_A(B, C) > 0 \Rightarrow A \subseteq B \cap C$  ou  $B \cup C \subseteq A$

Le terme de consistance traduit le caractère logique de la combinaison. Le terme faible y a été adjoint car les contraintes formulées sont souples. La règle disjonctive exclusive est un exemple de règle directe, qui n'est pas faiblement consistante. Dans la pratique, cette règle ainsi que d'autres règles  $\alpha$ -jonctives donnent en effet des résultats difficilement interprétables, comme Smets l'explique lui-même.

**Remarque :** Pour une règle de combinaison directe faiblement consistante,

$$\forall B \neq A, \psi_A(B, B) = 0 \quad (4.57)$$

**Preuve :** conséquence directe de l'équation 4.52

La première condition de la consistance faible implique que l'intégralité du produit  $m_1(A)m_2(A)$  soit transféré à  $A$ . Il n'y a en effet pas de meilleur candidat pour recevoir cette croyance. La seconde condition implique que  $m_1(B)m_2(C)$  ne peut être transférée à  $A$  si  $A$  n'est pas un sous-ensemble de  $B \cap C$  ou ne contient pas  $B \cup C$ . Cette contrainte rend impossible le transfert à un ensemble dont une partie est disjointe de  $B \cup C$  et permet un traitement équilibré des ensembles.

En combinant les propriétés de consistance faible, d'impartialité, d'associativité et de commutativité, une incompatibilité pour les règles de combinaison directe n'a pu être démontrée. Tout au plus, il a été constaté que le choix des fonctions  $\psi_A(\cdot)$  subit des contraintes, ce qui n'est pas surprenant. Ces résultats sont regroupés en annexe C.5. Dans le cas général et d'après divers

calculs que nous avons eu l'occasion de mener, il paraît peu probable qu'une règle combinant toutes ces propriétés ait un comportement satisfaisant en termes de fusion, surtout pour des cadres de discernement de petite taille.

Pour conclure, la notion d'impartialité semble particulièrement intéressante pour le cas où de nombreuses sources sont à fusionner. L'impartialité prévient alors le comportement abusif de certains sous-ensembles et évite une dérive du résultat de la fusion vers un élément absorbant. Smets dans le dernier paragraphe de conclusion de [Smets 06] soulève ce problème et explique qu'une renormalisation de temps à autres des FDM (à l'instar de la règle de Dempster) s'avère nécessaire. L'auteur parle d'"astuce de calcul" mais, de notre point de vue, le problème trouve son origine dans la non-impartialité de la règle conjonctive dont il défend les mérites.

L'ensemble des résultats de cette sous-partie montre l'influence de certains sous-ensembles dans le cœur du fonctionnement des règles de combinaison. La limite d'influence établie est représentée par la propriété d'impartialité. Toutefois, une règle alliant impartialité et certaines propriétés algébriques classiques sans pour autant gérer la fusion de manière aberrante n'a pu être obtenue. Une des règles parmi les plus satisfaisantes à cet égard est la règle intermédiaire qui est impartiale, commutative, faiblement consistante et quasi-associative.

#### 4.3.5 Autres propriétés

D'autres propriétés d'usage plus modeste existent :

- Continuité : une règle est dite continue si  $\forall m_1, m_2 \in \mathfrak{B}^\Omega$ , pour  $m'_1$  proche de  $m_1$ ,  $m_1 \odot m_2$  est proche de  $m'_1 \odot m_2$ . Dans la référence d'où nous tirons cette définition [Smarandache 04a], rien n'indique ce qu'il faut entendre par "proche", nous estimons donc cette propriété mal définie.
- Plausibilité d'un faux positif : Soit  $m_\odot = m \odot Z^0$ . Smets estime que  $\forall m \in \mathfrak{B}^\Omega, \forall Z \subseteq \Omega$ , la plausibilité calculée d'après  $m_\odot$  et évaluée en  $\bar{Z}$  doit être nulle. Cette propriété est liée à la notion de conditionnement, en effet si  $\odot$  est la règle conjonctive,  $m_\odot$  est la FDM  $m$  conditionnée par  $Z$ . L'objectif de cette propriété recoupe celui de la conjonction des informations.
- Résistance au raffinement : Cette propriété est remplie si le résultat d'une combinaison puis d'un raffinement est le même que le résultat du même raffinement sur les opérandes suivi de la même combinaison. Dans notre cadre applicatif, nous n'utilisons pas de raffinement.
- Duplication du conditionnement : Deux conditionnements successifs identiques doivent être équivalents à un seul conditionnement. En d'autres termes,  $\forall m \in \mathfrak{B}^\Omega, \forall Z \subseteq \Omega$ ,  $(m \odot Z^0) \odot Z^0 = m \odot Z^0$ . Des remarques équivalentes à celles faites sur l'idempotence peuvent être formulées par rapport à cette propriété.
- Conditionnement itéré : cette propriété correspond au fait que  $\forall Z_1, Z_2 \subseteq \Omega, \forall m \in \mathfrak{B}^\Omega$   $(m \odot Z_1^0) \odot Z_2^0 = m \odot (Z_1 \cap Z_2)^0$ . Cette propriété découle également d'une volonté de procéder à une combinaison de nature conjonctive. Son intérêt dépend du contexte.
- Conditionnement conflictuel : Il ne s'agit pas vraiment d'une propriété ici, mais plutôt d'un examen du résultat de la fusion suivante :  $(Z^0 \odot \bar{Z}^0) \odot m$ . La valeur de cette fusion varie assez largement selon les règles. Le premier terme rappelle ce que Smets appelle *le contre-exemple de la bombe* où deux FDM en conflit total sont agrégées (cf. [Smets 06]).

**Tableau 4.1** : Différentes règles de combinaison et leurs propriétés - partie n°1. × : la règle possède la propriété correspondante.

règle de combinaison	propriétés algébriques				propriétés liées à la nature de l'élément d'évidence
	commutativité	associativité	quasi-associativité	idempotence	distinction non requise
conjonctive	×	×			
disjonctive	×	×			
disjonctive exclusive	×	×			
conjonctive prudente	×	×		×	×
disjonctive hardie	×	×		×	×
Zhang	×				
Yager	×		×		
Dubois et Prade	×		×		
Intermédiaire	×		×		
adaptative de Florea	×		×		
mixte de Martin	×				
PCR	×				
DPCR	×				
MDPCR	×				
moyenne de Delmotte	×		×	×	
mixte Delmotte	×		×		
Murphy	×		×	×	
WAO	×		×		

Il appartient à l'utilisateur de définir s'il est préférable que le résultat de cette fusion s'approche de  $\emptyset^0$  ou de  $m$ .

Ces propriétés sont soit définies de façon imprécise, soit intéressantes dans un contexte différent du notre, ou encore recourent éventuellement d'autres propriétés évoquées précédemment. Conséquemment, nous les excluons de l'étude.

#### 4.3.6 Comparatif des règles

Les tables 4.1 et 4.2 résument l'existence de différentes propriétés pour les règles sélectionnées dans le cadre de notre étude.

Toutes les propriétés étudiées ou introduites au cours de cette étude trouvent une justification en terme de fusion d'informations. La difficulté à combiner de trop nombreuses propriétés et aussi l'impossibilité d'une coexistence de la conjonction et de la disjonction tendent à montrer qu'une règle de combinaison unique ne peut satisfaire à l'ensemble des problématiques de fusion de données. Dans la partie suivante, les choix des propriétés opportunes en cas de sources non fiables sont examinés et la présente étude nous aidera à déterminer quel schéma de combinaison apporte une réponse appropriée à ces besoins.

### 4.4 Stratégie de combinaison contextuel

En termes de suivi d'objets, nous rappelons que notre objectif est de combattre l'échec des algorithmes face à des situations précises liées à des événements perturbateurs. Chaque source d'information peut être sujette à ces erreurs, certains événements ayant plus d'influence que d'autres selon les sources. Concrètement, en termes de fusion d'informations, une source erronée ne réagit pas à la présence réelle de l'objet dans la scène ou bien indique une mauvaise position de l'objet qui, le plus souvent, correspond à un autre objet visuellement proche de celui

**Tableau 4.2** : Différentes règles de combinaison et leurs propriétés - partie n°2. × : la règle possède la propriété correspondante. ★ : la règle possède la propriété correspondante sous réserve qu'un ou plusieurs paramètres soient correctement choisis.

règle de combinaison	propriétés liées à la nature de la combinaison			propriétés liées aux sous-ensembles		
	comportement conjonctif	comportement disjonctif	comportement intermédiaire	impartialité	consistance faible	définie pour toute FDM
conjonctive	×				×	×
disjonctive		×			×	×
disjonctive exclusive			×			×
conjonctive prudente	×				×	
disjonctive hardie		×			×	
Zhang	×				×	×
Yager	×				×	×
Dubois et Prade	×				×	×
Intermédiaire			×	×	×	×
adaptative de Florea			★	★	×	×
mixte de Martin			×	×	×	×
PCR	×				×	×
DPCR			★	★	×	×
MDPCR			★	★	×	×
moyenne de Delmotte			×	★	×	×
mixte de Delmotte			★	★	×	×
Murphy			×	×	×	×
WAO	×			×	×	×

ciblé. Il s'agit donc typiquement d'une situation où des sources non fiables doivent être fusionnées.

Dans cette partie, nous examinons dans un premier temps comment se comportent ces sources non fiables, et quelles propriétés sont nécessaires aux règles de combinaison pour pouvoir agréger de telles sources. Enfin nous proposons une solution adaptée à nos besoins, appelée stratégie de combinaison contextuel.

#### 4.4.1 Combinaison de sources anormales

Comme tendent à le suggérer les résultats du chapitre précédent, trois catégories de sources se distinguent principalement en suivi d'objets :

- les sources **défaillantes**, qui correspondent à une forme d'indétermination majeure de la source en question. Elles fournissent une quantité très faible d'information mais qui n'est pas pour autant négligeable. La défaillance peut être momentanée (événement perturbateur passager, capteur saturé) ou définitive (capteur endommagé).
- les sources **leurrées**, qui correspondent à une forme sévère de non fiabilité des sources en question. Ces sources peuvent fournir une quantité importante de mauvaise information, par exemple en confondant l'objet-cible avec un autre élément de la scène dans le cadre du suivi d'objets. Les informations erronées sont plus dommageables pour le processus de fusion qu'un manque d'information. Des informations inexacts sont en général engendrées par un événement particulier dans la scène, et leur présence prend fin avec l'événement. Elles peuvent être plus persistantes quand le modèle de l'objet est imprécis et ne le distingue pas clairement du reste de la scène.
- les sources **normales**, qui regroupent l'ensemble des autres sources. La part d'indétermi-

nation et de non fiabilité de ces sources est négligée. Par opposition, les sources leurrées ou défaillantes sont dites **anormales**.

C'est donc par rapport à ces trois comportements que nous devons aborder le problème de la fusion en présence de sources anormales.

De nombreux auteurs se sont intéressés à cette question et nous présentons ici quelques approches existantes parmi les plus significatives. Pour ce qui est du caractère indéterminé de certaines FDM, seules les règles de Zhang et de Martin prennent en compte ce phénomène dans la définition même des règles. Comme expliqué en section 4.2, c'est en intégrant la cardinalité des ensembles au sein de la formule que l'imprécision des hypothèses influence le processus de combinaison. En effet, plus une hypothèse est imprécise, plus le cardinal de l'ensemble auquel elle correspond est important.

L'imprécision au sens des valeurs prises par les FDM est un problème fondamentalement différent. Deneux [Deneux 99] et Lamata et Moral [Lamata 94] utilisent d'autres objets mathématiques que les FDM pour décrire les croyances. Ces objets sont appelés FDM imprécises et associent à un ensemble un sous-intervalle de  $[0, 1]$  au lieu d'une unique valeur appartenant à  $[0, 1]$ . Plus l'intervalle associé à un ensemble est grand, plus la croyance associée à cet ensemble possède une valeur imprécise. La méthode a été étendue par Smarandache et Dezert [Smarandache 05a] à des sous-ensembles quelconques de  $[0, 1]$ .

Concernant les sources non fiables, certains auteurs ont intégré directement la fiabilité des sources à l'intérieur des règles de combinaison à l'aide de coefficients de fiabilité. C'est notamment le cas pour les règles de Delmotte que nous avons introduites en section 4.2 (*cf.* équations (4.14) et (4.14)). Un coefficient de fiabilité est défini pour chaque source, et généralement sa valeur est comprise entre 0 et 1, permettant ainsi une évaluation graduée de la fiabilité d'une source. Les coefficients peuvent être également utilisés à l'extérieur des règles de combinaison, comme le propose Appriou [Appriou 97]. Dans cette approche, les coefficients sont intégrés au niveau de la constructions des FDM à l'aide de deux modèles de FDM différents.

Pour ces approches, l'obtention de coefficients de fiabilité reflétant correctement cette notion de fiabilité est un enjeu majeur. Il existe principalement trois manières d'obtenir de tels coefficients :

- par des connaissances *a priori*, elles-mêmes obtenues souvent par apprentissage (voir [Elouedi 04, Lefevre 00] par exemple),
- par des informations acquises parallèlement aux informations des sources elles-mêmes, comme de l'information contextuelle (voir [Fabre 98] pour exemple),
- par une estimation *a posteriori*. Cette estimation est entre autres possible à l'aide de méthodes de consensus entre les résultats de la fusion et les FDM à l'origine de la fusion (voir [Delmotte 96] pour exemple).

Une discussion plus approfondie sur cette question est disponible dans [Rogova 08].

L'approche majoritairement retenue par les auteurs traitant des sources non fiables est la gestion du conflit. De nombreuses règles relatives à cet aspect ont été introduites en sec-

tion 4.2 [Yager 87a, Dubois 88, Inagaki 91, Smarandache 05b, Martin 07, Florea 06]. L'approche proposée par Lefevre *et al.* [Lefevre 02] est à souligner car la redistribution du conflit  $y$  est traitée de manière plus globale à l'aide d'une méthode d'optimisation. Les auteurs déterminent une règle optimale en minimisant une fonction de coût. Cela ne fonctionne que si des données étiquetées sont disponibles et que des contraintes précises se dégagent sur les critères à optimiser.

Le conflit peut être vu comme un indicateur de non fiabilité, mais ne permet pas de détecter toutes les situations relatives à de la non fiabilité. Prenons l'exemple de deux personnes représentant deux agents, chargés de donner leur avis sur un problème simple : déterminer la couleur d'un objet. Admettons que le cadre de discernement soit  $\{\text{rouge}, \text{vert}\}$ . Si un agent choisit  $\{\text{rouge}\}$  et l'autre  $\{\text{vert}\}$ , il y a conflit et la décision après fusion aura un certain niveau de non fiabilité. Si maintenant les deux agents votent pour la même couleur mais que ces mêmes agents souffrent de Daltonisme<sup>4</sup>, aucun conflit n'apparaît et pour autant la croyance établie ne sera pas fiable. Cet exemple correspond assez bien aux situations de fouillis en suivi d'objets.

D'autre part, la fiabilité peut également être traitée à l'aide de l'opération d'affaiblissement. Il suffit pour cela d'affaiblir une source détectée comme non fiable. L'affaiblissement incorpore de l'incertitude au sein d'une FDM en augmentant le crédit alloué à  $\Omega$ . Dans une combinaison de nature conjonctive, la FDM a ainsi moins d'impact sur le processus de fusion. Dans une combinaison de nature disjonctive, c'est l'inverse mais le résultat de la fusion sera lui aussi une FDM indécise, indiquant ainsi la présence d'informations non fiables. Toutefois, l'affaiblissement revient d'une certaine manière à censurer une source et peut engendrer une perte d'information.

Il existe des formes d'affaiblissement différentes de celle présentée en 3.4.6. Ainsi la règle DPCR use d'un affaiblissement plus modéré en transférant la croyance vers l'union des ensembles plutôt qu'à  $\Omega$  directement (*cf.* équation (4.8)). Mercier *et al.* [Mercier 08] proposent un affaiblissement adapté non seulement aux différentes sources, mais aussi selon les hypothèses du cadre de discernement. Même si cette opération exige plus d'information contextuelle, l'affaiblissement est de meilleure qualité et la fusion est plus efficace. De manière générale, l'affaiblissement ne doit pas être opposé aux autres méthodes de gestion des sources non fiables car elles sont complémentaires. Smets [Smets 06] propose une approche hiérarchique dans laquelle le conflit est géré par affaiblissement mais aussi par le choix d'une règle de combinaison adaptée. Notons tout de même, que l'obtention des coefficients d'affaiblissement pertinents pour le problème traité n'est pas une tâche triviale à l'image de l'obtention des coefficients de fiabilité.

Enfin, dans [Lefevre 00], les auteurs montrent que la redistribution du conflit à l'aide de poids sur les jeux de masses peut s'exprimer de manière équivalente à l'aide d'affaiblissements sur ces mêmes jeux suivi d'une combinaison par la règle conjonctive. Les deux méthodes principale de gestion du conflit évoquées sont donc liées.

Dans le cadre des études menées dans ce mémoire, la quantité d'information *a priori* ou contextuelle est trop faible pour mettre en place les approches citées ci-dessus. L'information *a priori* n'est utilisée que pour la construction des FDM et concerne la paramétrisation des modèles

---

<sup>4</sup>Anomalie génétique altérant la perception des couleurs, notamment rouge et verte ; aussi connue sous le nom de dyschromatopsie.

de FDM (cf. équation (3.83) et (3.90)). L'information contextuelle est suffisante pour classer les sources parmi les trois états cités précédemment : défaillante, leurrée ou normale. Tout au plus, certaines peuvent être adaptées, mais fonctionnent moins bien car elles ne sont pas prévues à l'origine pour un tel cadre de travail. Par exemple, il n'est pas possible de déterminer de manière précise des coefficients de fiabilité ou d'affaiblissement, mais il reste possible de fixer une valeur pour ces coefficients conditionnellement à l'état d'une source.

La vision en trois états du problème de fusion en présence de sources non fiables ou indécisées nécessite une réponse plus spécifique. Pour mettre au point ce processus adapté, il nous faut préalablement déterminer les exigences relatives au problème.

### Exigences majeures

La classification des FDM en trois états permet de séparer le problème de fusion en différents sous-problèmes, qu'il est plus aisé de résoudre séparément. Grâce à l'information contextuelle, nous proposons alors d'examiner pour chaque type de couple de sources quel est le comportement attendu d'une règle :

- Cas n°1 : si les deux sources sont fiables, cela signifie que leurs FDM respectives représentent un élément d'évidence qui contient une forme de vérité. Dans ce cas, une règle de nature conjonctive est typiquement celle qui pourra extraire le plus efficacement l'information pertinente communes aux deux sources.
- Cas n°2 : si l'une des sources est défaillante et que l'autre ne l'est pas, le résultat attendu est de conforter la source quelconque sans que la source défaillante viennent perturber le résultat. Une source défaillante doit donc s'apparenter à un élément neutre de la règle. Une source défaillante n'apporte aucune information, sa FDM est donc proche de l'ignorance totale  $\Omega^0$ . La règle doit donc avoir cette FDM comme élément neutre. Si les deux sources sont défaillantes ou bien que la source défaillante est combinée à une source leurrée, le raisonnement reste valable.
- Cas n°3 : si l'une des sources est leurrée et que l'autre n'est pas défaillante, cela signifie que chacune contient une part de vérité mais qu'il n'est pas possible d'identifier quelle partie de leur élément d'évidence correspond à cette "vérité". Dans ce cas assez extrême, une règle de nature purement disjonctive est typiquement adaptée pour ne pas privilégier ou écarter d'information.

Les trois cas discutés recouvrent l'ensemble des situations possibles (à condition que les règles soient commutatives, ce qui est toujours le cas).

Il faudra en outre veiller à ce que, pour chaque règle choisie, il ne soit pas possible d'appliquer cette règle à une FDM accordant un poids trop important à un sous-ensemble absorbant de la règle. En effet, dans un tel cas cette FDM aura un poids beaucoup trop prépondérant dans la fusion de données. Il suffit d'une source accordant un crédit important à un sous-ensemble absorbant d'une règle pour que tout le résultat de la fusion soit compromis.

### Exigences mineures

Un certain nombre de critères de second ordre par rapport à ceux des paragraphes précédents méritent d'être cités. Ces critères peuvent nous permettre d'affiner notre recherche de solution parmi l'ensemble des solutions répondant aux exigences majeures.

Pour limiter l'impact sur le temps de calcul du recours aux fonctions de croyance, une méthode quasi-associative apparaît indispensable. Ce critère est d'autant plus important que le nombre de sources est grand. Plus de détails sur le temps de calcul seront donnés dans le chapitre suivant (*cf.* 5.2.4).

Comme vu en 4.3.4, l'impartialité d'un processus de combinaison s'avère intéressant dès lors que le nombre de sources à fusionner augmente. Dans nos expériences, ce nombre ne dépasse pas 5 sources, mais dans l'absolu il est toujours préférable de prévoir un processus impartial.

### Stratégies de combinaison

D'après l'analyse des résultats de l'étude de la section 4.3, il n'existe pas de règle de combinaison qui satisfasse globalement aux exigences majeures. En effet, dans certaines situations la règle doit être purement disjonctive et dans d'autres situations elle doit être conjonctive, ce qui est impossible. Les règles aux comportements intermédiaires offrent une disjonctivité ou une conjonctivité partielle.

Il est néanmoins possible de mettre au point des systèmes plus généraux afin de contourner cette impossibilité. Sachant les états des sources, une règle peut être choisie parmi un ensemble prédéfini de règles  $\mathfrak{R}$ , et donc appliquée uniquement à un couple de FDM donné, tandis qu'une autre règle peut être sélectionnée et appliquée à un autre couple de FDM dont les états diffèrent. Nous appelons ce processus **stratégie de combinaison**.

Par exemple, suite à la classification des FDM parmi les trois états, une stratégie peut proposer de fusionner trois sources de la manière suivante :  $(m_1 \odot_1 m_2) \odot_2 m_3$  avec  $\mathfrak{R} = \{\odot_1, \odot_2\}$ . Les règles contenues dans  $\mathfrak{R}$  peuvent être choisies de manière à répondre aux exigences majeures au cas par cas (une règle par situation). De cette manière, le cumul des propriétés n'est plus un obstacle.

Par ailleurs, il est clair que le parenthésage des opérations de combinaison est très important, et que donc une stratégie ne peut être associative. En revanche, si les règles qui le composent sont au moins quasi-associatives, alors la stratégie est elle aussi quasi-associative.

De plus, le fait d'utiliser conjointement différentes règles peut avoir un avantage du point de vue de l'impartialité. Nous définissons ici un concept que nous nommons **dualité des règles**.

**Définition :** *Les règles  $\odot$  et  $\odot_d$  sont duales si*

*(i) l'ensemble des sous-ensembles absorbants<sup>5</sup> pour  $\odot$  est égal à l'ensemble des sous-ensembles*

---

<sup>5</sup>Une règle commutative ne peut avoir plus d'un sous-ensemble absorbant. Conséquemment, la définition de la dualité se simplifie pour de nombreuses règles.

de rejet pour  $\odot_d$  ;

(ii) l'ensemble des sous-ensembles de rejet pour  $\odot$  est égal à l'ensemble des sous-ensembles absorbants pour  $\odot_d$ .

En conséquence, si un couple de règles duales est inclus dans  $\mathfrak{R}$  et que ces deux règles sont utilisées en même temps, alors la stratégie ne possède pas de sous-ensemble qui soit globalement de rejet ou absorbant. En effet, l'absorption de l'une des règles est compensée par le rejet de l'autre et inversement. Nous qualifions une telle stratégie de **pseudo-impartiale**. La pseudo-impartialité n'est valable que dans des cas restreints :

- un couple de règles duales est contenu dans  $\mathfrak{R}$ ,
- au moins deux sources sont classifiées dans deux états différents,
- chacune des règles duales correspond à un de ces états.

Toutefois, si tous les états des sources sont identiques, la stratégie est alors équivalente à une seule règle et l'impartialité n'est pas nécessairement requise. Dans un tel cas, le plus important est d'éviter qu'un sous-ensemble absorbant engendre une perte d'information, ce qui prévu par les exigences majeures. Au final, la stratégie est pseudo-impartiale dans les situations qui l'exigent réellement.

Nous allons à présent proposer une stratégie de combinaison adaptée à notre contexte applicatif grâce à l'information contextuelle apportée par l'algorithme du filtre particulière.

#### 4.4.2 Utilisation du contexte dans une stratégie de combinaison de sources non fiables

Globalement, une stratégie de combinaison se décline en une étape de classification suivie d'une étape de fusion. Nous présentons ici ces deux étapes.

##### Étape de classification de FDM

D'après l'étude de cas menée en 4.4.1 (exigences majeures), il faut d'abord détecter les FDM défaillantes puis, parmi les FDM non-défaillantes, celles étant leurrées. Nous avons donc besoin pour cela d'un test de défaillance, suivi d'un test de leurre. La mise en oeuvre de ces tests est présentée ci-après.

– **Test de défaillance** Concernant la classification, les sources défaillantes sont les plus faciles à détecter. Comme évoqué précédemment, une source totalement défaillante produit une FDM égale à  $\Omega^0$ . Il faut donc comparer  $\Omega^0$  aux FDM pour tester la défaillance. La comparaison de FDM n'est pas une tâche triviale et des travaux ont été menés afin de définir des similarités ou des distances entre FDM, notamment par Diaz [Diaz 06] et Joussetme [Joussetme 01]. Toutefois, comme l'une des deux FDM à comparer possède une répartition élémentaire des croyances, une solution plus simple est préférable. Le seul élément focal de  $\Omega^0$  étant  $\Omega$ , une FDM quelconque  $m$  peut être classé "défaillante" simplement en seuillant  $m(\Omega)$ .

$$\Omega^0(\Omega) - m(\Omega) = 1 - m(\Omega) < t_{def} \Rightarrow m \text{ est "défaillante"}. \quad (4.58)$$

avec  $t_{def}$  un seuil déterminé *a priori*. Pour une classification plus fiable, il convient d'utiliser l'information contextuelle du filtre particulière. Une FDM est définie pour chaque source mais

aussi pour chaque particule  $X_t^{(i)}$ . Notons cette FDM  $m \left[ S_j, X_t^{(i)} \right] (\cdot)$ . La décision sur  $S_j$  doit être globale, nous proposons donc de classer  $S_j$  défaillante si :

$$1 - \min_{i=1..N} m \left[ S_j, X_t^{(i)} \right] (\Omega) < t_{def} \quad (4.59)$$

Au final, la source  $S_j$  est dite défaillante si elle n'apporte de discrimination pour aucune des particules échantillonnées. Comme tout seuillage, la valeur du seuil peut être relativement délicate à choisir. De manière générale,  $t_{def}$  est choisi entre 0,001 et 0,05. La valeur doit être surtout suffisamment élevée pour que des FDM proches de  $\Omega^0$  ne soit pas classées comme défaillantes, car la règle disjonctive est alors vouée à multiplier des valeurs faibles et peut atteindre la limite calculatoire du système. De plus, dans un tel cas, une seule source pourrait endommager la fusion (*cf.* explications relatives à la deuxième phase de l'étape de fusion 4.4.2).

En pratique, des valeurs relativement proches de ce seuil donnent des résultats quasiment similaires en terme de suivi, ce qui montre que le réglage de ce paramètre n'est pas primordial pour la qualité globale de l'approche. Notons que c'est le même principe de seuillage qui permet d'éliminer les sources défaillantes pour l'approche probabiliste du FPFb modifié, cela dit la valeur du seuil doit souvent être choisie bien plus faible pour éviter une perte de suivi au risque de prendre une décision risquée sur la base de faibles valeurs de probabilité (*cf.* 3.5.4).

– **Test de leurre** Parmi les sources non défaillantes, il reste à se doter d'une "mesure de leurre" afin de pouvoir les classer en "fiables" ou "leurrées". Dans notre contexte, cette mesure est possible grâce au fait que plusieurs FDM sont calculées pour chaque particule. Comme nous en faisons le constat dans la partie expérimentale du chapitre précédent 3.5.4, en cas de leurre une source indique plusieurs positions comme potentiellement candidate pour être celle de l'objet recherché, et conséquemment la densité  $p(X_t|Y_t)$  va présenter plusieurs modes centrés sur ces positions candidates. La variance  $var_t$  des particules pondérées par les poids  $w_t^{(i)}$  peut alors faire office de mesure de leurre :

$$var_t = \sum_{i=1}^N w_t^{(i)} \left( \bar{X}_t - X_t^{(i)} \right)^2 \quad (4.60)$$

La classification est alors possible par simple seuillage :

$$var_t > t_{var} \quad (4.61)$$

$$(4.62)$$

avec  $t_{var}$  un seuil prédéfini.

La détermination d'un seuil adéquat a ici plus de conséquences sur l'ensemble de l'approche, mais elle est plus facile car la valeur de la mesure varie nettement entre une situation avec ou sans leurre.

Dans nos expériences,  $t_{var}$  est fixé de manière dynamique de sorte que  $t_{var} = a \min(height_t, width_t)$  avec  $a$  un facteur choisi dans [1, 2]. Si la variance dépasse cette valeur, alors deux positions de l'objet distinctes sont retenues par le filtre, ces positions étant au minimum "côte à côte".

### Étape de fusion de FDM conditionnellement aux classes des FDM

Nous disposons à présent de sources rangées dans les trois catégories : fiables, défaillantes ou leurrées. Il faut désormais définir le processus de fusion. Cela commence par le choix des règles formant l'ensemble  $\mathfrak{R}$ . Examinons les trois cas présentés dans la sous-partie précédente :

- Cas n°1 : la règle doit être de nature conjonctive, dans le cadre d'un monde ouvert, toutes les règles de nature conjonctive ou intermédiaire sont des candidates.  $\emptyset$  est potentiellement un sous-ensemble absorbant pour ces règles, mais d'après les modèles de FDM définis, le crédit attribué à cet ensemble est limité à ce stade du processus.
- Cas n°2 :  $\Omega^0$  doit être élément neutre de la règle,  $\odot$  et  $\oslash$  sont des candidates. Le sous-ensemble absorbant de ces règles est  $\emptyset$ . De même que pour le cas n°1, le crédit attribué à cet ensemble est limité de par le modèle de construction des FDM.
- Cas n°3 : la règle doit être de nature purement disjonctive, dans le cadre d'un monde ouvert,  $\odot$  et  $\oslash$  sont des candidates. Ces règles ont pour sous-ensemble absorbant  $\Omega$ , et les FDM accordant un crédit important à cet ensemble tombent dans le cas n°2.

Nous éliminons toutes les règles définies pour un monde fermé, car la valeur de la FDM de sortie pour  $\emptyset$  peut être exploitée par l'algorithme après fusion. Il apparaît que deux règles suffisent pour constituer notre ensemble de règles. Pour des raisons de commodité, nous choisissons  $\mathfrak{R} = \{\odot, \oslash\}$ . Il serait intéressant d'observer les résultats en utilisant les règles conjonctive prudente et/ou disjonctive hardie mais, faute de temps, ces expériences n'ont pu être menées durant ces travaux de thèse.

La première phase de l'étape de fusion est à présent réalisable puisqu'une règle a été affectée à chaque classe de FDM. Pour les classes "normale" ou "défaillante", la même règle a été choisie : la règle conjonctive. Par associativité de cette règle, les FDM normales et défaillantes peuvent être combinées ensemble. Pour la classe "leurrée", c'est la règle disjonctive qui doit être appliquée sur l'ensemble de FDM classées dans cette catégorie.

Notons que, pour gagner du temps, les sources défaillantes peuvent être purement évincées du processus de fusion. Néanmoins, elles restent informatives comme les expériences le montreront. Quand bien même, la classification en "défaillante" est indispensable, car sinon une grande partie du crédit pourrait être absorbée par  $\Omega$  si une source défaillante était classée leurrée. C'est aussi pour cela que le test de défaillance précède celui de "leurre".

Pour la deuxième phase de fusion, il reste donc deux FDM seulement à fusionner : une première pour les sources leurrées, et une seconde pour les autres sources. Pour savoir quelle règle utiliser entre ces deux FDM, nous avons besoin de connaître la classe de ces FDM. Par définition, le résultat de la fusion des sources leurrées constitue lui-même une source leurrée, mais il peut arriver qu'elle soit en plus défaillante. Concernant l'autre résultat, il peut constituer une source défaillante ou normale. En somme, la dernière combinaison requiert la règle conjonctive si l'une des deux FDM est défaillante (cas n°2). Dans la cas contraire, la règle disjonctive doit être utilisée (cas n°3). La distinction entre ces deux cas se fait à l'aide de deux derniers tests de défaillance dont les résultats sont mis en commun par un "ou logique". Le processus de fusion

est à présent complet et est résumé par le schéma de la figure 4.1 et par l'algorithme 8.

**Figure 4.1** : Schéma décrivant la stratégie de combinaison contextuelle (SCC).

---

 Algorithme : Stratégie de combinaison contextuelle
 

---

```

pour  $j = 1$  à  $M$  faire
  si Test de défaillance = VRAI alors
    Classer en  $m[S_j]$  dans la classe défaillante  $\{Défaillante\}$ 
  sinon
    si Test de leurre = VRAI alors
      Classer en  $m[S_j]$  dans la classe leurrée  $\{Leurrée\}$ 
    sinon
      Classer en  $m[S_j]$  dans la classe normale  $\{Normale\}$ 
    fin si
  fin si
fin pour
Utiliser la règle conjonctive  $\odot$  sur l'ensemble des FDM normales ou défaillante :  $m_{\odot} =$ 
 $\odot_{k=1}^{\#(\{Normale\})+\#(\{Défaillante\})} m[S_j \in \{Normale\} \cup \{Défaillante\}]$ 
Utiliser la règle disjonctive  $\oplus$  sur l'ensemble des FDM leurrées :  $m_{\oplus} =$ 
 $\oplus_{k=1}^{\#(\{Leurrée\})} m[S_j \in \{Leurrée\}]$ 
si  $\#(\{Normale\}) + \#(\{Défaillante\}) = 0$  alors
  Retourner  $m_{\odot}$ 
sinon si  $\#(\{Leurrée\}) = 0$  alors
  Retourner  $m_{\oplus}$ 
sinon
  si Test de défaillance pour  $m_{\odot}$  ou pour  $m_{\oplus} =$  VRAI alors
    Retourner  $m_{\odot} \odot m_{\oplus}$ 
  sinon
    Retourner  $m_{\odot} \oplus m_{\oplus}$ 
  fin si
fin si
Fin

```

---

Algorithme 8: Stratégie de combinaison contextuelle.

La spécificité de cette stratégie par rapport aux règles reposant en grande partie sur la prise en compte du contexte, nous avons baptisé cette approche stratégie de combinaison contextuelle (SCC).

L'approche proposée rappelle dans la forme celle de Smets dans [Smets 07]. Cependant, Smets n'utilise que deux états de source et une seule règle à la fois. Sur le fond, la SCC diffère des méthodes présentées en 4.4.1 en quelques points principaux :

- Sans compter l'étape de classification, la SCC n'est pas plus coûteuse qu'une seule règle de combinaison si toutes les sources sont disponibles en même temps. En effet le nombre de combinaisons nécessaires est le même pour la SCC que pour une règle. Si les sources arrivent à des instants différents, une source arrivante est intégrée au calcul du résultat de sa classe par associativité. Ensuite, il suffit de recalculer la combinaison des résultats des classes, ce qui revient à une seule combinaison entre deux opérandes.
- La SCC ne requiert que des données contextuelles peu précises, et l'*a priori* reste l'affaire des modèles de FDM et donc n'est pas vraiment imputable à la SCC.
- La SCC nécessite le réglage de deux paramètres au niveau de la classification des FDM. Ces paramètres se règlent toutefois sans difficulté.
- La SCC traite à la fois le problème de l'indétermination et de la non fiabilité.
- D'autres approches peuvent adapter une règle aux opérandes, mais tous les opérandes sont traités de manière globale :  $\odot_{i=1}^M m_i$ . Avec la SCC, la combinaison devient :

$$\left(\odot_{i=1}^{M_1} m_i\right) \oplus \left(\oplus_{j=1}^{M_2} m_j\right)$$

avec  $M_1$  le nombre de sources classées "normales" ou "défaillantes" et  $M_2$  ne nombre

de sources classées "leurrées". Cette version de la SCC est celle où  $\left(\bigoplus_{i=1}^{M_1} m_i\right)$  n'est pas classée défaillante. Grâce à ce procédé, les propriétés adéquates sont présentes localement pour chaque opérande et globalement cumulées.

Enfin, notons que, de plus,  $\bigoplus$  et  $\bigodot$  sont duales ; la stratégie est donc pseudo-impartiale.

## 4.5 Expérimentations

Nous avons pour but dans cette partie de démontrer l'intérêt de la SCC par rapport aux autres règles de combinaison dans le cadre du suivi d'objets par le FPPC. La forme et le paramétrage du FPPC est identique pour toutes les méthodes de combinaison testées, seule l'étape de fusion diffère, ainsi la comparaison sur une séquence d'image donnée des méthodes de combinaison est équitable (*cf. Nota benne* en 3.5.4 concernant ce point).

L'objectif affiché est également de vérifier que la SCC correspond bien au cumul des avantages que présentaient le FPPB et le FPPC dans la fin du chapitre précédent (*cf. 3.5.4*). La SCC doit apporter une réponse plus globale au problème du suivi d'objets.

La SCC a été justifiée dans le cadre de combinaisons de sources non fiables. Ce type de sources est fréquemment rencontré dans le cadre du suivi d'objets comme nous le montrions déjà dans les expériences de la partie 3.5.4. Nous reprenons ici ces mêmes expériences afin de clairement identifier l'apport de la SCC. Le filtre particulière ayant recours à la SCC a été baptisé, filtre particulière avec fusion crédibiliste contextuelle (FPPCC).

Nous rappelons que les tests sont effectués avec trois sources : une source texture-couleur (matrices de cooccurrences pondérées), une source de forme (cartes de symétrie), une source de mouvement (algorithme de Lefaix [Lefaix 02]). Pour plus de clarté dans les graphes présentés, seuls les résultats concernant les règles conjonctive et disjonctive, la SCC et la fusion bayésienne sont représentés dans un premier temps. Les résultats concernant les autres règles sont discutés dans un second temps.

### 4.5.1 Cas de sources défaillantes

Nous rappelons qu'une source est défaillante quand elle ne fournit plus d'information. A ce titre, la séquence de la "balle de tennis" présentée en 3.5.4 présente une source définitivement défaillante (la source texture) et une source momentanément défaillante (la source mouvement). La source texture défaille après changement d'illumination et n'est plus rétablie par la suite. La source mouvement défaille quand la balle est immobile.

Dans le chapitre précédent, le FPPC associé à la règle conjonctive donnait déjà pleine satisfaction, la SCC doit donc faire aussi bien que cette autre approche. Il convient donc de vérifier que l'étape de classification se passe convenablement. La figure 4.2 présente les résultats de la classification source par source en fonction du temps. La source texture-couleur passe logiquement de la classe "normale" à la classe "défaillante" au niveau de l'image 50 qui correspond au changement d'illumination brutal de la scène. La source de forme est quasiment constamment jugée "normale". Seulement sur quelques images, elle passe au statut de "défaillante" en raison de changements d'illumination locaux et temporaires qui perturbent l'apparente symétrie de la

**Figure 4.2** : Résultats de l'étape de classification des FDM de la SCC. **1** : classe normale, **2** : classe défaillante, **3** : classe leurrée.

balle. Enfin, la source de mouvement est alternativement classée "normale" ou "défaillante" selon que l'objet suivi se déplace suffisamment rapidement ou pas.

Les résultats de la classification sont en adéquation avec les événements décrivant la scène, ce qui prouve l'efficacité de la détection de sources défaillantes. Étant donné que sur cet exemple aucune source n'est classée "leurrée", la SCC est équivalente à la règle conjonctive. Les résultats en termes de suivi sont donc rigoureusement identiques à ceux présentés en figure 3.6 dans le chapitre précédent. Pour comparatif, nous donnons l'évolution du taux de suivi dans la figure 4.3.

**Figure 4.3** : Évaluation des performances en présence de défaillances pour différentes méthodes de fusion.

### 4.5.2 Cas de sources leurrées

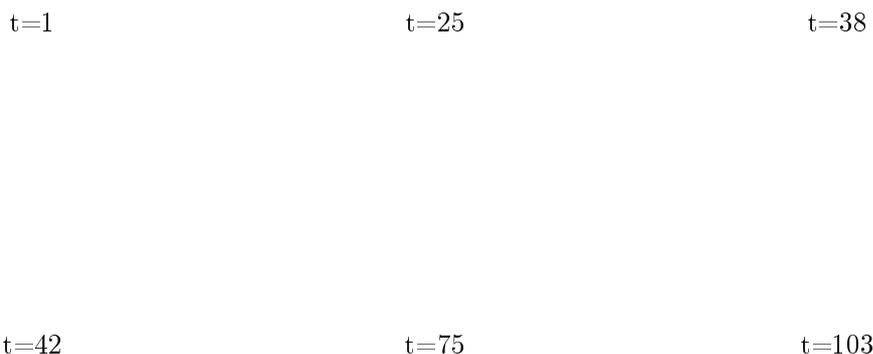
Nous rappelons qu'une source est leurrée quand elle fournit potentiellement une proportion importante d'informations erronées.

**Premier exemple** Nous reprenons ici la vidéo "deux voitures" présentée en 3.5.4. Des sources d'information s'y trouvent leurrées en raison d'une situation de fouillis entre deux véhicules. Sur cette vidéo, ni le FPPC associé à la règle conjonctive, ni le FPPB modifié ne donnait entière satisfaction. Les résultats de l'étape de classification de la SCC sont présentés dans la figure 4.4. En examinant le premier graphe de la figure 4.4, il apparaît que la source texture-couleur alterne

**Figure 4.4 :** Résultats de l'étape de classification des FDM de la SCC. **1** : classe normale, **2** : classe défaillante, **3** : classe leurrée.

entre deux états : "normale" et "leurrée". Les détections en tant que "leurrée" s'intensifient après l'image 28, qui correspond au moment où la seconde voiture entre dans l'ombre portée de la voiture qui est suivi par l'algorithme. La seconde voiture passe alors d'une couleur d'ensemble blanche à une couleur grise, très similaire à celle de la voiture suivie, d'où cette soudaine aggravation du fouillis. Vers la fin de la séquence, le nombre de détections en tant que "source leurrée" baisse de nouveau, car la seconde voiture quitte la scène. Toutefois, la source conserve une part d'information erronée, car le bitume est lui aussi de couleur grise et le véhicule suivi subit des changements d'illumination locaux au niveau de la carrosserie, ce qui au fil du temps rend le modèle texture-couleur moins discriminant.

Concernant la source de forme, le second graphe de la figure 4.4 montre que cette source alterne elle aussi entre les états "normale" et "leurrée", car les deux véhicules ont des formes quasi-identiques. Plus le second véhicule se rapproche, plus le nombre de détections en tant que



**Figure 4.5 :** suivi par FPFCC (3 sources).

"leurrée" augmente. Ce même nombre chute brutalement après que le second véhicule a disparu de la scène (image 60).

Enfin la source de mouvement est successivement détectée "leurrée" ou "défaillante". La vidéo étudiée est particulièrement difficile à traiter pour la source de mouvement, comme nous l'évoquions au chapitre précédent. Les détections en tant que "leurrées" s'expliquent par la présence de nombreux objets-tiers aux mouvements indépendants dans la scène (véhicule constituant le fouillis, véhicules en contre-sens, piétons). Les détections en tant que "défaillante" s'expliquent par un mouvement du véhicule suivi d'intensité équivalente à celui du véhicule où est positionnée la caméra.

Sur cet exemple la SCC va en particulier utiliser à plusieurs reprises les règles disjonctive et conjonctive en même temps. La figure 4.5 montre les résultats obtenus par le FPFCC. Les résultats sont nettement meilleurs que ceux de la figure 3.8 correspondant aux performances du FPFCC associé à la règle conjonctive. Ce constat est renforcé par l'examen des taux de suivi de la figure 4.6. Comme espéré, les graphiques de la figure 4.6 montrent que la SCC procure une fusion plus efficace que toutes les autres méthodes représentées sur la figure. La SCC parvient notamment à maintenir un niveau moyen de suivi satisfaisant en s'appuyant sur la discrimination apportée par la règle disjonctive et à gommer les erreurs de détection du FPFCC modifié grâce à la règle conjonctive. C'est uniquement grâce à l'incorporation de l'information contextuelle apportée par l'étape de classification que la SCC parvient à ce résultat.

**Deuxième exemple** Nous reprenons ici la vidéo "chien et ballon" présentée en 3.5.4. Cette vidéo contient notamment une occultation et de plus les modèles des sources sont peu robustes, ce qui occasionne des anomalies dans les informations fournies par les sources. Dans le chapitre

**Figure 4.6 :** Évaluation des performances en présence d'informations erronées pour différentes méthodes de fusion.

précédent, seule l'approche par FPFB modifié permettait de reprendre le suivi après occultation. Les résultats de l'étape de classification de la SCC sont présentés dans la figure 4.7. La source

**Figure 4.7 :** Résultats de l'étape de classification des FDM de la SCC. **1** : classe normale, **2** : classe défaillante, **3** : classe leurrée.

texture-couleur passe soudainement du statut "normale" à "leurrée" aux alentours de l'image 25, c'est-à-dire au moment où l'occultation atteint un seuil critique. Il serait alors logique que cette source soit détectée en tant que "défaillante", puisque l'objet n'est plus présent dans la

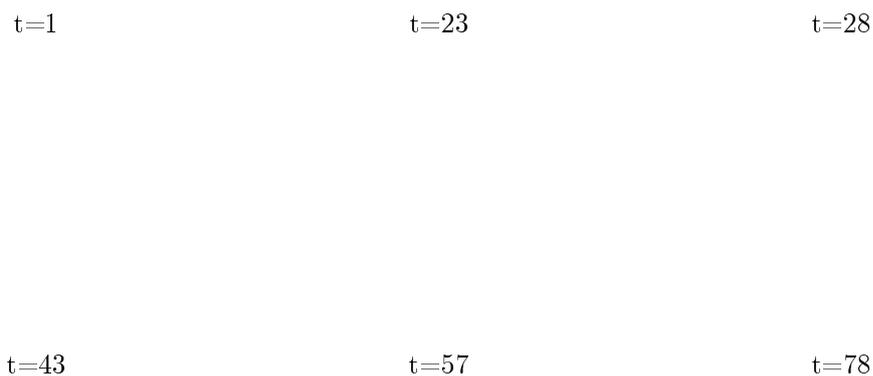
scène. L'objet suivi possède trois couleurs dominantes : le jaune, le noir et le vert. Or, dans cette scène, un objet-tiers (le chien) a pour couleur dominante le noir et le fond de l'image (de l'herbe) a pour couleur dominante le vert. La source texture-couleur assimile partiellement ces éléments au modèle de l'objet-cible et continue donc à donner du crédit malgré la disparition de l'objet-cible. Naturellement le crédit alloué est nettement plus faible que pour le véritable objet ce qui explique, qu'une fois l'objet réapparu dans le giron du filtre particulière, cette source repasse en mode "normal". Notons que pour un groupe d'images situé en fin de séquence, la source repasse de nouveau en "leurrée". Cela s'explique par le bruit et les changements d'apparence du ballon (*cf.* figure 3.10).

La source de forme n'est absolument pas adaptée à l'exemple étudié, puisque dès que le ballon subit une rotation trop importante, la source de forme est incapable de le reconnaître et passe donc de "normale" à "défaillante".

La source de mouvement est quasiment constamment détectée comme "leurrée". En effet, deux objets sont principalement en mouvement dans cette scène : le chien et le ballon. Quand ceux-ci sont en mouvement, ils constituent une forme de fouillis pour la source mouvement qui se trouve classée en "leurrée". Quand ils sont immobiles, la source ne dispose pas d'information et est classée "défaillante". De temps à autres, seul le chien se déplace réduisant ainsi la variance pondérée des particules, la source passe en mode "normal". Dans ce cas, cette classification est logique bien que regrettable, puisque la source fournit ici des informations uniquement erronées. L'architecture de la SCC ne prévoit pas un tel cas d'erreur, mais le modèle des FDM prend lui en compte ce danger potentiel, car la source de mouvement crédite  $\{\omega_1, \omega_3\}$  et non  $\{\omega_1\}$ . Les désagréments d'une telle situation sont donc limités par la cadre des fonctions de croyance.

Comme pour les exemples précédents, le comportement de l'étape de classification est globalement en accord avec la description événementielle des vidéos testées. Notons que la classification n'a pas besoin d'être excessivement précise pour améliorer les performances du système de fusion. Seule une mauvaise classification persistante sur plusieurs images consécutives engendrerait potentiellement une baisse de performance de la SCC. La méthode simple de classification proposée en 4.4.2 est suffisamment robuste pour éviter une telle situation. Cette remarque montre également que la paramétrisation de cette étape revêt une certaine souplesse.

Sur l'exemple de la séquence "chien et ballon", la SCC en s'appuyant sur la règle disjonctive parvient à reprendre le suivi après occultation, comme le montre la figure 4.8. La figure 4.9 confirme les bonnes performances de la SCC sur cette séquence. En s'appuyant également sur la règle conjonctive, la performance moyenne en termes de suivi est améliorée comparé au FPFb modifié, car une erreur de positionnement autour de l'image 70 est évitée, à l'instar de l'exemple précédent. Notons que la SCC sur cet exemple se repose essentiellement sur la règle disjonctive, comme les résultats de la classification l'imposent. Les courbes correspondantes sont pourtant nettement distinctes en raison de la source de forme, qui comme dans la séquence "balle de tennis" annihile les efforts de combinaison en raison de taux de détection quasi-nuls.



**Figure 4.8 :** suivi avec FPFCC (3 sources).

**Figure 4.9 :** Évaluation des performances en présence d'informations erronées pour différentes méthodes de fusion.

### 4.5.3 Comparatif global des règles et de la stratégie de combinaison

Nous proposons à présent les taux de suivi moyens ainsi que les écarts-types correspondants pour d'autres règles de combinaisons. La table 4.3 rassemble ces données pour les trois séquences expérimentales évoquées précédemment. Certaines règles étudiées nécessitent une paramétrisation. C'est notamment le cas des règles de Delmotte qui requièrent la définition de coefficients

**Tableau 4.3** : Moyenne et écart-type de  $r$  pour différentes techniques de combinaison. Les résultats les plus significatifs sont indiqués en gras. \* : méthodes utilisant de l'information contextuelle..

technique de combinaison	séquence de la balle de tennis		séquence des deux voitures		séquence du chien et du ballon	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
conjonctive	<b>0,7803</b>	0,1510	0,6390	0,1330	0,3195	0,3481
disjonctive	0,3376	0,4060	0,6634	0,1391	0,3283	0,3256
conjonctive prudente	<b>0,7756</b>	0,1546	0,6382	0,1345	0,3170	0,3474
Zhang	<b>0,7685</b>	0,1506	0,6433	0,1293	0,3227	0,3455
Yager	<b>0,7803</b>	0,1510	0,6390	0,1330	0,3192	0,3482
Dubois et Prade	<b>0,7803</b>	0,1510	0,6390	0,1330	0,3306	0,3728
WAO	<b>0,7803</b>	0,1510	0,6397	0,1326	0,3306	0,3728
Intermédiaire	<b>0,7781</b>	0,1516	0,7109	0,1335	0,3192	0,3482
adaptative de Florea	<b>0,7803</b>	0,1510	0,6390	0,1330	0,3433	0,3785
mélangeante de Martin	<b>0,7779</b>	0,1483	0,6393	0,1314	0,3345	0,3716
PCR	<b>0,7803</b>	0,1510	0,6390	0,1330	0,3306	0,3728
DPCR	<b>0,7803</b>	0,1510	0,6390	0,1330	0,3311	0,3722
MDPCR	<b>0,7779</b>	0,1483	0,6393	0,1314	0,3305	0,3736
moyenne de Delmotte*	0,7075	0,2347	0,6463	0,1277	0,3454	0,3777
mélangeante de Delmotte*	0,3376	0,4060	0,6384	0,1291	0,3396	0,3755
FPFB modifié*	0,5514	0,3933	<b>0,7390</b>	0,1180	<b>0,5954</b>	0,2818
Stratégie de combinaison contextuelle*	<b>0,7803</b>	0,1510	<b>0,7287</b>	0,0797	<b>0,6043</b>	0,2662

de fiabilité. L'étape de classification de la SCC peut être assimilée à l'estimation "binaire" de coefficients de fiabilité. Afin d'utiliser un niveau équivalent d'information contextuelle dans les différentes méthodes, il est possible d'associer un coefficient fixe à chaque classe de FDM. Suite à plusieurs manipulations, nous avons choisi  $R_i = 0,3$  si la source est leurrée et  $R_i = 1$  dans le cas contraire. Ce choix permet de réduire significativement l'impact des sources leurrées sans pour autant les éliminer totalement puisqu'elles peuvent s'avérer déterminantes par moments. Les sources défaillantes ne sont pas modifiées, car leur importance est déjà minime par nature.

Concernant les règles DPCR et MDPCR, nous avons choisi le paramètre d'affaiblissement  $\epsilon = 0,5$ , afin de répartir le conflit de façon équitable entre les différents termes qui composent ces règles. Cela permet d'accentuer la différence de traitement d'information par rapport aux autres règles.

Sur les deux premières vidéos ("balle de tennis" et "deux voitures"), on constate qu'un nombre important de règles donnent exactement les mêmes résultats. Il s'agit des règles fondées sur la redistribution du conflit, car dans ces deux vidéos aucun conflit n'est généré. Il n'y a pas en effet de modèle défini pour que la source texture-couleur crédite  $\{\omega_2\}$ . Dans la troisième séquence, ce modèle est défini et du conflit apparaît. Toutefois, ce conflit reste de faible amplitude, et les règles concernées donnent des résultats très sensiblement différents. Ces résultats montrent que ces mêmes règles n'offrent pas une réponse adéquate pour notre problème. Toutes ces méthodes reposent en grande partie sur la règle conjonctive, et fournissent donc une réponse équivalente à la règle conjonctive.

**Tableau 4.4 :** Moyenne et écart-type de  $r$  pour différentes techniques de combinaison et avec un affaiblissement contextuel. Les résultats les plus significatifs sont indiqués en gras.  $\star$  : les sources défaillantes sont évincées du processus de fusion..

technique de combinaison	séquence de la balle de tennis		séquence des deux voitures		séquence du chien et du ballon	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
conjonctive	<b>0, 7803</b>	0, 1510	0, 6397	0, 1326	0, 3450	0, 3783
disjonctive	0, 6501 $\star$	0, 3248	0, 6505	0, 1497	0, 5752 $\star$	0, 2904
conjonctive prudente	<b>0, 7756</b>	0, 1546	0, 6389	0, 1335	0, 3437	0, 3789
Zhang	<b>0, 7685</b>	0, 1506	0, 6479	0, 1273	0, 3446	0, 3732
Yager	<b>0, 7803</b>	0, 1510	0, 6397	0, 1326	0, 3452	0, 3786
Dubois et Prade	<b>0, 7803</b>	0, 1510	0, 6397	0, 1326	0, 3350	0, 3721
WAO	<b>0, 7803</b>	0, 1510	0, 6397	0, 1326	0, 3300	0, 3740
Intermédiaire	<b>0, 7781</b>	0, 3692	0, 6453	0, 1298	0, 3450	0, 3778
adaptative de Florea	<b>0, 7803</b>	0, 1510	0, 6397	0, 1326	0, 3450	0, 3783
mélangeante de Martin	<b>0, 7779</b>	0, 1483	0, 6452	0, 1297	0, 3345	0, 3716
PCR	<b>0, 7803</b>	0, 1510	0, 6397	0, 1326	0, 3350	0, 3721
DPCR	<b>0, 7803</b>	0, 1510	0, 6397	0, 1326	0, 3338	0, 3727
MDPCR	<b>0, 7779</b>	0, 1483	0, 6452	0, 1297	0, 3339	0, 3724
moyenne de Delmotte	<b>0, 7758<math>\star</math></b>	0, 1567	0, 6515	0, 1249	0, 3449	0, 3753
mélangeante de Delmotte	<b>0, 7976<math>\star</math></b>	0, 1499	0, 6420	0, 1273	0, 3450	0, 3783

La règle de Zhang reste elle-aussi de nature conjonctive et n'apporte que peu de différences, malgré une meilleure prise en compte des sources indéscises. La règle mixte de Martin offre elle aussi une telle prise en compte tout en étant de nature mixte entre conjonction et disjonction. Pourtant, les performances de cette règle suivent celles de la règle conjonctive. Les performances de la MDPCR sont très proches de celles de la règle mixte de Martin pour les mêmes raisons que celles évoquées dans le paragraphe précédent, à savoir l'absence ou l'influence négligeable du conflit.

Les autres règles que la règle mixte de Martin et dont la nature est intermédiaire en conjonction et disjonction semblent toutes être prioritairement influencées par la règle conjonctive et à ce titre ne parviennent pas non plus à relever les défis soulevés par notre application de suivi d'objets. La règle de la moyenne de Delmotte fait exception car en l'absence de conflit elle est équivalente à la règle disjonctive, ce qui dans la première expérience n'est pas un bon choix.

Pour une comparaison plus équitable, il convient d'insérer dans chacune des méthodes étudiées le même niveau d'information contextuelle, ce qui n'est pas le cas dans les expériences précédentes (*cf.* légende de la table 4.3). Pour atteindre cet objectif, nous proposons d'ajouter un affaiblissement conditionnel, reposant sur l'étape de classification de la SCC. Ainsi, la quantité d'informations traitées par les différentes méthodes est équivalente. La table 4.4 rassemble les résultats obtenus après ajout de cette étape d'affaiblissement. L'affaiblissement contextuel est réalisé de manière analogue à l'obtention de coefficients de fiabilité. Un affaiblissement est opéré sur les FDM qui sont classées "leurrées" avec un coefficient  $\alpha = 0,7$ . De même qu'évoqué précédemment, un tel niveau d'affaiblissement réduit nettement la portée d'une FDM sans

pour autant la supprimer du processus de fusion. Notons qu'affaiblir une source "défaillante" n'a pas de sens, car cela la rendrait encore plus proche de l'ignorance totale  $\Omega^0$ . En appliquant ce même affaiblissement aux règles de Delmotte, cela revient à tester une autre valeur de coefficient  $R_i = 0,09$  pour les sources leurrées.

A l'image de ce qui est fait pour le FPFB, les sources défaillantes peuvent être éliminées pour les règles ayant recours au moins en partie à la règle conjonctive. Cette sélection contextuelle peut prévenir des erreurs telles que celle présentée en figure 4.3 pour la règle disjonctive. L'apport de cette sélection est contrasté. Dans la table 4.4, nous présentons pour chaque méthode concernée le meilleur des résultats obtenus dans les deux cas (avec ou sans cette sélection).

Il n'est pas possible de tester toutes les méthodes pour toutes les valeurs possibles du coefficient  $\alpha$  ou des coefficients  $R_i$ . D'après nos expérimentations, des valeurs proches de celles choisies (à 0,1 près) donnent des performances équivalentes. Les performances sont plus largement modifiées quand les coefficients sont choisis très faibles ou très élevés. S'ils sont choisis très faibles, cela revient à annihiler certaines sources, ce qui n'est pas convenable comme le montre l'expérience sur la séquence "deux voitures", où même des FDM classées défaillantes permettent d'améliorer les performances. Si les coefficients sont choisis très hauts, alors les performances se rapprochent de celles présentées dans la table 4.3, puisque l'affaiblissement devient trop léger.

De manière assez surprenante, les performances des règles semblent globalement diminuées par l'affaiblissement contextuel. En réalité, un affaiblissement engendre d'une certaine manière une perte d'information en injectant de l'indécision au sein des FDM affaiblies. Sans un modèle d'évaluation des coefficients d'affaiblissement plus précis, l'affaiblissement gère la quantité d'information disponible de manière moins efficace que la SCC. L'information contextuelle, qui provient dans notre contexte du filtre particulière, ne permet pas de définir de manière précise l'affaiblissement adapté à telle ou telle situation.

Il en va de même pour les coefficients de fiabilité qui, en règle générale, sont définis dynamiquement dans l'intervalle  $[0, 1]$ . Notons tout de même que la règle mixte de Delmotte réalise la meilleure performance sur la première séquence.

La méthode qui concurrence le mieux le FPFCC reste le FPFB modifié. Elle est aussi l'approche classique de fusion dans les filtres particulières. Nous résumons ici les principales différences entre ces deux approches :

- Quand la SCC est équivalente à la règle disjonctive et qu'il n'y a pas de conflit généré, le principe des opérations de fusion de la fusion bayésienne et de la SCC sont semblables. Le FPFB modifié est donc en quelque sorte contenu dans le modèle du FPFCC.
- La SCC étant définie dans le cadre des fonctions de croyances, des modèles plus complexes de la scène peuvent être utilisés. C'est notamment le cas dans la séquence "chien et ballon", où la texture-couleur de la tête du chien est modélisée, ce qui rend la source texture-couleur plus fiable.
- La SCC nécessite le réglage de deux paramètres contre un seul pour le FPFB modifié, mais le choix des paramètres pour la SCC est moins critique que pour le FPFB modifié.
- Le FPFB modifié utilise un mode de fusion essentiellement disjonctif. Dans la situation

décrite dans l'exemple 2 de la section 3.5.4 et que l'on retrouve présente dans la séquence "deux voitures", le FPFb modifié prend (en fonction de la valeur du seuil de rejet des sources défaillantes) soit une décision trop risquée, soit ne prend pas de décision. Dans cette même situation la SCC en s'appuyant sur un autre mode de combinaison parvient à un résultat intermédiaire en prenant une décision prudente. Il est intéressant de noter par ailleurs que bien que la disjonction soit une attitude prudente de fusion, la décision qui en résulte au niveau du filtre peut être risquée.

La conclusion la plus importante à retenir de l'ensemble des résultats est la suivante : à niveau d'information traité équivalent, la SCC est la seule méthode de fusion qui donne satisfaction sur l'ensemble des trois séquences, en atteignant des taux moyens parmi les plus élevés et avec un taux de suivi instantané jamais inférieur à 0,5 comme le prouvent les figures 4.3, 4.6 et 4.9. A ce titre, l'approche proposée, qui se compose des différents extracteurs d'information, du filtre particulière et de la SCC répond aux objectifs fixés en introduction de ce mémoire, à savoir maintenir un suivi de qualité malgré la présence d'événements perturbateurs.

## 4.6 Conclusion

Dans ce chapitre, une présentation des principales règles de combinaison existantes dans le cadre du formalisme des fonctions de croyance a été effectuée. Elle s'est axée autour de l'étude des différentes propriétés que peuvent revêtir les règles, et du sens qu'ont ces propriétés en termes de fusion de données. Les besoins en propriétés de règle de combinaison ont été énoncés dans le cadre de sources anormales. Dans le contexte du suivi d'objets, les sources anormales correspondent aux situations engendrées par les événements parasites que nous cherchons à combattre. Pour répondre aux besoins énoncés, nous avons alors proposé la stratégie de combinaison contextuelle. Cette stratégie a été comparée aux autres modes de fusion du formalisme des fonctions de croyance afin de valider son intérêt. La comparaison a été établie en termes de qualité de suivi d'objets via l'algorithme FPFc.

Les résultats fournis par le FPFCC améliorent ceux obtenus dans le chapitre précédent. Ceci n'est pas étonnant étant donné que les méthodes de fusion utilisées qui avaient retenu notre attention précédemment sont des sous-cas de la SCC. La SCC permet de sortir de l'impasse que constituaient les conclusions du chapitre 3. Nous avons à présent dessiné les contours d'une approche de suivi destinée à contrer les effets indésirables des événements parasites grâce à un processus de fusion de données élaboré. Chaque étape de l'algorithme de suivi est à présent explicitée et justifiée, il convient à présent de tester cette approche à plus large échelle. Nous proposons l'étude de l'algorithme FPFCC dans le cadre du suivi de la face arrière du véhicule précédent dans le chapitre 5. Ce chapitre applicatif permettra de vérifier la robustesse de l'approche dans diverses situations de circulation automobile.

# Chapitre 5

## Application au suivi du véhicule précédent

### Sommaire

---

<b>5.1</b>	<b>Introduction</b>	<b>161</b>
<b>5.2</b>	<b>Remarques préliminaires</b>	<b>162</b>
5.2.1	Remarques sur les extracteurs	163
5.2.2	Fusion et véhicules intelligents	164
5.2.3	Modifications algorithmique mineures	165
5.2.4	Remarques sur la complexité	165
5.2.5	Résumé du raisonnement scientifique formulé	166
5.2.6	Protocole expérimental	166
<b>5.3</b>	<b>Robustesse du suivi dans le temps</b>	<b>169</b>
<b>5.4</b>	<b>Environnement d'acquisition</b>	<b>171</b>
5.4.1	Conditions climatiques	171
5.4.2	Conditions d'éclairage	176
5.4.3	Milieu routier	180
5.4.4	Géométrie de la route	183
<b>5.5</b>	<b>Influence du nombre de sources</b>	<b>186</b>
<b>5.6</b>	<b>Remarques et bilan des expérimentations</b>	<b>187</b>
<b>5.7</b>	<b>Conclusion</b>	<b>190</b>

---

### 5.1 Introduction

Les chapitres précédents nous ont permis de définir et de justifier une approche (FPFCC) ayant pour but d'apporter une réponse fiable au problème du suivi de la face arrière du véhicule précédent présenté en introduction du mémoire. Dans le présent chapitre, nous mettons en œuvre cette approche spécifiquement dans le cadre du contexte applicatif de ces travaux de thèse. Nous déclinerons différentes situations particulières relatives au suivi de véhicules afin de montrer que le FPFCC fournit un suivi de qualité indépendamment de nombreuses contraintes. Nous mettrons particulièrement l'accent sur les conditions climatiques, le type de route empruntée et les conditions d'éclairage.

Ce chapitre à vocation applicative propose de nombreux tests issus de diverses séquences. Le système d'acquisition des séquences d'image utilisé pour ces tests est peu complexe, il se compose de :

- une caméra CDD couleur Marlin F080C (Allied Vision Technologies, AVT) munie d'un objectif 12mm ;
- un ordinateur portable Packard Bell EasyNote E3247 (processeur AMD Athlon à 2GHz) qui enregistre les informations fournies par la caméra à l'aide du logiciel SmartView 1.8 de chez AVT ;
- un convertisseur électrique alimentant la carte PCMCIA via l'allume-cigare du véhicule (cette carte interfaçant la caméra et l'ordinateur).

La figure 5.1 montre en images le système d'acquisition. Ce système peut être reproduit sans

**Figure 5.1 :** Système d'acquisition embarqué.

grande difficulté et les résultats des algorithmes sont peu dépendants des éléments matériels constituant ce système. Seule la qualité de la caméra constitue un critère ne pouvant être négligé. La qualité de la caméra utilisée est tout à fait standard ; elle a été utilisée avec un réglage automatique du gain, du shutter et de la balance des blancs. Ces procédés permettent de limiter quelque peu l'influence des changements d'illumination. Sans ces procédés, la caméra sature en pratique très rapidement dans les blancs. Enfin, les calculs et les traitements des séquences s'effectuent à l'aide d'un second ordinateur portable Lenovo 3000 N200 (processeur Intel celeron core duo à 1,46 GHz).

## 5.2 Remarques préliminaires

Les chapitres précédents se sont attachés à examiner l'état de l'art et à justifier notre approche principalement sur le plan théorique. Dans ce chapitre, nous apportons une justification plus pratique à nos travaux, et quelques ajustements algorithmiques spécifiques à l'application sont nécessaires. Nous commençons donc ce chapitre par quelques remarques et rappels sur

l'architecture proposée.

### 5.2.1 Remarques sur les extracteurs

La justification des extracteurs adaptés pour le suivi de la face arrière des véhicule découle des conclusions des chapitre 1 et 2 (*cf.* 2.8). Nous avons alors déjà sélectionné certaines méthodes d'extraction intéressantes pour notre problème. Nous affinons ici cette sélection en examinant quelques travaux de références dans le domaine des véhicules intelligents.

Il est difficile de trouver un véritable compromis entre le temps de calcul requis et l'efficacité d'un extracteur. Dans ces travaux de thèse, nous avons privilégié les extracteurs ayant une efficacité reconnue pour le suivi de véhicule afin d'avoir à disposition des sources d'information fournissant un minimum d'information pertinente et discriminante. Toutefois, le critère du temps de calcul a été pris en compte en évinçant des méthodes excessivement coûteuses en temps de calcul. Nous utiliserons donc de 2 à 5 extracteurs :

- extracteurs texture-couleur : cooccurrences couleur pondérées (source  $S_1$ ) et densités de couleur (source  $S_2$ ).
- extracteurs de forme : cartes de symétrie (source  $S_3$ ) et cartes de symétrie des contours (source  $S_4$ ).
- extracteur de mouvement : algorithme de Lefaix (source  $S_5$ ).

Ces cinq méthodes répondent aux critères d'efficacité minimale et de temps de calcul acceptable. Les cooccurrences couleur pondérées ont prouvé leur intérêt et leur applicabilité aux séquences routières en annexe A. De plus, les cooccurrences ont été utilisées avec succès en suivi de véhicule par Handmann *et al.* [Handmann 00]. Une étude de Sun *et al.* [Sun 06] datant de 2006 dresse un état des lieux des extracteurs les plus efficaces pour la détection de véhicules dans des images. Les auteurs préconisent entre autres l'usage des symétries, de la couleur, de l'ombre des véhicules, des contours et des textures. Les cooccurrences couleur pondérées permettent déjà d'exploiter les textures et les couleurs. Les densités de couleurs viennent en appui pour renforcer cet axe d'analyse, mais en focalisent une densité sur la partie intérieure de l'image analyse, il est également possible d'utiliser la présence d'ombres. Les cartes de symétrie captent naturellement les propriétés de symétrie. Elles ont été mises en œuvre notamment par Broggi *et al.* [Broggi 00] et par Hoofmann *et al.* [Hoffmann 04]. Les cartes de symétrie de contours permettent de cumuler les propriétés de contour et de symétrie. Cette idée est d'ailleurs recommandée par Broggi *et al.* qui expliquent que cela permet de mieux résister à d'éventuels changements d'illumination locaux et aussi de mieux distinguer le véhicule d'autres zones fortement symétriques comme une image contenant intégralement du bitume. Enfin nous rajoutons une source de mouvement dont l'exploitation est rendue possible par la dimension temporelle des séquences d'images. Cette source offre une information de nature différente et a été appliquée à ces séquences routières (*cf.* [Lefaix 02]).

Nous utiliserons de façon privilégiée au moins trois sources de natures différentes à l'instar des expériences des chapitres précédents. Toutefois, si pour une raison particulière une source est sans intérêt pour une séquence donnée, nous nous réservons la possibilité de l'exclure des traitements afin de gagner en temps d'exécution, même si une catégorie de source n'est alors plus représentée.

### 5.2.2 Fusion et véhicules intelligents

Le besoin en fusion de données dans les systèmes de transport intelligents n'est pas nouveau. Un rapport de 1998 du groupe thématique Fusion du GdR Isis [Aldon 98] présentait déjà un compte-rendu des projets et publications proposées par différents grands laboratoires français en rapport avec la fusion de données et les véhicules intelligents. Ce rapport établit que la fusion est généralement effectuée sous l'une des formes suivantes :

- au niveau signal : les signaux sont directement couplés pour ne former plus qu'un seul signal à traiter (fusion d'images par exemple).
- au niveau primitive : l'obtention de primitives est conditionnée par l'incorporation d'informations contextuelles à l'aide de capteurs intelligents ou bien des primitives mixtes sont directement générées.
- au niveau décisionnel : c'est le niveau d'abstraction le plus élevé et qui permet de fusionner des informations en provenance de toutes sortes de capteurs, mais nécessite une modélisation du problème de fusion.

Toujours d'après ce rapport, seule la troisième classe de méthodes permet de fusionner des sources hétérogènes, c'est-à-dire une fusion telle que nous l'entendons dans ce mémoire, sans *a priori* sur la nature d'une source. Les deux autres catégories sont trop dépendantes du matériel utilisé pour acquérir les données et donc ne s'adaptent pas à tous types de sources.

De manière plus générale, il est également constatable que dans le domaine des véhicules intelligents de nombreux travaux sont destinés à la localisation d'un véhicule (détermination des coordonnées géométriques du véhicule dans un référentiel) et non pas à un suivi visuel (détermination de la région-image correspondant au véhicule suivi). Pour une application de localisation, les trois grandes théories de l'incertain évoquées au chapitre 3 sont souvent utilisées. Pour une application de détection ou de suivi visuel de véhicules, les auteurs ont principalement recours aux probabilités et quelques fois aux fonctions de croyance. Steux *et al.* [Steux 02] utilisent les probabilités via un réseau bayésien *ad hoc* sur des sources semblables à celles mentionnées dans la sous-section précédente mais aussi sur d'autres sources issues de détections laser. Zhao *et al.* [Zhao 07] ont récemment proposé un exemple d'utilisation des fonctions de croyance pour la détection de véhicules. La fusion est réalisée à l'aide de primitives visuelles extraites d'images issues de caméras fixes mais aussi à l'aide de signaux magnétiques.

Notons qu'en marge de ce discours, d'autres auteurs préfèrent utiliser une stratégie de fusion, c'est-à-dire donc différents extracteurs intervenant séquentiellement à plusieurs étapes algorithmiques comme dans [Toulminet 06]. Le choix de ces méthodes est souvent imposé par un contexte fortement contraint en temps de calcul.

En outre, la quasi-totalité des approches utilisent au moins un capteur non-visuel ou bien la stéréo-vision. A notre connaissance, les fonctions de croyances n'ont pas encore été utilisées dans le cadre de séquences issues de caméras embarquées uniquement sur la base d'attributs visuels et sans apprentissage majeur. A ce titre, ces travaux de thèse se détachent de l'état de l'art également sur le plan applicatif.

### 5.2.3 Modifications algorithmique mineures

Quelques modifications mineures ont été ajoutées au FPFCC afin de faciliter son usage. Ces modifications ont un intérêt essentiellement pratique et n'ont pas vocation à être justifiées vis-à-vis des théories évoquées dans les chapitres précédents.

Pour faciliter la convergence du filtre particulaire, la probabilité pignistique  $BetP$ , qui permet d'évaluer la vraisemblance  $p(Y_t|X_t)$ , est tronquée automatiquement. Cette opération s'effectue de la manière suivante :

$$BetP' = BetP - s \text{ si } BetP > s \quad (5.1)$$

$$BetP' = 0 \text{ sinon} \quad (5.2)$$

avec  $BetP'$  la valeur tronquée de  $BetP$  et  $s$  le seuil. Le seuil  $s$  est donné automatiquement par la formule suivante :

$$s = \frac{1}{\#(\Omega)} \quad (5.3)$$

La valeur de  $s$  correspond en fait à ce que vaut la probabilité pignistique en cas d'ignorance totale (équiprobabilité des hypothèses de  $\Omega$ ). Cette valeur n'apporte aucune information pour le filtrage. De plus, c'est le rapport entre les poids  $w_t^{(i)}$  du filtre qui importe. Les poids non-informatifs étant réduits à 0, le rapport est maximisé et donc le filtre converge plus aisément.

D'autre part, les séquences en entrées des extracteurs subissent un filtrage 2D gaussien de taille 3 à 7 afin d'atténuer certaines aspérités dues à des bruits de nature diverses. Ce filtrage est un pré-traitement classique en traitement d'images et est peu coûteux en temps de calcul.

Enfin pour diminuer le temps de calcul, certaines séquences sont sous-échantillonnées, réduisant ainsi la taille des images à traiter. Globalement, l'algorithme fonctionne bien pour un objet dont la surface pixellaire est supérieure à 25 pixels ( $5 \times 5$  pixels). Au delà de 400 pixels ( $20 \times 20$  pixels), le sous-échantillonnage est particulièrement intéressant car l'objet suivi n'est pas mieux reconnu pour des surfaces supérieures tandis que la plupart des extracteurs voient leur temps de calcul fortement augmenter. Les images fournies par le système d'acquisition sont de taille  $1024 \times 768$ . Elles sont sous-échantillonnées de 1 à 3 octaves.

### 5.2.4 Remarques sur la complexité

Étant donné la succession d'algorithmes que représente notre approche, il est impossible de fournir une complexité globale, car de trop nombreux paramètres interviennent. Nous donnons toutefois quelques éléments fondamentaux de cette complexité :

- La complexité est linéaire avec  $N$  le nombre de particules. Néanmoins, l'algorithme du filtre particulaire est hautement parallélisable, puisque l'évolution de chaque particule est indépendante.
- La plupart des extracteurs dépendent linéairement de la surface pixellaire traitée (*cf.* remarque en fin de sous-section précédente). Les cartes de symétrie et l'algorithme de détection de mouvement de Lefaix sont les plus lents des extracteurs. Pour les cartes de symétrie, l'algorithme ne peut être appliqué globalement à l'image  $I_t$ . Il faut donc calculer une carte par particule. Pour les autres extracteurs il convient de calculer les

attributs pour l'image  $I_t$  entière puis d'analyser les résultats localement pour chaque image relative à la particule  $X_t^{(i)}$ .

- Le temps de calcul imputable à la fusion (SCC ou autres) est négligeable par rapport au temps requis par le reste des algorithmes de l'approche. En pratique, ce sont les extracteurs qui consomment le plus de capacité de calcul.

Il est également difficile de donner une estimation globale du temps de calcul tant celui-ci dépend de nombreux paramètres, mais pour une taille d'objet d'environ  $10 \times 10$  avec 500 particules, le FPFCC possède un temps de calcul par image de l'ordre de la seconde. Ce temps reste au moins 15 fois supérieur à la barrière du temps réel, mais avec une implémentation hardware, une parallélisation et une optimisation du code, cette barrière ne semble pas inaccessible.

### 5.2.5 Résumé du raisonnement scientifique formulé

Au fil des chapitres précédents, plusieurs étages algorithmiques ont été conçus ou justifiés. Nous proposons à l'aide de la figure 5.2 une vue globale de l'architecture de l'approche proposée dans le cadre de ces travaux de thèse.

### 5.2.6 Protocole expérimental

L'architecture proposée possède trois niveaux principaux comme le montre la figure 5.2 : le niveau extraction, le niveau suivi et le niveau fusion. Comme évoqué en 3.5.4, nous rappelons que tous les tests pour une séquence donnée sont effectués à paramétrage égal des algorithmes d'extraction et du filtre particulaire. Nous cherchons à démontrer l'intérêt de la fusion d'informations dans un contexte de suivi d'objets et donc seuls les paramètres des techniques de fusion sont susceptibles d'être modifiés. Pour comparer les techniques de fusion, il est primordial qu'elles aient accès aux mêmes informations, d'où une utilisation éventuelle du contexte si elle revêt un intérêt. Suite à l'étude du chapitre précédent, quatre méthodes de fusion sont conservées pour les tests à suivre :

- la SCC, qui est l'approche proposée et une des contributions principales de ces travaux de thèse,
- la règle conjonctive sur laquelle se repose la SCC, et qui est une règle de combinaison de référence,
- la règle disjonctive avec élimination contextuelle des sources muettes, sur laquelle se repose également la SCC, et dont les résultats des chapitres précédents montrent une nette complémentarité par rapport à la règle conjonctive,
- la fusion bayésienne avec seuillage contextuelle, qui constitue la méthode de référence en terme de fusion explicite dans les filtres particulières.

Pour évaluer le caractère générique de notre méthodes, les paramètres intrinsèques de la SCC ne sont pas changés d'une séquence à l'autre dans ce chapitre. Plus une méthode donne de bons résultats dans un maximum de séquence, moins celle-ci est donc dépendante du paramétrage. Les valeurs de ces paramètres sont :  $t_{def} = 10^{-3}$  et  $a = 2$ . Cette même valeur du paramètre  $t_{def}$  est conservée pour le FPFCC disjonctif. Pour le FPFCC modifié deux valeurs de ce seuil sont testées ( $10^{-3}$  et  $10^{-6}$ ) et le meilleur des deux résultats est présenté. Nous estimons que ces deux valeur singulièrement différentes sont suffisantes pour avoir un aperçu des meilleurs résultats possiblement obtenus par le FPFCC modifié (*cf.* discussion en 3.5.4 sur le réglage de ce paramètre). A ce titre, l'approche bayésienne est avantagée, mais comme le montreront les expériences, non seulement l'approche bayésienne ne donne pas ses meilleurs résultats pour une valeur de  $t_{def}$

**Figure 5.2 :** Architecture proposée pour le suivi de la face arrière du véhicule précédent.

fixe, mais en plus les résultats du FPFCC sont meilleurs sur la quasi totalité des séquences testées.

Néanmoins, les performances globales de l'architecture dépendent de tous les paramètres. Il est donc impossible d'affirmer que si la méthode  $A$  ne fonctionne pas sur une séquence avec un réglage donné, alors il n'existe aucun jeu de paramètres permettant de la faire fonctionner. Dans ce chapitre, les paramètres sont choisis de sorte qu'aucun algorithme sous-jacent ne favorise une méthode par rapport à une autre. De plus, ces mêmes paramètres font en sorte qu'un algorithme ne fonctionne pas grâce à un heureux concours de circonstances, notamment en choisissant un écart-type suffisamment élevé pour la densité de transition.

Le choix du nombre de source  $M$  permet d'accentuer les différences entre les approches de fusion. En effet, comparativement, si une méthode de fusion nécessite moins de sources qu'une autre pour des résultats d'une qualité équivalente, alors cette méthode est meilleure. Les paramètres  $N$  et  $M$  sont également choisis de sorte à limiter le temps de calcul.

Pour des raisons pratiques, les tests présentés ici ne sont pas analysés à l'aide de mesure du taux de suivi, qui nécessite l'obtention d'une vérité-terrain excessivement longue à définir pour l'ensemble des séquences à tester. Une évaluation qualitative des résultats est proposée à la place, permettant d'obtenir une approximation du taux de suivi moyen. Un expert évalue le taux de suivi en classant chaque image en cinq catégories :

- catégorie  $r_1$  :  $r < 0,2$
- catégorie  $r_2$  :  $0,2 \leq r < 0,4$
- catégorie  $r_3$  :  $0,4 \leq r < 0,6$
- catégorie  $r_4$  :  $0,6 \leq r < 0,8$
- catégorie  $r_5$  :  $r \geq 0,8$

Les résultats sont ensuite présentés sous forme d'histogrammes simples et d'histogrammes cumulés. Notons que cette méthode est moins précise que celle du taux de suivi, et que conséquemment de faibles écarts de performance doivent être interprétés avec prudence.

Neuf séquences et quelques extraits de séquences sont étudiées dans ce chapitre. Chaque séquence a été choisie pour tester une situation particulière. Toutefois, les séquences correspondent toutes à plusieurs situations testées dans différentes catégories. Par exemple, la séquence n°1 est utilisée pour tester la pérennité du suivi dans le temps, mais elle a été acquise sur autoroute et pourrait constituer un exemple valable de test sur autoroute. Ainsi, les résultats obtenus sur une séquence peuvent être interprétés à différents niveaux. A cet effet, un bilan des expériences est disponible en 5.6. Notons que, dans de rares cas, d'après l'analyse qualitative que l'on peut faire des résultats, certaines mauvaises performances obtenues ne dépendent pas du critère testé, nous précisons alors la cause principale de ces erreurs.

Nous rappelons également qu'une seule imagerie de l'objet est utilisée pour alimenter les extracteurs en information *a priori*. Si d'autres algorithmes de suivi obtiennent de très bons résultats, notamment en terme de durée du suivi, c'est souvent grâce à une quantité beaucoup plus importante d'informations *a priori*.

Enfin, nous ne discuterons pas à chaque test des raisons amenant telle ou telle source à devenir défaillante ou leurrée. Ce type de commentaires ne sera donné que si le contexte s’y prête afin de gagner en clarté. De même, les séquences sont brièvement présentées, et un descriptif plus complet est donné en 5.6.

### 5.3 Robustesse du suivi dans le temps

Nous entendons par robustesse du suivi dans le temps, la capacité à maintenir le suivi sur une période de temps la plus longue possible. De notre point de vue, il s’agit du critère de qualité le plus important d’un algorithme de suivi. A l’évidence, suivre un même objet dans les mêmes conditions sans aucune perturbation extérieure pendant des heures ne prouve en rien l’efficacité d’un algorithme de suivi. Néanmoins, sur des séquences routières hautement dynamiques où des événements perturbateurs parasitent le suivi à des instants aléatoires, le temps pour lequel un suivi est maintenu est alors primordial. Il est par exemple bien plus difficile de suivre un même véhicule pendant des kilomètres (une seule séquence traitée) que de suivre différents véhicules sur quelques centaines de mètres (plusieurs séquences traitées).

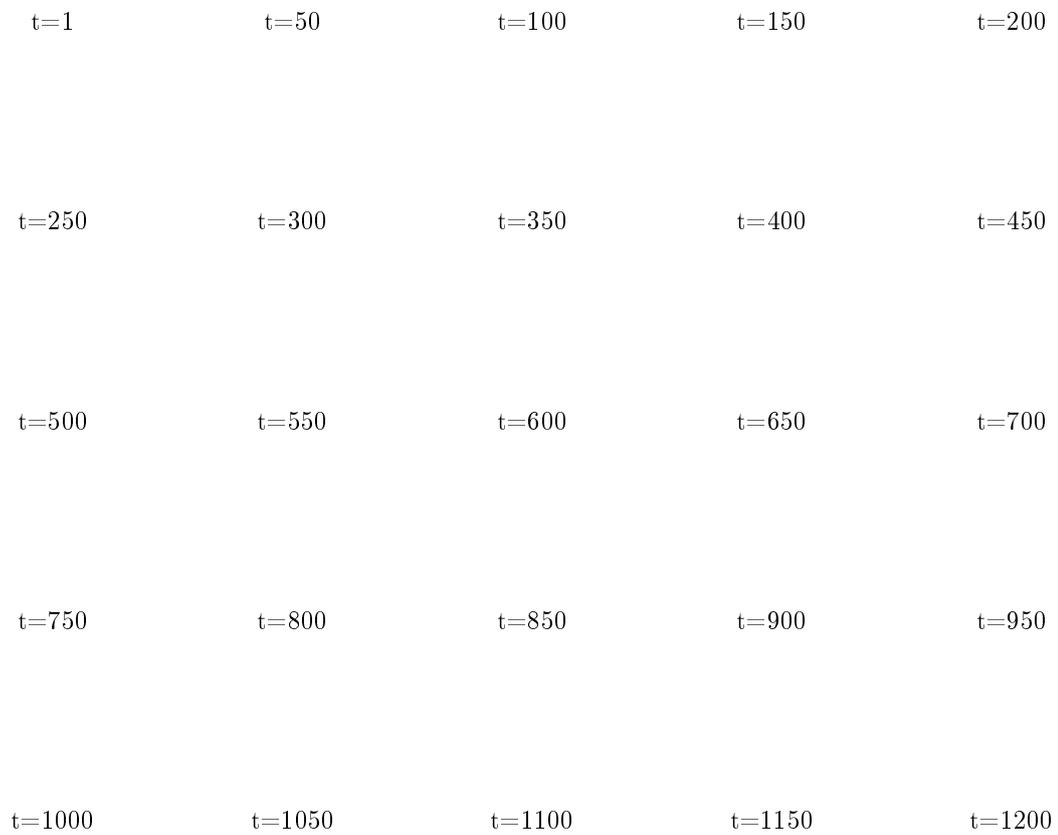
Un phénomène inattendu rend pourtant ces tests délicats à mener. En effet, dans la pratique il est très rare que le véhicule précédent soit le même pendant plus d’un kilomètre. Cette valeur dépend du milieu routier ; sur autoroute elle peut être plus élevée mais en ville bien plus courte. Nous présentons ici une séquence test acquise sur autoroute. Sur autoroute, il est plus facile de suivre un véhicule pendant quelques temps et, de plus, le dynamisme de la séquence est important du fait de la vitesse de déplacement élevée. Cette acquisition a été réalisée à l’aide d’un véhicule complice sur un trajet déterminé à l’avance.

La séquence test n°1 correspond à un parcours d’environ 5,6km à une vitesse moyenne de 120km/h. Elle compte 1250 images équivalent à 2mn 46s<sup>1</sup>. Étant donné le fort ensoleillement de cette scène, le véhicule subit des changements d’illumination très importants, rendant les sources de type texture-couleur quasiment inopérantes pendant plusieurs parties de la séquence. La couleur même du véhicule est proche d’un ton gris-bleu peu éloigné de la couleur du bitume, d’où un modèle texture-couleur de toutes façons peu précis et sujet à de nombreuses situations de leurre. Comme dans toutes séquences routières, les sources de type forme ou mouvement sont elles-aussi sujettes aux situations de leurre quand un véhicule-tiers s’approche trop près du véhicule suivi et entre dans le champ d’action du filtre particulière.

La figure 5.3 montre la qualité de suivi obtenu par le FPFCC sur cette séquence. Le caractère hautement dynamique de la séquence test n°1 est visible à travers les images qui composent la figure 5.3. Selon la position du véhicule par rapport au soleil, celui-ci passe du gris-bleu au noir puis quasiment au blanc. Le fond de la scène évolue rapidement de même que les conditions de trafic. Un fouillis est observable sur l’image correspondant à  $t = 600$  de la figure 5.3. La figure 5.4 présente les statistiques des performances pour les quatre méthodes étudiées dans ce chapitre. Comme cela était prévisible, les performances du FPFCC disjonctif et du FPFB modifié

---

<sup>1</sup>Les fréquences d’acquisition ne sont pas systématiquement les mêmes pour les différentes séquences présentées (cf. table 5.1).



**Figure 5.3** : suivi par FPFCC. 5 sources :  $S_1$  à  $S_5$  (cf. 5.2.4), séquence n°1.

**Figure 5.4 :** Performance moyenne pour la séquence test n°1 selon quatre approches de fusion différentes.  $t_{def} = 10^{-6}$  pour le FPFB modifié. A gauche, histogrammes pour chaque approche, la couleur correspond à la catégorie de qualité de suivi, de  $r_1$  en bleu foncé à  $r_5$  en vert. A droite, mêmes données présentées sous forme d’histogrammes cumulés rassemblés sur un même graphe.

sont très proches. Ces deux algorithmes ont des statistiques particulièrement mauvaises car ils perdent trace de l’objet à hauteur de l’image 600 et cumulent donc de très nombreuses images en catégorie  $r_1$ . Cette perte de suivi est due à une situation de fouillis autour de l’image 600 entre un camion et le véhicule suivi. A cet instant, l’ensoleillement est tel que les sources de type texture-couleur sont totalement inopérantes. Les autres sources étant trop imprécises, un fouillis apparaît. Les deux approches FPFCC disjonctif et FPFB modifié font alors un pari risqué en mettant tout le poids de la décision sur des valeurs faibles de croyances et se trompent d’objet. Le FPFCC conjonctif est insensible à cette situation. Le FPFCC donne un comportement intermédiaire en se focalisant sur le camion pendant quelques images avant de revenir sur le véhicule-suivi.

En revanche, le FPFCC perd momentanément l’objet à deux reprises dans cette séquence. Celui-ci manque en effet de réactivité quand les modèles sont peu discriminant et que l’objet se déplace plus rapidement dans l’image (changement de file). Le FPFCC donne de loin le meilleur résultat sur cette séquence particulièrement difficile à traiter. Les résultats obtenus sur cette séquence corroborent les conclusions des expérimentations du chapitre précédent au sens où l’attitude intermédiaire du FPFCC permet d’augmenter nettement les performances en terme de longévité du suivi.

En adoptant un seuil  $t_{def}$  plus élevé pour le FPFB, les performances de ce dernier sont encore plus mauvaises car le suivi est perdu en amont de la situation de fouillis. Avec un seuil plus faible, la décision de suivre le mauvais objet est accentuée.

## 5.4 Environnement d’acquisition

### 5.4.1 Conditions climatiques

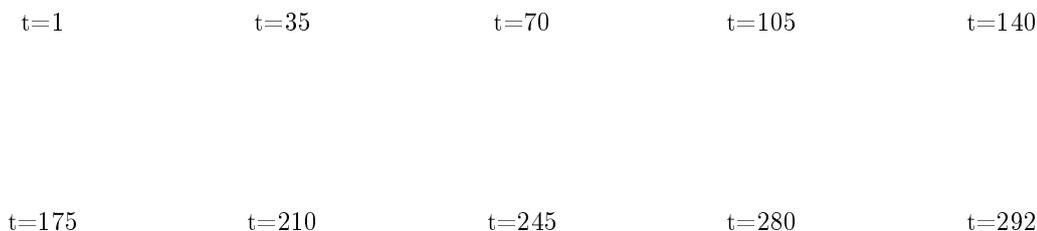
Les conditions climatiques sont l’un des principaux facteurs de disparité des séquences routières. Elles peuvent notamment engendrer des changements d’illumination globaux ou locaux et ainsi perturber de nombreuses sources. Nous déclinons ici ces conditions en trois grandes

catégories : un temps ensoleillé, un temps gris et un temps pluvieux. Ces catégories regroupent l'essentiel des conditions climatiques rencontrées en Europe du nord. Nous n'abordons pas le temps neigeux ou le brouillard.

### Temps ensoleillé

Contrairement à ce qu'on pourrait penser de prime abord, un temps ensoleillé n'est pas l'idéal pour appliquer un algorithme de suivi sur des véhicules. En effet, comme dans la séquence "deux voitures" étudiée dans les chapitres précédents, un fort ensoleillement a tendance à modifier l'apparence visuelle d'un véhicule à cause de la lumière réfléchiée par la carrosserie du véhicule. Cela est d'autant plus vrai pour des véhicules dotés de peintures métallisées. Un fort ensoleillement n'affecte pas seulement l'objet suivi mais aussi le capteur qui peut être plus ou moins saturé selon son orientation par rapport au soleil. Le phénomène s'amplifie quand le soleil est rasant.

Nous présentons ici la séquence test n°2 de 292 images où un véhicule à l'origine de couleur bordeaux apparaît presque noir à l'image à cause de l'ensoleillement. La vidéo a été acquise sur autoroute et la vitesse de déplacement est donc relativement élevée (autour de 100 km/h de moyenne). La figure 5.5 présente les résultats du FPFCC pour cinq sources d'information. D'après cette figure, le FPFCC produit un suivi satisfaisant sur la quasi-totalité de la séquence



**Figure 5.5** : suivi par FPFCC. 5 sources :  $S_1$  à  $S_5$  (cf. 5.2.4), séquence n°2.

malgré cet ensoleillement pénalisant. Cette performance est confirmée par la figure 5.6 qui résume également les performances relatives aux autres méthodes de suivi testées dans ce chapitre. Sur cette séquence, l'approche bayésienne est légèrement meilleure si le seuil de détection des sources défaillantes est choisi beaucoup plus faible que celui du FPFCC. Le risque pris en se reposant sur les petites valeurs de probabilité est ici payant, notamment parce qu'il n'y a pas de situation de fouillis trop forte. Avec le même seuil que FPFCC, le FPFB modifié perd le suivi du véhicule à l'instar du FPFC avec la règle disjonctive. Le FPFC avec la règle conjonctive est légèrement plus instable que le FPFCC. Les quelques erreurs commises par le FPFCC sont dues au fait que les trop faibles valeurs fournies par les sources imposent l'utilisation de la règle conjonctive et donc l'approche manque de réactivité sur quelques images.

Comme nous l'évoquions à propos de la séquence n°1 présentée dans la section précédente,

**Figure 5.6 :** Performance moyenne pour la séquence test n°2 selon quatre approches de fusion différentes.  $t_{def} = 10^{-6}$  pour le FPFCC modifié. À gauche histogrammes simples, à droite histogrammes cumulés.

la séquence n°2 contient également de très importants changements d'illumination dus à un fort ensoleillement. Elle constitue donc aussi une séquence permettant de tester la résistance des approches à l'ensoleillement. La figure 5.7 montre l'étendue des changements engendrés par l'ensoleillement sur l'apparence du véhicule suivi. En dépit de ces difficultés, les résultats de la

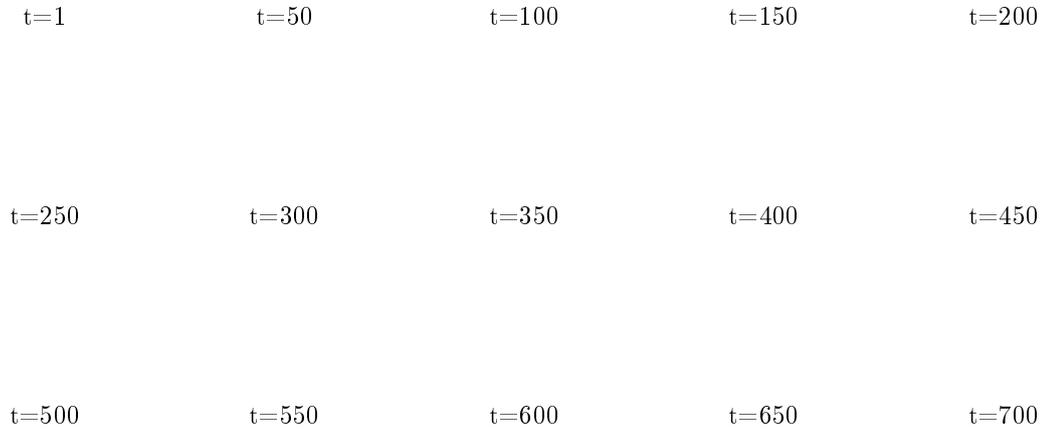
**Figure 5.7 :** Variations d'apparence du véhicule suivi dans la séquence test n°1.

section 5.3 montrent que le FPFCC résiste convenablement aux conséquences d'un ensoleillement important.

### Temps gris

Un temps gris constitue sans doute le type de conditions climatiques le plus favorable au suivi d'objets. L'illumination varie peu sous ces conditions, et ce quelle que soit l'orientation de la caméra par rapport au soleil. Ainsi, les propriétés visuelles de l'objet sont plus persistantes au long de la séquence et les modèles plus discriminants.

Nous étudions ici une séquence acquise par temps gris sur autoroute. La séquence test n°3 comporte 721 images correspondant environ à 1mn 33s. La vitesse de déplacement est d'environ 110km/h sur une distance d'environ 2,8km. La figure 5.8 donne un aperçu visuel des performances du FPFCC sur cette séquence. À l'exception d'une image parmi celles présentées dans la figure 5.8, le FPFCC produit un suivi de qualité très satisfaisante. L'image pour laquelle l'estimation de la position est mauvaise correspond à un moment où le FPFCC utilise la règle conjonctive alors que l'objet se déplace assez rapidement dans l'image. La règle conjonctive peine alors à ré-



**Figure 5.8** : suivi par FPFCC, 3 sources :  $S_1$ ,  $S_2$  et  $S_4$  (cf. 5.2.4), séquence n°3.

agir suffisamment vite, mais l'erreur est corrigée à peine quelques images plus tard. Sur tout le reste de la séquence, le FPFCC ne commet aucune erreur comme en atteste la figure 5.9 présentant les performances moyennes des quatre méthodes testées dans ce chapitre. D'après cette

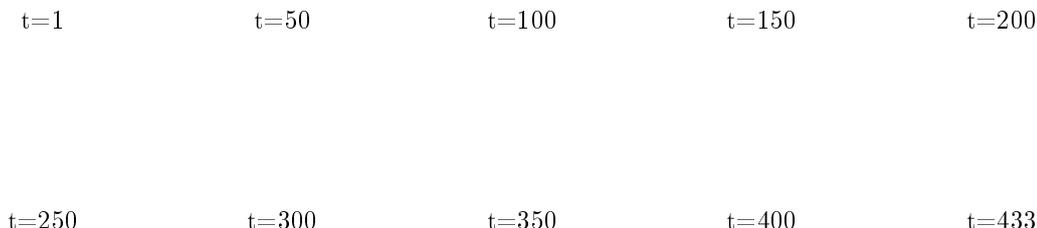
**Figure 5.9** : Performance moyenne pour la séquence test n°3 selon quatre approches de fusion différentes.  $t_{def} = 10^{-6}$  pour le FPFB modifié. A gauche histogrammes simples, à droite histogrammes cumulés.

figure, les performances du FPFB modifié et du FPFCC sont très proches à ceci près qu'aucune image ne tombe dans la catégorie  $r_1$  pour le FPFB modifié. Globalement, les performances par temps gris de ces deux algorithmes donnent satisfaction et les différences sont en moyenne trop faibles pour conclure sur la supériorité d'une approche par rapport à l'autre.

## Temps pluvieux

Sans surprise, la pluie est un événement climatique qui va fortement perturber le bon déroulement des algorithmes de suivi. Premièrement, la présence de gouttes d'eau sur le pare-brise dénature fortement la forme de l'objet perturbant ainsi les sources d'information de ce type. Deuxièmement, le battement des essuie-glaces obstrue naturellement l'objectif de la caméra, même si nous avons constaté qu'il est extrêmement rare que l'essuie-glace soit exactement dans le champ de la caméra lorsqu'une image est acquise. En revanche, le passage de l'essuie-glace laisse une empreinte sombre sur certaines images engendrant un semblant de changement d'illumination. Enfin, les projections d'eau soulevées par le véhicule précédent sont très pénalisantes à haute vitesse car elles peuvent réduire l'objet suivi à une simple tache dans l'image, constituant quasiment une occultation. Toutes les sources se trouvent alors affectées par ce phénomène. Les équimentiers pneumatiques s'intéressent depuis longtemps à ce problème appelé "spray". Notons qu'à cause de ces projections, le fait de placer la caméra à l'extérieur de l'habitacle présente peu d'intérêt.

La séquence test n°4 est adaptée à l'évaluation des algorithmes de suivi par temps de pluie. Elle a été acquise sur autoroute sur un parcours d'environ 2km. Elle comprend 433 images correspondant environ à 1mn 56s. La figure 5.10 illustre les performances du FPFCC sur cette séquence et permet d'apprécier la qualité des images à traiter dans des conditions de pluie. Il



**Figure 5.10** : suivi par FPFCC. 4 sources :  $S_1$  à  $S_4$  (cf. 5.2.4), séquence n°4.

apparaît au travers de cette figure que le dynamisme de la séquence est beaucoup plus élevé dans la première partie car en effet dans la seconde partie un ralentissement dû à la densité du trafic réduit très fortement le déplacement du véhicule. La vitesse de déplacement varie donc énormément dans cette séquence (de 110 à 5km/h). On retrouve également sur cette figure l'ensemble des problèmes décrits en amont (gouttes d'eau, ombres dues aux battements d'essuie-glace et spray).

Les performances du FPFCC sont passables sur cette séquence. Ce sont d'ailleurs là les plus mauvais résultats du FPFCC sur l'ensemble des séquences présentées dans ce mémoire. La figure 5.11 montre les performances moyennes obtenues par les quatre algorithmes sur la séquence n°4.

Il apparaît que le nombre d'images classées en catégorie  $r_1$  est beaucoup plus important

**Figure 5.11 :** Performance moyenne pour la séquence test n°4 selon quatre approches de fusion différentes.  $t_{def} = 10^{-6}$  pour le FPFBB modifié. À gauche histogrammes simples, à droite histogrammes cumulés.

que précédemment quelle que soit l'approche utilisée. Malgré tout, les performances ne sont pas catastrophiques car les pertes de suivi sont momentanées. Du point de vue des bonnes détections, c'est aussi le FPFCC qui donne les meilleurs résultats, mais la différence avec ceux du FPFBB modifié est en moyenne trop faible pour véritablement conclure sur cet aspect. Au regard des erreurs de positionnement du véhicule, les différences sont plus flagrantes et le FPFCC donne des résultats sensiblement meilleurs que le FPFBB modifié. Concernant le FPFCC disjonctif, celui-ci reste un ton en dessous du FPFCC tandis que le FPFCC conjonctif donne des résultats d'une qualité très nettement inférieure.

D'après cette expérience, un temps pluvieux constitue un cas limite pour l'architecture proposée. La qualité du suivi distillé par le FPFCC est dégradée même si notre approche résiste mieux que les approches concurrentes face à ces conditions.

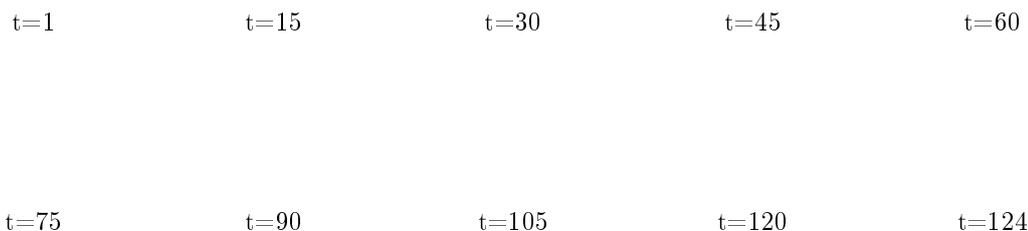
#### 5.4.2 Conditions d'éclairage

De même que les conditions climatiques, les conditions d'éclairage influent sur la qualité du suivi en perturbant l'apparence du véhicule suivi. Comme évoqué dans le chapitre 1, la lumière est un maillon essentiel de la chaîne d'acquisition en imagerie. Dans l'énorme majorité des séquences, l'éclairage est apporté naturellement par le soleil qui fournit une lumière blanche est diffuse. Dans de plus rares cas, l'éclairage n'est pas naturel et le spectre lumineux est réduit. L'éclairage peut être aussi plus ou moins directionnel comme celui des phares d'un véhicule. Nous présentons dans cette section quelques situations particulières où l'éclairage joue un rôle prépondérant par rapport à la qualité du suivi obtenu.

#### Ponts et tunnels

Nous entamons cette section par une séquence correspondant à la traversée d'un tunnel long d'environ 700m. La séquence n°5 comprend 124 images correspondant environ à 33s. L'entrée dans le tunnel intervient à hauteur de l'image 20. Le changement d'illumination subséquent est alors très brutal et rappelle à ce titre la séquence "balle de tennis" étudiée dans les chapitres précédents.

La vitesse moyenne est d'environ 70km/h. La figure 5.12 montre comment se comporte le FPFCC face à cette situation particulière.



**Figure 5.12** : suivi par FPFCC. 3 sources :  $S_1$ ,  $S_2$  et  $S_4$  (cf. 5.2.4), séquence n°5.

D'après cette figure, il apparaît que le FPFCC résiste bien au défi que représente cette séquence. Nous devons préciser toutefois que l'unique image du véhicule suivi utilisée pour nos modèle de FDM est prise à  $t = 25$  soit quelques secondes après l'entrée dans le tunnel. A ce stade, le capteur ne s'est pas encore acclimaté au nouvel environnement et cette image revêt une teinte globale rougeâtre singulièrement différente de la teinte verdâtre observable sur les images intérieures au tunnel de la figure 5.12. Cette image est donc en terme d'illumination éloignée à la fois des images intérieures ou extérieures au tunnel.

Comme prévu, la source texture-couleur est inopérante dans une telle situation, malgré le fait que le véhicule soit de couleur noire. La source de symétrie des contours fonctionne mal sur cette séquences car les contours ne sont pas assez francs dans la majorité des images. Les sources de symétrie est la plus fiable, même si comme à l'accoutumée elle reste imprécise par moments. Quant à la source de mouvement, elle n'apporte pas d'information suffisamment pertinente sur cette séquence et est écartée du processus. A cet égard, cette séquence est particulièrement difficile à traiter et les résultats obtenus peuvent être qualifiés de satisfaisants.

Nous précisons néanmoins que les paramètres des algorithmes d'extraction et des modèles de FDM ont été difficiles à régler pour obtenir le résultat présenté. Plus exactement, quelques changements assez minimes de certains paramètres engendrent une perte de suivi environ à mi-parcours du tunnel. Les performances moyennes du FPFCC et des autres approches testées sont regroupées dans la figure 5.13.

Les résultats relatifs à cette figure donnent un net avantage à l'approche FPFCC par rapport à ces concurrentes. Indépendamment des deux valeurs testées pour le seuil du FPFB modifié, l'approche bayésienne perd le suivi au cours de la séquence, d'où ces mauvaises statistiques au final. La perte de l'objet intervient encore plus tôt dans la séquence pour le FPFCC conjonctif ou disjonctif. Il n'est pas impossible que les conclusions soient différentes pour des réglages différents des modèles et des extracteurs, toutefois dans le cadre des expérimentations menées avec le FPFB modifié ne semble pas en mesure d'assurer la pérennité du suivi sur cette séquence.

Si le FPFCC surmonte avec succès cette épreuve, des modifications mériteraient d'être

**Figure 5.13 :** Performance moyenne pour la séquence test n°5 selon quatre approches de fusion différentes.  $t_{def} = 10^{-6}$  pour le FPFCC modifié. A gauche histogrammes simples, à droite histogrammes cumulés.

apportées en vue de rendre l'ensemble de l'architecture plus robuste. Il est assez aisé de détecter un changement d'illumination aussi brutal en calculant par exemple quelques statistiques simples sur le niveau moyen d'illumination ou en exploitant des différences d'images consécutives. Une base de données d'images de l'objet plus vaste pourrait alors être utilisée en fonction du contexte. Les performances des extracteurs seraient grandement améliorées, et les réglages des modèles de FDM seraient plus souples.

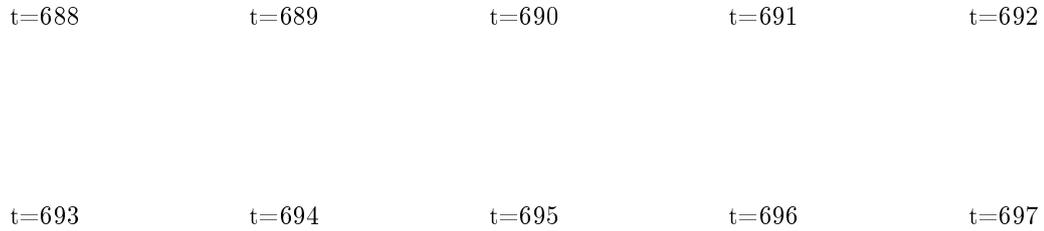
### Ombres portées

Les ombres portées d'objets jouxtant la scène perturbent l'éclairage d'un objet et donc son rendu visuel. Les changements d'illumination imputables aux ombres portées sont généralement plus localisés et plus modérés que ceux relatifs à une entrée dans un tunnel. De plus, la traversée d'une ombre portée est souvent de courte durée et l'inertie du filtre particulière aide alors à surmonter la difficulté. Cet événement étant plus anecdotique, nous ne présentons pas de séquence complète pour son étude mais simplement un extrait de la séquence n°1.

La figure 5.14 montre que le FPFCC résiste convenablement au phénomène d'ombres portées. Les images sur cette séquence sont consécutives afin de mieux apprécier l'étendue et la nature des modifications de l'éclairage. Les performances moyennées sur 10 images ne présentent quasiment aucun intérêt statistique, c'est pourquoi nous ne présentons pas non plus les histogrammes habituels. De plus, les autres approches fonctionnent toutes relativement bien en présence d'ombres portées, aucune ne se détache donc par rapport aux autres vis-à-vis de ce critère.

### Acquisitions de nuit

Les séquences acquises de nuit sont singulièrement différentes des séquences acquises de jour étant donné que la quantité d'information disponible est à l'évidence beaucoup plus restreinte. L'architecture proposée n'est pas adaptée à de telles séquences au sens où aucune disposition n'a été prise pour prendre en compte les spécificités des acquisitions de nuit. De plus, la plupart des hypothèses et des choix méthodologiques formulés tout au long du mémoire prévoyaient implici-

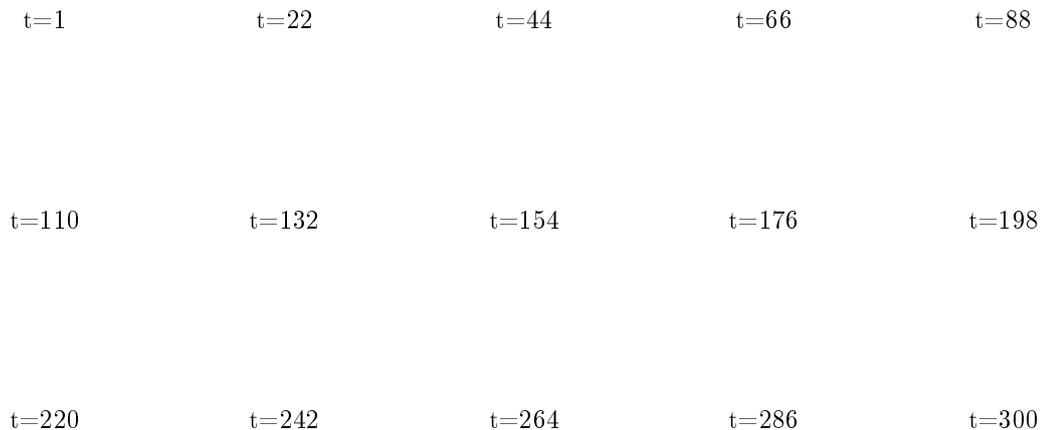


**Figure 5.14** : suivi par FPFCC. 5 sources :  $S_1$  à  $S_5$  (cf. 5.2.4).

tement un éclairage de jour. Toutefois, il est intéressant de tester l'approche dans un tel contexte plus difficile afin de voir si nos conclusions restent pertinentes.

La séquence n°6 a été acquise sur voie rapide à une vitesse moyenne de 90km/h et sur une distance d'environ 2,3km. La source de lumière principale de la scène est constituée par les phares du véhicule où est disposée la caméra, mais l'éclairage public intervient également dans la deuxième partie de la séquence. Cette séquence comporte 300 images équivalent à 1mn 20s. L'objet suivi un poids lourd.

La figure 5.15 donne les résultats obtenus par le FPFCC sur la séquence n°6. D'après cette



**Figure 5.15** : suivi par FPFCC. 4 sources :  $S_1$  à  $S_4$  (cf. 5.2.4), séquence n°6.

figure, les résultats délivrés par le FPFCC sont satisfaisants malgré une imprécision du positionnement de la boîte englobante un peu plus marquée sur la coordonnée verticale. Cette imprécision accrue provient du fait que le haut du poids lourd est moins bien éclairé et se distingue consé-

quement fort mal du fond de l'image. Pour les mêmes raisons, la source texture-couleur est assez peu précise. Sur cette séquence, le modèle de FDM du bitume pour la source  $S_1$  aide à mieux positionner la boîte englobante. La meilleure source est constituée par les cartes de symétrie, tandis que la source de mouvement est quasiment sans effet étant donné que l'algorithme de Lefaix n'a pas suffisamment de points à mettre en correspondance pour estimer le mouvement global. La source de mouvement n'est donc pas utilisée sur cette séquence.

la figure 5.16 présente les performances moyennes obtenues sur cette séquence. Il apparaît

**Figure 5.16 :** Performance moyenne pour la séquence test n°6 selon quatre approches de fusion différentes.  $t_{def} = 10^{-3}$  pour le FPFb modifié. À gauche histogrammes simples, à droite histogrammes cumulés.

que dans ce cas limite, le FPFCC donne des résultats nettement meilleurs que les autres approches car aucune image ne se classe dans les catégories  $r_1$  ou  $r_2$ . Concernant le FPFb modifié, il s'agit du plus mauvais résultat obtenu sur l'ensemble des séquences. Le suivi de l'objet est perdu assez rapidement par le FPFb modifié et encore plus tôt en optant pour un seuil plus faible. Le comportement conjonctif est ici bénéfique et le FPFCC sait en tirer profit.

Le cas particulier des acquisitions de nuit est donc bien toléré par le FPFCC. Néanmoins pour plus de facilité, il convient de prévoir quelques éléments supplémentaires au sein de l'architecture spécifiques à ce type de séquences notamment au niveau des extracteurs. Comme pour le cas du franchissement d'un tunnel, l'illumination globale de la scène permet de distinguer aisément des acquisitions de jour de celles de nuit. Cette information contextuelle peut être alors exploitée au sein des extracteurs.

### 5.4.3 Milieu routier

Nous entendons par milieu routier la catégorie de route sur lesquelles évolue le véhicule suivi. Nous distinguons trois milieux singulièrement différents : le milieu urbain, le milieu autoroutier et le milieu rural. Selon ces environnements routiers, différentes difficultés se présentent pour le suivi d'un véhicule notamment au niveau :

- de la densité du trafic (plus il y a de véhicules, plus le risque de fouillis est grand),
- de la vitesse de déplacement (plus la vitesse est élevée plus la scène est amenée à changer rapidement d'aspect),

- du mouvement du véhicule précédent (en ville, les trajectoires sont plus rectilignes et le véhicule suivi est plus proche de la caméra),
- de l'encombrement du fond de la scène (en ville, le fond est plus densément parsemé d'objets divers pouvant rendre les extracteurs moins discriminants).

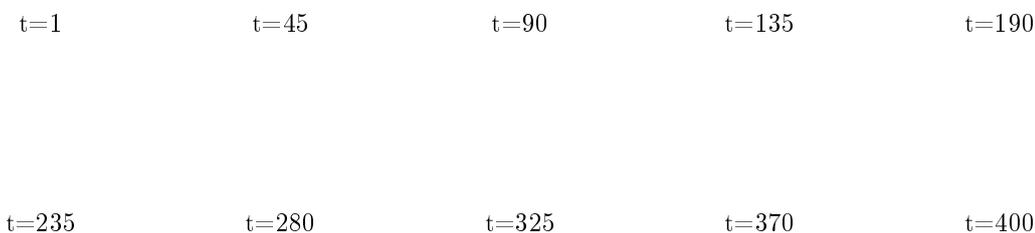
Nous présentons dans cette sous-section quelques séquences représentatives de ces différents cas de figure.

### Trafic urbain

Le trafic urbain se caractérise par une forte densité de trafic mais avec une vitesse de déplacement faible en moyenne et des trajectoires simples. De plus, le fond de l'image est densément peuplé d'objets divers (piétons, vélos, immeubles, véhicules stationnés, signalisation). Les extracteurs s'y trouvent donc moins précis que dans les autres environnements. Étant donné que les déplacements entre deux images sont moins intenses, il est possible de penser que le suivi sera malgré tout plus aisé à obtenir. En réalité, ces séquences restent délicates à traiter à cause notamment du phénomène de "stop-and-go"<sup>2</sup>. Les constructeurs automobiles cherchent à exploiter ce phénomène bien connu afin de limiter la consommation en carburant des véhicules (extinction du moteur en cas d'arrêt, récupération de l'énergie du freinage). En termes de suivi, ce phénomène se traduit par la présence répétée de changements d'échelles de l'objet, qui constitue un des événements perturbateurs que nous évoquons en 2.7.

La séquence n°7 possède toutes les propriétés requises pour l'évaluation des différentes approches en milieu urbain. Elle a été acquise en milieu urbain en dehors des horaires de pointe de trafic afin que la séquence soit un minimum dynamique. La vitesse de déplacement du véhicule varie de 0 à 50km/h. La séquence comprend 400 images correspondant à environ 1mn 47s.

Sur la figure 5.17 sont représentées les estimations de la position du véhicule suivi fournies par le FPFCC. D'après cette figure, la qualité du suivi délivrée par le FPFCC est satisfaisante.



**Figure 5.17** : suivi par FPFCC. 2 sources :  $S_2$  et  $S_3$  (cf. 5.2.4), séquence n°7.

Le phénomène de stop-and-go que nous évoquons est également observable au travers de cette figure. La figure 5.18 confirme les bonnes performances du FPFCC en donnant les performances moyennes de notre approche ainsi que celles des autres approches testées.

<sup>2</sup>Arrêts et redémarrages réguliers et répétés.

**Figure 5.18 :** Performance moyenne pour la séquence test n°7 selon quatre approches de fusion différentes.  $t_{def} = 10^{-6}$  pour le FPFB modifié. À gauche histogrammes simples, à droite histogrammes cumulés.

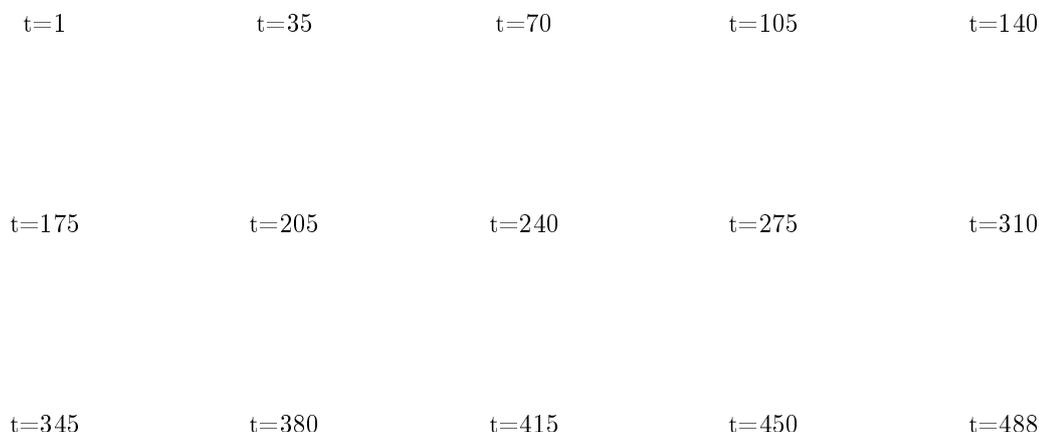
Sur cette séquence, seule le FPFCC conjonctif donne de mauvais résultat. Le changement d'échelle nécessite une adaptation rapide de l'algorithme que ne peut fournir cette approche. Concernant les trois autres approches, elles sont toutes dans des niveaux de performances équivalents. À cause des changements d'échelle, un écart-type plus élevé a été octroyé aux composantes 3 et 4 de  $p(X_t|X_{t-1})$ . Ces composantes sont responsables de la variabilité des dimensions des boîtes englobantes  $X_t^{(i)}$  échantillonnées par le filtre particulaires. Malheureusement cette augmentation nécessite également l'élévation du nombre de particules  $N$  et donc du coût calculatoire de l'ensemble de la procédure.

### Trafic sur voies rapides

Le trafic sur voie rapide se caractérise par une vitesse élevée de déplacement, et donc des séquences hautement dynamiques. Le trafic y est cependant généralement plus fluide qu'en ville et le fond de la scène plus épuré.

La séquence n°8 est utilisée pour tester notre approche sur autoroute. La vitesse moyenne de déplacement est de l'ordre de 110km/h. La séquence compte 488 images équivalent à 1mn 5s. La distance parcourue est d'environ 2km. La figure 5.19 montre les résultats obtenus à l'aide du FPFCC sur cette séquence.

Au regard de cette figure, de bons résultats sont fournis par l'algorithme FPFCC. En  $t = 205$ , seule la partie inférieure du véhicule est détectée en raison d'un changement d'illumination dû à une ombre portée et au fort ensoleillement. Les bonnes performances du FPFCC sont corroborées par la figure 5.20 qui donne les histogrammes des performances moyennes des quatre approches testées. La figure 5.20 montre une qualité de suivi nettement supérieure de la part du FPFCC par rapport à toutes les autres approches, et ce en termes d'erreur de positionnement et de qualité de positionnement. Toutefois, d'après notre analyse, ces différences ne s'expliquent pas par le milieu routier testé mais par les changements d'illumination imputables à l'ensoleillement. Cela dit, la vitesse de déplacement des véhicules accentue les difficultés apportées par l'ensoleillement.



**Figure 5.19** : suivi par FPFCC. 2 sources :  $S_2$  et  $S_4$  (cf. 5.2.4), séquence n°8.

### Trafic sur route de campagne

Un trafic sur route de campagne se traduit notamment par un environnement peu encombré (faible trafic, fond de scène simple) et une vitesse modérée. Ce type de route est probablement celui pour lequel il est le plus facile de suivre un véhicule.

La séquence n°9 a été acquise sur route nationale. Elle comprend 822 images correspondant à 3mn 39s. La vitesse de déplacement est en moyenne d'environ 90km/h. La distance parcourue est approximativement de 6,3km. La figure 5.21 présente le suivi obtenu par le FPFCC sur cette séquence

La qualité des résultats sur cette séquence est très satisfaisante comme il est possible de le constater sur la figure 5.21. Malgré quelques images très isolées où seule une faible fraction de l'objet est détectée (comme sur l'image à  $t = 280$ ), il s'agit là d'une des meilleures performances du FPFCC. Sur cette séquence relativement longue, le FPFCC tire profit de la simplicité d'un environnement rural. Les performances moyennes pour les quatre approches sont données par la figure 5.22.

D'après cette figure, le FPFCC montre une nouvelle fois une qualité de suivi clairement supérieure au FPFB modifié ou au FPFCC conjonctif. Le FPFCC disjonctif produit une qualité de suivi quasi identique. Comme pour la séquence précédente, les échecs des approches sont largement dus à des changements d'illumination plutôt qu'à l'environnement routier.

De manière générale, on peut dire que l'environnement rural ne pose pas de difficulté particulière et a même tendance à simplifier la tâche du FPFCC.

#### 5.4.4 Géométrie de la route

Dans cette sous-section, nous examinons l'influence de la géométrie de la route sur la qualité du suivi. Pour simplifier, nous réduisons cette géométrie à la possibilité de variations

**Figure 5.20 :** Performance moyenne pour la séquence test n°8 selon quatre approches de fusion différentes.  $t_{def} = 10^{-6}$  pour le FPFCC modifié. À gauche histogrammes simples, à droite histogrammes cumulés.

d'altitude (montée ou descente) ou de variations planaires (virages). Comme la présence de ces changements géométriques est très localisée, nous ne présentons ici que de courts extraits de séquences correspondant aux événements à tester.

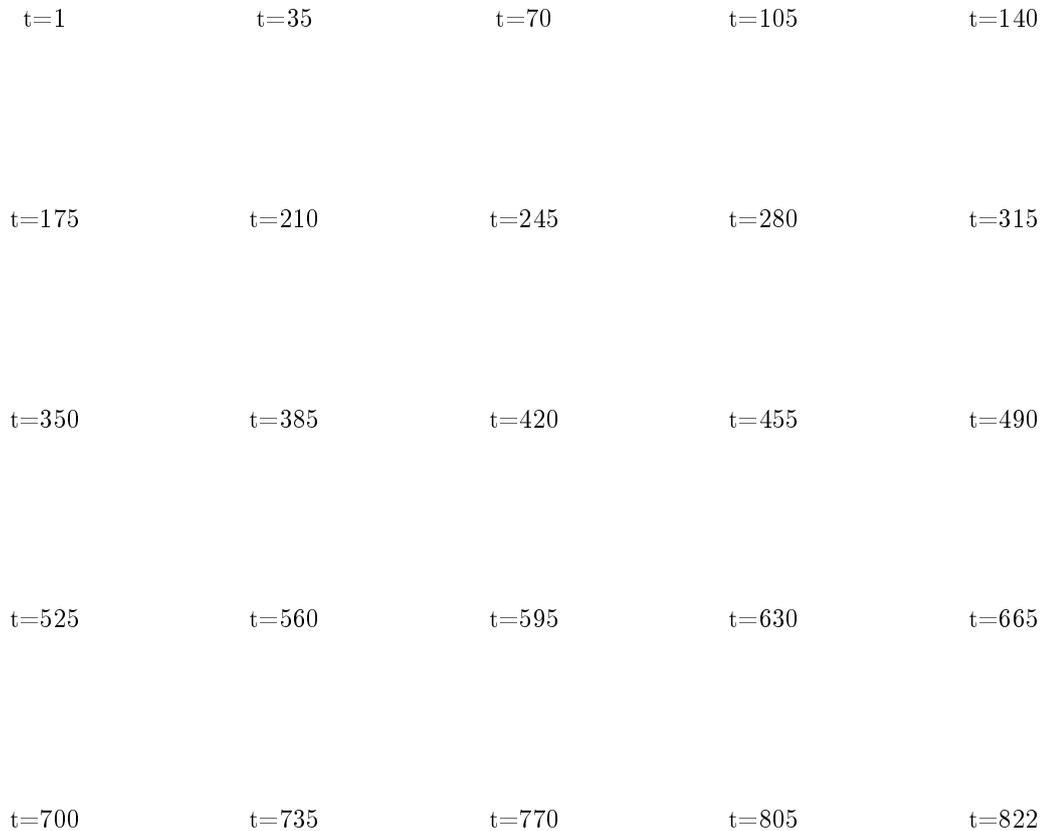
### Route avec dénivelé

La présence d'un dénivelé important va engendrer un mouvement vertical dans le plan image important de manière assez soudaine. Cet événement ne constitue pas une réelle difficulté en soi, il suffit d'augmenter la zone de recherche du filtre particulaire dans la dimension verticale. La figure 5.23 montre les résultats obtenues par le FPFCC sur un exemple de descente brutale. D'après la figure, le suivi est maintenu de manière très satisfaisante malgré ce dénivelé. Comme pour le cas des ombres portées, il n'est pas raisonnable d'étudier des performances moyennes sur seulement 10 images. Les autres approches se comportent tout aussi bien que le FPFCC sur cet exemple. Même le FPFCC conjonctif parvient à réagir suffisamment rapidement.

Étant donné qu'une montée ne présente pas plus de difficulté algorithmiquement parlant, nous ne montrons pas de test correspondant. Notons tout de même que comme pour le cas d'un milieu urbain, il s'avère nécessaire d'augmenter le nombre de particules  $N$  afin d'assurer la convergence du filtre. Plus la zone de recherche est grande, plus le besoin en particule se fait pressant. Il est regrettable de devoir laisser une zone de recherche si vaste en permanence pour le filtre, car cela ralentit le processus de suivi. Comme pour d'autre événement particulier, il est possible d'imaginer des solutions algorithmiques pour détecter un dénivelé et d'incorporer cette information contextuelle à l'architecture. Le nombre  $N$  pourrait être alors augmenté temporairement seulement. La détection d'un dénivelé est possible par exemple par l'intermédiaire d'algorithmes de détection de la ligne d'horizon [Fefilatyeu 06].

### Virages

À l'image des dénivelés, les virages vont demander d'étendre la zone de recherche de l'objet mais cette fois dans la dimension horizontale du plan image. La figure 5.24 montre les résultats obtenus par le FPFCC en présence d'un virage serré. Le FPFCC donne entière satisfaction sur cet

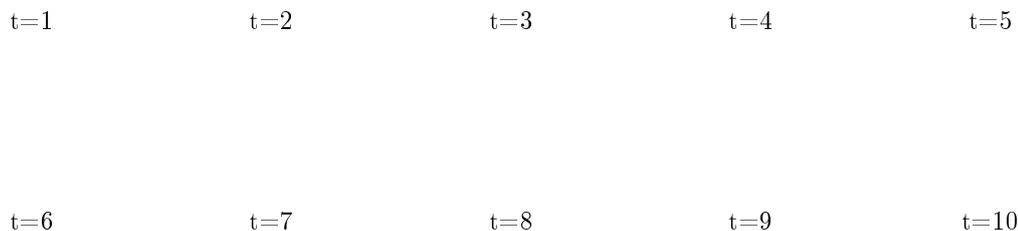


**Figure 5.21** : suivi par FPFCC. 3 sources :  $S_1$ ,  $S_3$  et  $S_4$  (cf. 5.2.4), séquence n°9.

exemple. Une nouvelle fois, il n'est pas opportun de donner des statistiques sur les performances dans cette expérience. Les autres approches donnent également de bons résultats à l'exception du FPFCC conjonctif qui perd le véhicule en fin de virage. Une zone de recherche plus étendue permettrait au FPFCC conjonctif de converger de nouveau vers l'objet suivi. Néanmoins, augmenter la zone de recherche est ici doublement pénalisant :  $N$  doit être augmenté et conséquemment le temps de calcul aussi, mais le risque de fouillis est également fortement accru. Sur l'exemple de la figure 5.24, on constate notamment qu'un véhicule-tiers de couleur proche de celui suivi constitue une telle situation. En augmentant trop fortement la zone de recherche, les algorithmes convergent tous vers le mauvais véhicule.

De même que pour les dénivelés, il est donc très regrettable de maintenir constamment une zone de recherche trop vaste de l'objet. A cet effet, un algorithme de détection des lignes de marquage au sol [Tarel 06] peut procurer une information contextuelle sur la courbure de la route. En utilisant cette information, la zone de recherche pourrait être augmentée temporairement seulement. D'autres perspectives à cet effet sont également évoquées en conclusion du mémoire.

**Figure 5.22 :** Performance moyenne pour la séquence test n°9 selon quatre approches de fusion différentes.  $t_{def} = 10^{-6}$  pour le FPFB modifié. A gauche histogrammes simples, à droite histogrammes cumulés.

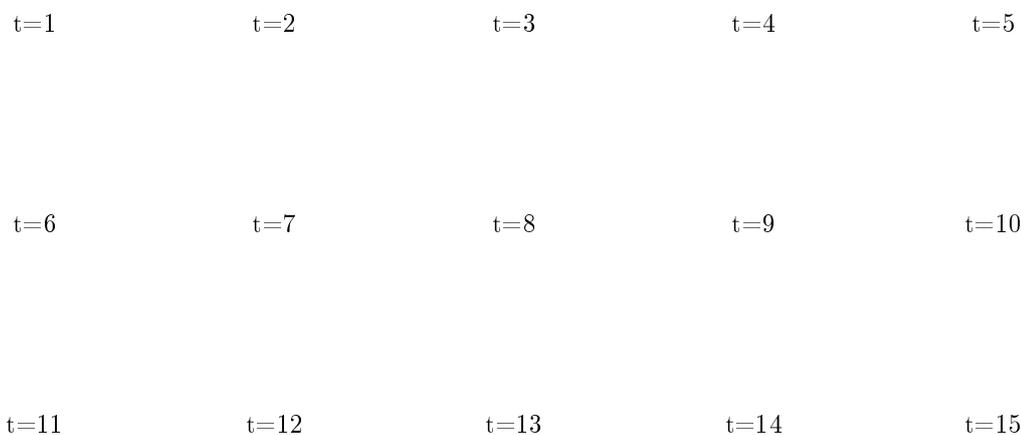


**Figure 5.23 :** suivi par FPFCC. 2 sources :  $S_2$  et  $S_3$  (cf. 5.2.4), extrait d'une séquence avec dénivelé.

## 5.5 Influence du nombre de sources

L'architecture proposée repose en partie sur la fusion de données, or un des principes de la fusion de données veut que la qualité des décisions tirées des différentes sources croisse avec le nombre de sources utilisées. Il est clair que le paramètre noté  $M$  dans les chapitres précédents et symbolisant le nombre de sources utilisées a une influence cruciale sur la qualité des résultats. Nous proposons dans cette partie un test visant à observer dans quelle mesure le principe évoqué ci-dessus est respecté par le FPFCC.

A l'évidence, la séquence permettant d'illustrer l'influence du nombre de sources doit être suffisamment difficile à traiter pour que l'apport de la fusion de données puisse être sensible. Nous avons donc sélectionné pour cette partie la séquence n°1 dans laquelle toutes les sources ont des raisons d'échouer individuellement. La figure 5.25 représente les résultats obtenus en faisant varier le nombre de sources  $M$  de 1 à 5. Chaque ligne de la figure correspond à une valeur de  $M$ . Pour des raisons évidentes toutes les combinaisons de sources possibles n'ont pas été représentées.



**Figure 5.24** : suivi par FPFCC. 2 sources :  $S_2$  et  $S_3$  (cf. 5.2.4), extrait d'une séquence avec virage.

*Stricto sensu*, le FPFCC ne respecte pas le principe. En effet le deuxième test de la troisième ligne ( $S_1, S_3$  et  $S_5$ ) donne clairement de moins bon résultat que le FPFCC uniquement pourvu de la source  $S_1$ . Néanmoins, il est également possible de constater que le meilleur résultat de chaque est supérieur ou égal à celui de la ligne précédente. En d'autres termes, en augmentant le nombre de source, le risque d'échec est minimisé et la qualité moyenne du suivi est maximisée, ce qui est bien là l'essentiel. En conclusion, le FPFCC respecte partiellement le principe évoqué ci-dessus, et l'intérêt de la fusion est une nouvelle fois démontré.

Un autre élément doit lui aussi être pris en compte dans ces expériences : la nature d'une source. Un second principe de la fusion de données stipule que des sources de nature variées apportent des informations complémentaires, et donc une fusion de meilleure qualité. Les deux premiers résultats de la deuxième ligne correspondent à des fusion de sources de mêmes natures. Il apparaît que leurs résultats sont inférieurs au meilleur résultat observé pour  $M = 2$  obtenu avec deux source de natures variées ( $S_2$  et  $S_4$ ).

## 5.6 Remarques et bilan des expérimentations

Comme nous le disions en début de chapitre, les séquences utilisées présentent des caractéristiques diverses de sorte que certaines peuvent être interprétées dans différentes catégories de tests. Pour résumer et recouper ces caractéristiques, nous proposons d'analyser la table 5.1

L'hétérogénéité des séquences-test ressort au travers de cette table. Cette diversité donne plus de poids aux conclusions et aux interprétations des résultats effectuées dans les paragraphes précédents de ce chapitre.

Il est également visible que l'architecture proposée procure la meilleure qualité de suivi dans la quasi-totalité des cas examinés. De plus, parmi ces cas, le FPFCC est régulièrement la

M=1

M=2

M=3

M=4

M=5

**Figure 5.25 :** Performance moyenne pour la séquence test n°1 avec le FPFCC pour différentes fusions de sources. M varie de 1 à 5.

seule approche capable de fournir la qualité maximale constatée. En ce sens, les résultats expérimentaux du chapitre 5 valident les choix théoriques émis tout au long du mémoire et confirment les résultats préliminaires présentés à la fin du chapitre précédent.

De manière synthétique, les performances intrinsèques de l'approche FPFCC indiquent que :

- la difficulté à pérenniser le suivi d'un véhicule croît avec la durée de la séquence.
- l'architecture est sensible aux conditions climatiques. Elle donne satisfaction en cas de temps gris ou d'ensoleillement important, mais tolère moins bien un temps pluvieux. Une pluie intense conjuguée à une vitesse élevée de déplacement constitue donc un cas limite de l'approche.
- l'architecture est sensible aux conditions d'éclairage, bien qu'elle fournisse de bons résultats dans les expériences présentées. Cette sensibilité se traduit par le fait qu'il est plus délicat de déterminer un paramétrage adapté permettant d'obtenir de bonnes performances.
- l'architecture est peu sensible au milieu routier, tout au plus un parcours en ville pose

**Tableau 5.1 :** Récapitulatifs des résultats et des caractéristiques des séquences test. \* : mesures n’ayant pas fait l’objet d’un relevé précis ; elles sont présentées à titre indicatif. En outre, nous ne pouvons certifier que le logiciel d’acquisition respecte rigoureusement la cadence d’acquisition indiquée.

Critères	seq. n°1	seq. n°2	seq. n°3	seq. n°4	seq. n°5	seq. n°6	seq. n°7	seq. n°8	seq. n°9
durée	1250 images 2mn 46s	292 images 39s	700 images 1mn 33s	433 images 1mn 56s	124 images 33s	300 images 1mn 20s	400 images 1mn 47s	488 images 1mn 5s	822 images 3mn 39s
cadence d’acquisition	7,5im/s	7,5im/s	7,5im/s	3,75im/s	3,75im/s	3,75im/s	3,75im/s	7,5im/s	3,75im/s
distance parcourue*	5,6km	1,1km	2,8km	3km	750m	2,3km	1,5km	2km	6,3km
vitesse de déplacement*	120km/h	100km/h	110km/h	de 5 à 110 km/h	70km/h	90km/h	de 0 à 50km/h	110km/h	90km/h
conditions climatiques	fort ensoleillement	fort ensoleillement	temps gris	forte pluie	temps gris	temps gris	temps gris	fort ensoleillement	ensoleillement modéré
milieu routier	autoroute	autoroute	autoroute	autoroute	voie rapide	voie rapide	ville	autoroute	route nationale
éclairage	jour	jour	jour	jour	jour/tunnel	nuit	jour	jour	jour
catégorie de véhicule suivi	monospace	berline	SUV	citadine	monospace	poils lourd	citadine	citadine	citadine
principales difficultés	ombres portées, fouillis, changements d’illumination	virage, fouillis, changements d’illumination	dénivelé, changements d’échelle	changements d’illumination, occultation légère (spray)	changements d’illumination, fouillis	fouillis, faible quantité d’information	changements d’échelle, fouillis	changements d’illumination, ombres portées, fouillis	changements d’illumination, changements d’échelle
meilleure(s) approche(s) en termes de suivi	FPFCC	FPFB modifié	FPFCC, FPFB modifié	FPFCC	FPFCC	FPFCC	FPFCC, FPFB modifié, FPFC disjonctif	FPFCC	FPFCC
temps de calcul par image	3s/im	4,7s/im	2,8s/im	2,3s/im	2,8s/im	5,2s/im	1,6s/im	3,2s/im	4,3s/im
$N$	200	600	600	600	600	600	500	500	600
$M$	5	5	3	4	3	4	2	2	3

quelques problèmes supplémentaires en terme de coût calculatoire afin de prendre en compte les changements d’échelle du véhicule.

- l’architecture est peu sensible à la géométrie de la route. Les résultats montrent qu’il suffit d’étendre la zone de recherche de l’objet dans une image afin de résoudre le problème au détriment toutefois du temps de calcul.
- l’architecture respecte l’essentiel du prédicat selon lequel l’augmentation du nombre de sources améliore les performances globales du système. Plus le nombre de sources utilisées est grand, plus les chances d’obtenir un suivi de longue durée et de qualité augmentent. Il s’agit là d’un des objectifs fondamentaux qui ont motivé toute notre approche.

Sur chacun de ces points, l’approche proposée montre naturellement des limites. Ces limites sont sujettes à être repoussées à l’aide de travaux de recherche annexes à ceux présentés ici. De nombreuses voies d’amélioration restent ouvertes et sont répertoriés en conclusion générale du mémoire.

Nous rappelons que le but du FPFCC est de prouver que le recours à une technique de fusion plus sophistiquée que la règle Bayésienne probabiliste permet d’améliorer les performances d’un algorithme de suivi de type filtre particulière. Ainsi, d’autres versions du filtre sont susceptibles de donner de meilleurs résultats sur les séquences que nous présentons, mais en ayant recours à d’autres solutions que la fusion de données explicite. A ce titre, nous estimons que le présent chapitre apporte les garanties suffisantes pour affirmer que les fonctions de croyances et la SCC apportent un gain réel au filtrage particulière.

Nous regrettons de ne pas être en mesure de comparer notre approche à d’autres versions du filtre particulière voire à d’autres algorithmes de suivi. Ces approches sont souvent trop hétérogènes pour pouvoir être comparées de façon directe et nécessitent un investissement important en terme de développement informatique. De plus, nombre de méthodes reposent sur des hy-

pothèses non vérifiées dans notre cas (information *a priori* abondante, capteurs multiples et de natures diverses). Étant donné la très faible quantité d'information *a priori* exploitée par notre approche, nous estimons que les résultats sont très encourageants en vue d'élaborer des systèmes plus complets et plus globaux.

## 5.7 Conclusion

Dans ce chapitre, nous avons étudié d'un point de vue pratique l'algorithme FPFCC issu des travaux menés dans les chapitres précédents. La mise en œuvre de cette approche a été explicitée. Nous nous sommes attaché à tester l'architecture proposées dans diverses situations particulières de circulation automobile. Ces différents tests ont permis de définir les limites de l'approche et de prouver son efficacité par rapport à d'autres méthodes faisant appel à la fusion de données et servant de référence.

Au vu des conclusions de ce chapitre, nous pouvons dire que l'objectif principal de ces travaux de thèse a été atteint. Cet objectif consistait à surmonter les principaux points de difficultés relatifs au suivi de la face arrière du véhicule précédent, que nous avons nommés "événements perturbateurs". Ces événements sont présents dans les séquences sur lesquelles le FPFCC a été testé et les résultats montrent une nette amélioration des performances due à l'usage d'une méthode de fusion adaptée à la problématique. Ce chapitre clôt donc le raisonnement scientifique formulé dans ce mémoire articulé autour de trois mouvements principaux : étude de l'état de l'art du domaine, réflexions et justifications théoriques puis démonstrations pratiques.

## Conclusion et Perspectives

La vision par ordinateur et plus précisément le suivi d'objets, thème abordé dans ces travaux de thèse, posent de nombreux défis en terme de traitement de l'information et de conception de systèmes dits intelligents. Contrairement aux idées reçues, il n'y a pas d'étape au sein d'un algorithme de suivi qui soit à proprement parlée "facile". La représentation de l'objet, la caractérisation de l'objet, et la manière de reconnaître un modèle d'objet à chaque instant d'une séquence, toutes ces étapes nécessitent une réflexion mûrie.

Ces travaux de thèse se destinent plus particulièrement au suivi de la face arrière du véhicule précédent le véhicule instrumenté dans les scènes routières. Dans ce mémoire, nous avons identifié d'emblée divers événements pouvant appartenir aux séquences routières et étant susceptibles de fortement perturber un algorithme de suivi. Ces événements, dits perturbateurs, ont des conséquences néfastes en cascade au niveau de toutes les étapes que nous évoquions à l'instant. Les scènes routières sont à cet égard peu évidentes à traiter par des approches classiques de suivi d'objets car la fréquence des événements perturbateurs y est élevée.

Pour répondre à la problématique posée, nous avons axé notre réflexion autour de trois axes principaux. Le premier axe est constitué d'une analyse de l'état de l'art des différentes étapes constitutives d'un algorithme de suivi. A l'issue de cette analyse, nous avons choisi de travailler autour du thème de la fusion de données, qui est une des principales solutions avancées par la communauté afin de lutter contre les événements perturbateurs et leurs erreurs inhérentes. Une réflexion théorique d'une part sur la manière d'insérer cette fusion au sein d'un algorithme de suivi et d'autre part sur la mise au point même d'un processus de fusion adapté à nos besoins constitue donc le second axe. Enfin, le troisième axe clôture le raisonnement scientifique en apportant une mise en œuvre pratique de la solution algorithmique avancée et de nombreux tests spécifiques aux conditions que l'on peut rencontrer dans les scènes routières.

Les deux premiers chapitres ont été consacrés au premier axe, c'est-à-dire à l'état de l'art. La séparation en deux chapitres de cet état de l'art se justifie par le fait qu'un algorithme de suivi se distingue par deux modules en interaction mais de natures différentes : la modélisation de l'objet et la procédure de suivi. La modélisation de l'objet a fait l'objet d'une étude selon trois voies principales d'investigation, correspondant à trois sensations perceptuelles distinctes : la texture-couleur, la forme et le mouvement. Dans chacune de ces catégories, nous avons examiné les méthodes les plus couramment employées en écartant celles non adaptées à notre cadre applicatif.

Les principales procédures de suivi mises au point à ce jour ont également été listées dans le second chapitre. Cependant, il n'est pas possible de définir une classification claire des procédures de suivi étant donné que nombres d'entre elles ont recours à des techniques hybrides. On notera tout de même quelques techniques spécifiques : l'optimisation d'une grandeur scalaire, la maximisation d'une probabilité ou le calcul d'un appariement de détection optimal.

A l'issue de ces deux chapitres, nous disposons d'une représentation de l'objet (boîte englo-

bante), d'une caractérisation d'objets à l'aide de cinq extracteurs d'information (cooccurrences couleur pondérées, densité de couleur, carte de symétrie, carte de symétrie des contours et algorithme de détection de mouvement de Lefaix) et d'une procédure de suivi (filtre particulière). Les cinq extracteurs sont amenés à être vus comme des sources d'information par la suite. Tout l'enjeu est de faire collaborer harmonieusement ces extracteurs à l'aide d'une fusion de données, car certains extracteurs résistent mieux que d'autres à certains événements perturbateurs. Ainsi, la pérennité du suivi peut être améliorée en conjuguant les qualités de chaque extracteur. Une nouvelle fois, les choix formulés à la fin du chapitre 2 sont justifiés par des besoins pratiques liés à notre application selon deux critères principaux : efficacité en terme de pouvoir discriminant et complexité calculatoire.

Le second axe est lui aussi divisé en deux chapitres. Comme nous le disions déjà auparavant, la réflexion théorique a consisté dans un premier temps à justifier du choix d'un cadre théorique de fusion de données pouvant s'insérer dans la procédure de suivi choisie, puis dans un deuxième temps à mettre au point un procédé de fusion répondant plus spécifiquement à nos besoins. Pour la première phase de ce plan, une présentation des différents travaux existant et mêlant filtres particuliers et fusion de données a été proposée. Cette présentation nous a permis de mettre en évidence que ce couplage se fondait principalement autour de deux techniques : une stratégie de fusion ou fusion implicite, ou bien une fusion explicite à l'aide de méthodes probabilistes. Dans le cadre de la fusion de données, d'autres méthodes peuvent être employées pour effectuer une fusion explicite. Entre autres, la théorie des fonctions de croyances offre plus de possibilités tout en étant compatible avec les probabilités. Nous avons donc proposé un schéma permettant l'usage d'une fusion par fonction de croyance au sein d'un filtre particulier.

Après une rapide étude, un constat intéressant émerge : le filtre crédibiliste et le filtre probabiliste offrent des atouts complémentaires. Nous nous sommes alors dirigés vers une étude plus théorique du cadre des fonctions de croyance qui puisse allier les atouts de ces deux filtres. Il nous a alors fallu définir nos besoins en termes de fusion de données. Ces besoins se traduisent par la nécessité de fusionner des sources imprécises et non fiables. En examinant les propriétés des règles de fusion de la théorie des fonctions de croyance, nous avons mis au point une nouvelle stratégie de combinaison, notée SCC, permettant d'atteindre l'objectif de cumul des atouts des deux filtres. D'après une étude menée en fin de chapitre 4, la SCC se distingue des autres règles de combinaison sur notre application de suivi. Le filtre particulier utilisant la SCC est noté FPFCC.

Le dernier axe de notre raisonnement a donc consisté à une étude plus pratique du FPFCC dans le cadre du suivi de la face arrière du véhicule précédent. Plusieurs tests ont été menés dans ce dernier chapitre afin de couvrir un large spectre des situations que l'on peut rencontrer quotidiennement en conduisant un véhicule. Ces conditions sont notamment liées à la météorologie, à l'éclairage, au type de route empruntée et à la géométrie de la route empruntée. Au regard des expérimentations menées dans ce chapitre, le FPFCC procure un net gain en robustesse en termes de suivi, prouvant ainsi l'intérêt de notre stratégie de combinaison au sein d'un filtrage particulier.

Si nous estimons avoir contribué à l'amélioration du suivi d'un véhicule grâce à la fusion

de données, il n'en reste pas moins que l'architecture globale résumée par la figure 5.2 est parfaite en de nombreux points. Tout d'abord au niveau du temps de calcul nécessaire à son exécution, de nombreuses pistes ont été évoquées dans le mémoire (parallélisation du filtre particulaire, implémentation hardware et sous-échantillonnage). Le sous-échantillonnage a déjà été mis en place dans nos tests et nous avons constaté qu'il réduit le temps de calcul de manière spectaculaire. Toutefois, ce sous-échantillonnage a été effectué à fréquence constante pour une séquence donnée dans nos expériences. Ainsi, pour des séquences dans lesquelles ont lieu des changements d'échelle de l'objet suivi, le sous-échantillonnage doit être suffisamment intense pour que le temps de calcul soit raisonnable quand la taille de l'objet est grande, ce qui se fait au détriment des images dans lesquelles l'objet est de petite taille et où le sous-échantillonnage supprime trop d'information. Il serait donc plus pertinent d'opérer un sous-échantillonnage proportionnellement à la taille de l'estimation précédente de la boîte englobante. Grâce à ce procédé, l'objet garderait une surface pixellaire quasi-fixe tout au long de la séance, le temps de calcul serait quasiment le même pour chaque image et le réglage des paramètres des extracteurs plus aisé.

Nous estimons qu'en conjuguant ces divers efforts, l'architecture proposée a de très bonnes chances d'atteindre un temps d'exécution proche du temps-réel, or cette condition s'avère nécessaire pour certaines exploitation souhaitée d'un suivi de véhicule comme la conduite autonome.

L'autre grand champ d'investigation est bien entendu la fiabilité du système. L'une des grandes limitations que nous avons posée dans ce mémoire concerne la quantité d'information *a priori*. L'algorithme présenté utilise une seule imagerie de l'objet suivi d'une taille moyenne de  $25 \times 25$  pixels, ce qui correspond à une quantité d'information très faible. Il serait dommage de se contenter de si peu d'information si l'architecture doit être employée à une plus large échelle. Il existe de nombreux extracteurs utilisant des techniques d'apprentissage et dont les performances dépassent largement celles des extracteurs employés dans nos tests. L'architecture proposée ne pose pas de condition particulière sur les extracteurs et ceux-ci sont donc interchangeables sans grandes difficultés. Il va de soi que des sources plus fiables améliorerait grandement nos résultats. Malgré des caractéristiques plus ou moins fiables et plus ou moins informatives, on constate que l'approche proposée se comporte convenablement. L'usage de telles caractéristiques permet, de plus, de mieux faire ressortir les limites de l'approche et d'accentuer le rôle de la fusion dans l'obtention des résultats.

La robustesse de l'approche peut être également améliorée par d'autres moyens, que nous évoquons déjà dans le dernier chapitre, à savoir l'utilisation de plus d'informations contextuelles. Il serait particulièrement intéressant de calculer en parallèle et à l'aide de modules spécifiques des données concernant l'éclairage, l'illumination ou la géométrie de la route. Ces données permettent d'adapter et de mieux diriger la puissance de calcul. Ces informations contextuelles peuvent être obtenues par des procédés algorithmiques sur la base même de l'image traitée, mais elles peuvent être également obtenue de manière plus directe et moins coûteuse en temps de calcul grâce à des capteurs spécifiques, ce qui nous amène tout droit à la voie d'amélioration suivante : l'augmentation des données observées.

Bien que la méthode de fusion joue un rôle majeure dans le processus que nous proposons la quantité d'information observée reste peu importante : une seule image par instant. Les sources pourraient être diversifiées à l'aide de capteurs de nature différentes, par exemple : un

radar, un laser, un GPS, une sonde crépusculaire, une caméra infra-rouge, un dispositif de stéréoscopie, ou encore une centrale à inertie. Certes l'incorporation de telles sources nécessiteraient un travail supplémentaire de modélisation pour l'obtention de masses de croyance, mais cette tâche ne paraît pas insurmontable tant le cadre des fonctions de croyance se prête facilement à la modélisation. De plus, comme nous l'avons dit, la diversité des sources est un atout majeure dans la fiabilité de la décision découlant de la fusion. L'avenir des systèmes de transport intelligents semble d'ailleurs se diriger vers une forte augmentation et diversification des données avec notamment la communication inter-véhicules ou la communication entre un véhicule et tous les éléments de l'environnement routier (*cf.* [Aycard 06] pour exemple de travaux).

En outre, l'augmentation du nombre de capteurs de même nature serait d'un secours important. Il est possible d'imaginer poster plusieurs caméras sur le véhicule. Certaines caméras pourraient acquérir un champ de vue plus large tandis que d'autres se focaliseraient sur des régions plus précises et pourraient ainsi offrir des informations complémentaires et hiérarchisées. De même, la direction des caméras pourrait varier, ce qui pourrait s'avérer intéressant en cas de virage serré par exemple ou de franchissement d'intersection. Le véhicule TerraMax utilisé par Broggi *et al.* dans leurs travaux de recherches ne comporte pas moins de 5 caméras rien qu'à l'avant du véhicule (*cf.* figure 5.26). Ce véhicule a été réalisé dans le cadre du challenge DARPA organisé par le département de la défense des États-Unis d'Amérique, en vue de mettre au point un véhicule en conduite autonome.

**Figure 5.26 :** Véhicule autonome TerraMax. Figure extraite de [Braid 06].

Enfin, au delà de l'amélioration des performances de l'approche proposée, cette dernière peut également être adaptée à des ambitions plus générales telles que le suivi multi-objets. Les besoins exprimés par les industriels vont en ce sens, notamment dans le cadre de la télésurveillance ou du traitement d'événements sportifs par équipes. Cette évolution pose cela dit quelques défis au sens où les besoins en capacité de calcul se trouveraient nécessairement accrus. Néanmoins, un tel algorithme apporterait plus de connaissances sur les solutions potentielles aux problèmes de suivis relatifs à chaque objet individuellement. En observant l'interaction de ces objets, il serait plus aisé de traiter les occultations ou les situations de fouillis.

Pour conclure, il apparaît que le domaine du suivi d'objets reste ouvert à des améliorations sur de nombreux points. Bien que de très grands pas en avant aient déjà été effectués, la recherche dans ce domaine est encore d'actualité et devrait le rester pour encore de nombreuses années avant d'atteindre la généricité couplée à la fiabilité exigible pour des applications de sécurité routière ou militaires. Nous espérons par les travaux présentés dans ce mémoire avoir contribué modestement à l'achèvement de ce projet.

# Annexes



## Annexe A

# Matrices de cooccurrences couleur pondérées

Cet annexe présente les matrices de cooccurrences couleur pondérées. Il s'agit d'une technique de caractérisation texture-couleur s'appuyant sur les matrices de cooccurrences d'Haralick [Haralick 73]. Ces matrices sont des attributs statistiques. Elles dénombrent les couples de pixels selon leurs deux couleurs. Les couples sont formés d'après une relation de voisinage définie d'une part en terme de distance séparant les deux pixels et en terme d'orientation d'autre part.

Une distance d'un pixel est souvent choisie, les pixels du couple sont donc adjacents. L'orientation est l'angle formé entre l'axe des abscisses et la droite formée par le couple de pixel. En général, l'orientation  $\theta$  est quantifiée et se borne aux quatre valeurs suivantes :  $\theta \in \{0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}\}$ . Les orientations supérieures à  $\pi$  n'ont pas d'utilité car les matrices sont symétriques. La distance est repérée par un vecteur  $\vec{d}$ , que l'on peut définir d'après l'angle  $\theta$  :  $\vec{d} = \begin{pmatrix} r \cos \theta \\ r \sin \theta \end{pmatrix}$ , avec  $r$  la norme de  $\vec{d}$ .

On rappelle la formule des cooccurrence dans le cas de couleurs :

$$M_{\vec{d}}(i, j) = \frac{\# \left( p \in D / I(\vec{p}) = c_1, I(\vec{p} + \vec{d}) = c_2 \right)}{\#(D)} \quad (\text{A.1})$$

où  $I$  est une image définie sur  $D$  que l'on cherche à caractériser,  $\vec{d}$  un vecteur de translation,  $(c_1, c_2)$  un couple de couleurs donné, et  $\#$  le cardinal d'un ensemble.

Il existe autant de matrices que l'on définit de vecteur  $\vec{d}$ . En général, on se limite à quatre matrices, pour les quatre orientations précédemment citées. Chaque cellule de chaque matrice est indexée par un couple de couleur  $(c_1, c_2)$ . Sur la figure A.1, la valeur  $a$  correspond au nombre de fois dans l'image où la couleur  $c_1$  est séparée de la couleur  $c_2$  par  $\vec{d}$ . Si  $N_c$  est le nombre de couleurs quantifiées, la taille d'une matrice est donc  $N_c \times N_c$ .

### A.1 Ajout de la pondération

Les matrices de cooccurrences couleur caractérisent d'une part la couleur et d'autre part la répartition spatiale des couleurs, c'est donc un attribut texture-couleur. Néanmoins nous dressons le constat suivant : si l'objet caractérisé (ou l'image caractérisée) subit un changement d'illumination, alors les couleurs vont changer. Il est donc fort possible que des couples de pixels qui étaient

**Figure A.1 :** Matrice de cooccurrence pour un vecteur  $\vec{d}$  donné..

de couleurs  $(c_1, c_2)$  deviennent  $(c'_1, c'_2)$ , et donc tombe dans une autre cellule de la matrice. Ainsi les matrices obtenues après changement d'illumination (même très faible) seront totalement différentes des matrices avant le changement. Il sera par la suite impossible de les faire correspondre.

Une telle sensibilité au changement d'illumination est inacceptable en suivi d'objets. Nous proposons donc de modifier la façon dont sont dénombrés les couples couleur en attribuant un poids ou un droit de vote spécifique à chaque couleur. Nous commençons par repérer les couleurs principales de l'objet devant être caractérisé. Ces couleurs sont appelées couleurs de références :  $\{c_{ref1}, \dots, c_{refn}\}$ . Le nombre  $n$  de couleurs de référence est en pratique assez limité (3 à 5).

Ensuite dans le processus de dénombrement, pour chaque couleur observée  $c_{obs}$ , la distance euclidienne avec chaque couleur de référence est calculée. La figure A.2 illustre cette opération. Parmi toutes cette distance, seule la distance minimale est conservée ainsi que l'étiquette  $\tilde{c}_{ref}$  de

**Figure A.2 :** Matrice de cooccurrence pour un vecteur  $\vec{d}$  donné..

la couleur de référence la plus proche de la couleur observée. Une fonction  $k()$  de type noyau<sup>1</sup> est appliquée à la distance minimale :

$$w_1 = k(\|c_1 - \tilde{c}_{ref1}\|) \quad (\text{A.2})$$

Cela revient à dire que la couleur observée  $c_1$  est associée à la couleur de référence  $c_{ref1}$  avec le

<sup>1</sup>Les propriétés de type noyau ne sont pas essentielles, mais la forme de la courbe de ces fonctions présente un intérêt.

score  $w_1$ . Au lieu d'incrémenter la cellule  $M_{\vec{d}}(c_1, c_2)$  de 1 comme pour les cooccurrences couleur normales, c'est la cellule  $M_{\vec{d}}(c_{ref1}, c_{ref2})$  qui est incrémentée de  $w_1 w_2$ .

La pondération a un double intérêt. D'une part les couleurs sont mieux quantifiées pour résister aux changements d'illumination, car des couleurs proches correspondent à la même cellule de la matrice. D'autre part, la taille de la matrice est fortement réduite et devient de taille  $n \times n$ .

Dans la suite de l'annexe, nous présentons quelques résultats concernant les matrices de cooccurrences couleur pondérées (MCCP) fondé sur des critères de suivi d'objets. L'algorithme de suivi est mis en œuvre :

- au niveau représentation : par une boîte englobante
- au niveau caractérisation : par les MCCP
- au niveau procédural : par un filtre particulière classique

Seule la partie caractérisation évolue dans nos tests, les paramètres du filtre particulière sont fixes pour toutes les expériences.

## A.2 Test comparatif sur une séquence routière

Nous testons ici les MCCP sur une séquence routière simple longue de 71 images. Durant cette séquences, l'objet suivi (la face arrière d'un monospace gris) subit des changements d'illumination d'amplitude modeste, principalement dus à la réflectance de la peinture métallisée de l'objet. Les MCCP ont été mises au point afin de supporter ce type de changement. La vitesse du véhicule est d'environ 50km/h. L'acquisition a été réalisée par temps ensoleillé et en milieu urbain.

Les MCCP sont étudiées pour trois fonctions noyaux  $k()$  différentes : le noyau rationnel, le noyau gaussien, et le noyau d'Epanechnikov. La figure A.3 représentent ces trois fonctions. Naturellement, chacune de ces fonctions vaut 1 en 0. Elles sont également toutes globalement de

**Figure A.3** : Différents noyaux utilisés pour les MCCP..

formes similaires, avec un pic central et des limites à l'infini valant 0. Un seul paramètre est à régler afin de contrôler la largeur du pic central. Ce paramètre est très important car c'est lui qui

(a)

(b)

(c)

(d)

(e)

**Figure A.4 :** (a) : suivi avec cooccurrences classiques, (b) : suivi avec M CCP Epanechnikov , (c) : suivi avec M CCP Gaussien , (d) : suivi avec M CCP rationnel, (e) suivi avec Gabor.

détermine jusqu'à quel point il est possible d'associer des couleurs les unes aux autres. Plus ce pic est large, plus est facile pour une couleur observée d'être associée à une couleur de référence avec un score élevé.

Le noyau d'Epanechnikov a été choisi car son support est borné, il est donc le plus sélectif des trois noyaux. Le noyau rationnel au contraire a été choisi car il converge lentement et est donc beaucoup moins sélectif. Enfin le noyau gaussien fait figure de solution intermédiaire entre les deux autres.

La figure A.4 présente visuellement les résultats sur la séquence des M CCP pour les trois noyaux, pour les cooccurrences couleurs classiques et pour les filtres de Gabor. Les filtres de Gabor constituent une autre méthode texture-couleur de référence, et ont été appliqués sur chaque plan couleur pour constituer des attributs.

D'après cette figure, il apparaît que les filtres de Gabor supportent mal les variations de luminosité de la scène. Les matrices de cooccurrences produisent un suivi partiel de l'objet sur plus de la moitié de la séquence. A l'inverse les M CCP se montrent tolérantes aux variations de luminosité et maintiennent un suivi tout au long de la séquence. Pour plus de précision, la figure A.5 donne l'évolution du taux de suivi (*cf.* 2.6.3 pour la définition de taux) pour cette séquence L'analyse des taux de suivi permet d'observer que les noyaux gaussiens et d'Epanechnikov produisent les meilleurs résultats, tandis que le noyau rationnel est légèrement en retrait.

**Figure A.5 :** Évolution du taux de suivi au cours d'une séquence pour les différentes approches texture-couleur examinées.

**Figure A.6 :** Changement artificiel de luminosité, première image de la séquence  $\gamma = 1$ , dernière image  $\gamma = 0,5$ .

Le noyau rationnel étant moins sélectif, il a tendance à mettre en correspondance des couleurs trop éloignées, d'où une certaine imprécision par rapport aux autres noyaux.

### A.3 Test de résistance aux changements d'illumination

Nous cherchons ici à déterminer les limites de notre approche en terme de résistance aux changements lumineux. Pour ce faire, nous prenons la même séquence que précédemment à laquelle nous appliquons une transformation  $\gamma$  pour intensifier sa luminosité au fur et à mesure de la séquence. La figure A.6 montre la première et la dernière image de cette séquence modifiée afin de mettre avant le changement significatif d'illumination provoqué par la transformation. La figure A.7 donne les taux de suivi pour les MCCP avec les trois noyaux différents.

Il est clair que le suivi est globalement de beaucoup moins bonne qualité que sur la séquence naturelle. Les noyaux d'Epanechnikov et gaussien sont les plus robustes. Le noyau d'Epanechnikov maintient le suivi avec une qualité satisfaisante ( $r > 0,8$ ) jusqu'à l'image 33 où  $\gamma$  vaut alors 0,7643. Le noyau rationnel étant moins sélectif, il aurait pu cette fois-ci être le plus robuste. En réalité l'expérience montre qu'il est inutile d'élargir trop la base du pic, car cela revient à

**Figure A.7 :** Évolution du taux de suivi en fonction du temps et de la correction  $\gamma$ .

**Figure A.8 :** Suivi avec M CCP gaussien pour un objet en rotation.

mettre en correspondance des couleurs qui ne devraient pas l'être. Autrement dit un pic trop large bruite les matrices, mais n'apporte pas de robustesse. Un pic trop mince revient à utiliser les cooccurrences classiques. Aux vues des ces conclusions, nous recommandons l'usage des noyaux d'Epanechnikov ou gaussien. Dans nos expériences, une largeur de support pour le noyau d'Epanechnikov et un écart-type pour le gaussien, tolérant des distances euclidiennes entre triplet colorimétrique de l'ordre de 20 à 40, offrent d'assez bons résultats.

#### A.4 Test d'invariance aux rotations et changements d'échelle

Cette partie montre l'invariance de notre méthode aux rotations d'axe orthogonal au plan image et aux changements d'échelle de l'objet caractérisé. Concernant les rotations, il suffit de calculer les matrices dans suffisamment de directions  $\theta$  pour devenir quasi-invariant. Dans la très grande majorité des cas, les quatre orientations classiques évoquées précédemment suffisent. La figure A.8 montre le maintien du suivi avec des M CCP malgré le mouvement rotationnel. Concernant les changements d'échelle, une normalisation des matrices suffit. Il s'agit en effet de faire en sorte que les résultats du dénombrement soient les mêmes quel que soit le nombre de pixels pris en compte. Toutefois, il faut remarquer que tous les couples de pixels ne sont pas présents dans les mêmes proportions à toutes les échelles. La différence se situe au niveau des éléments diagonaux des matrices qui symbolisent des couples de pixels de mêmes couleurs. Ces couples appartiennent donc à des régions de couleur quasi uniforme. A l'inverse les couples couleur des autres éléments de la matrice correspondent à des transitions de couleurs, et donc à des frontières de régions. Or d'une échelle à l'autre, si le nombre de couples sur une frontière est divisé par 2, le nombre des couples dans une région est divisé par 4. La figure A.9 illustre le phénomène. Cela vient de la différence de dimension entre frontières et surfaces. Pour pallier cet effet, les éléments non-diagonaux des matrices sont donc multipliés par 2 avant normalisation. La

**Figure A.9 :** Exemple d'un changement d'échelle d'un carré vert sur fond jaune. Surfaces et périmètres ne sont pas modifiés par le même rapport.

figure A.10 donne l'exemple d'un suivi maintenu avec les MCCP malgré un changement d'échelle de l'objet.

**Figure A.10 :** Suivi avec MCCP gaussien pour un objet changeant d'échelle.



## Annexe B

### Filtres de Kalman

Cet annexe présente le filtre de Kalman en complément des éléments donnés dans la section 2.4.1. La présentation proposée reprend celle d'Arnaud [Arnaud 04a]. Soit le système linéaire suivant :

$$X_t = F_t X_{t-1} + b_t + b'_t \quad (\text{B.1})$$

$$Y_t = G_t X_t + c_t + c'_t \quad (\text{B.2})$$

avec  $F_t$  et  $G_t$  deux matrices,  $b_t$  et  $c_t$  deux vecteurs et  $b'_t$  et  $c'_t$  deux vecteurs de bruits blancs gaussiens centrés. Les matrices et les vecteurs  $b_t$  et  $c_t$  sont des données du système connues *a priori*. On suppose de plus connues les matrices de covariances de  $b'_t$  et  $c'_t$  respectivement notées  $Q_t$  et  $R_t$ . Les bruits sont par hypothèse indépendants.  $X_t$  est le vecteur d'état dont nous cherchons à déterminer les composantes.  $Y_t$  est le vecteur d'observation. L'état initial  $X_0$  est également connu.

Le système peut alors être réécrit sous forme probabiliste :

$$X_0 \sim \mathcal{N}(X_0, \bar{X}_0, \Sigma_0) \quad (\text{B.3})$$

$$X_t | X_{t-1} \sim \mathcal{N}(X_t, F_t X_{t-1} + b_t, Q_t) \quad (\text{B.4})$$

$$Y_t | X_t \sim \mathcal{N}(Y_t, H_t X_t + c_t, R_t) \quad (\text{B.5})$$

avec  $\forall t$ ,  $\bar{X}_t$  l'espérance de  $X_t$  et  $\Sigma_t$  la covariance de  $X_t$ . L'espérance et la covariance suffisent pour caractériser entièrement la densité de filtrage, ce sont donc ces objets mathématiques qu'on cherche à estimer récursivement par le filtre de Kalman. On désigne dans la suite de cet annexe, une variable indexée par  $k|k'$ , comme la variable au temps  $k$  sachant les  $k'$  premiers états.

Le filtre s'exprime alors principalement en deux temps :

– Étape de prédiction

$$\bar{X}_{t|t-1} = F_t \bar{X}_{t-1|t-1} + b_t \quad (\text{B.6})$$

$$\Sigma_{t|t-1} = F_t \Sigma_{t-1|t-1} F_t^T + Q_t \quad (\text{B.7})$$

– Étape de correction

$$K_t = \Sigma_{t|t-1} G_t^T (G_t \Sigma_{t|t-1} G_t^T + R_t) \quad (\text{B.8})$$

$$\bar{X}_{t|t} = \bar{X}_{t|t-1} + K_t [Y_t - (G_t \bar{X}_{t|t-1} + c_t)] \quad (\text{B.9})$$

$$\Sigma_{t|t} = (Id - K_t G_t) \Sigma_{t|t-1} \quad (\text{B.10})$$

$K_t$  est une matrice intermédiaire appelé gain de Kalman permettant de favoriser la prédiction ou l'observation selon ses valeurs. Le terme  $\nu_t = Y_t - G_t F_t (\bar{X}_{t|t-1} + c_t)$  est appelé innovation. Lorsqu'on présente le filtre de Kalman en utilisant la théorie de l'estimation, la matrice  $\Sigma_{t|t}$  est souvent appelée matrice de covariance de l'erreur. La trace de cette matrice est égale à l'erreur au sens des moindres carrés entre l'état et son estimé.

## 2. Filtre de Kalman étendu

Le filtre de Kalman étendu autorise l'emploi de modèles non-linéaires :  $f_t()$  et  $g_t()$  (cf. présentation du filtre Bayésien en 2.4.1). On utilise alors la formule de Taylor pour approximer les fonctions  $f_t$  et  $g_t$  par des modèles non-linéaires.

$$F_t \approx \left. \frac{df}{dX_t} \right|_{X_t=\hat{X}_{t-1}} \quad (\text{B.11})$$

$$G_t \approx \left. \frac{dg}{dX_t} \right|_{X_t=\hat{X}_{t-1}} \quad (\text{B.12})$$

$F_t$  et  $G_t$  sont ici les matrices Jacobiennes des fonctions  $f_t$  et  $g_t$ . L'algorithme devient :

– Étape de prédiction

$$F_t = \left. \frac{df}{dX_t} \right|_{X_t=\hat{X}_{t-1}} \quad (\text{B.13})$$

$$\bar{X}_{t|t-1} = f_t(\bar{X}_{t-1|t-1}) \quad (\text{B.14})$$

$$\Sigma_{t|t-1} = F_t \Sigma_{t-1|t-1} F_t^T + Q_t \quad (\text{B.15})$$

– Étape de correction

$$G_t \approx \left. \frac{dg}{dX_t} \right|_{X_t=\hat{X}_{t-1}} \quad (\text{B.16})$$

$$K_t = \Sigma_{t|t-1} G_t^T (G_t \Sigma_{t|t-1} G_t^T + R_t)^{-1} \quad (\text{B.17})$$

$$\bar{X}_{t|t} = \bar{X}_{t|t-1} + K_t [Y_t - g_t(\bar{X}_{t|t-1})] \quad (\text{B.18})$$

$$\Sigma_{t|t} = (Id - K_t G_t) \Sigma_{t|t-1} \quad (\text{B.19})$$

## 3. Filtre de Kalman sans parfum

La transformation sans parfum est intégrée au processus. Une étape d'échantillonnage apparaît notamment. L'algorithme devient :

– Échantillonnage déterministe du nuage pondéré :  $\left\{ X_{t-1|t-1}^{(i)}, w_{t-1|t-1}^{(i)} \right\}_{i=1..N}$  à partir de l'estimation précédente  $\bar{X}_{t-1|t-1}$ ,  $\Sigma_{t-1|t-1}$

– Étape de prédiction

$$X_{t|t-1}^{(i)} = f_t \left( X_{t-1|t-1}^{(i)} \right) \quad (\text{B.20})$$

$$\bar{X}_{t|t-1} = \sum_{i=1}^{N_{point}} w_{t-1|t-1}^{(i)} X_{t|t-1}^{(i)} \quad (\text{B.21})$$

$$\Sigma_{t|t-1} = \sum_{i=1}^{N_{point}} w_{t-1|t-1}^{(i)} \left( X_{t|t-1}^{(i)} - \bar{X}_{t|t-1} \right) \left( X_{t|t-1}^{(i)} - \bar{X}_{t|t-1} \right)^T + Q_t \quad (\text{B.22})$$

$$Y_{t|t-1}^{(i)} = g_t \left( X_{t-1|t-1}^{(i)} \right) \quad (\text{B.23})$$

$$\bar{Y}_{t|t-1} = \sum_{i=1}^{N_{point}} w_{t-1|t-1}^{(i)} Y_{t|t-1}^{(i)} \quad (\text{B.24})$$

– Étape de correction

$$\Sigma_{YY} = \sum_{i=1}^{N_{point}} w_{t-1|t-1}^{(i)} \left( Y_{t|t-1}^{(i)} - \bar{Y}_{t|t-1} \right) \left( Y_{t|t-1}^{(i)} - \bar{Y}_{t|t-1} \right)^T + R_t \quad (\text{B.25})$$

$$\Sigma_{XY} = \sum_{i=1}^{N_{point}} w_{t-1|t-1}^{(i)} \left( X_{t|t-1}^{(i)} - \bar{X}_{t|t-1} \right) \left( Y_{t|t-1}^{(i)} - \bar{Y}_{t|t-1} \right)^T \quad (\text{B.26})$$

$$K_t = \Sigma_{XY} \Sigma_{YY}^{-1} \quad (\text{B.27})$$

$$\bar{X}_{t|t} = \bar{X}_{t|t-1} + K_t [Y_t - \bar{Y}_{t|t-1}] \quad (\text{B.28})$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - K_t \Sigma_{YY} K_t^T \quad (\text{B.29})$$



# Annexe C

## Preuves

Cet annexe compile des preuves relatives à la section 4.3.4 et donne quelques résultats supplémentaires par rapport aux propriétés des règles de combinaison liées aux sous-ensembles.

### C.1 Preuve de la proposition 1

Soit  $\odot$  une règle de combinaison directe. Si  $\odot$  est commutative, alors  $\forall A \subseteq \Omega, \forall m_1, m_2 \in \mathfrak{B}^\Omega$  :

$$m_1 \odot m_2 (A) = m_2 \odot m_1 (A) \quad (\text{C.1})$$

$$\sum_{B, C \subseteq \Omega} \psi_A (B, C) m_1 (B) m_2 (C) = \sum_{B', C' \subseteq \Omega} \psi'_A (B', C') m_1 (B') m_2 (C') \quad (\text{C.2})$$

Les indices (relatifs à l'index des sous-ensembles) peuvent être redéfinis de sorte à obtenir :

$$\sum_{B, C \subseteq \Omega} (\psi_A (B, C) - \psi_A (C, B)) m_1 (B) m_2 (C) = 0 \quad (\text{C.3})$$

Comme ce résultat est valable  $\forall m_1, m_2 \in \mathfrak{B}^\Omega$ , il est toujours possible de choisir  $m_1 = B^0$  et  $m_2 = C^0$ , conséquemment  $\forall A, B, C \subseteq \Omega$ , il vient :

$$\psi_A (B, C) - \psi_A (C, B) = 0 \quad (\text{C.4})$$

$$\psi_A (B, C) = \psi_A (C, B) \quad (\text{C.5})$$

Réciproquement, la preuve est triviale :

$$m_1 \odot m_2 (A) = \sum_{B, C \subseteq \Omega} \psi_A (B, C) m_1 (B) m_2 (C) \quad (\text{C.6})$$

Par hypothèse  $\psi_A (B, C) = \psi_A (C, B)$ , d'où :

$$m_1 \odot m_2 (A) = \sum_{B, C \subseteq \Omega} \psi_A (C, B) m_1 (B) m_2 (C) \quad (\text{C.7})$$

$$m_1 \odot m_2 (A) = \sum_{B, C \subseteq \Omega} \psi_A (C, B) m_1 (B) m_2 (C) \quad (\text{C.8})$$

$$m_1 \odot m_2 (A) = m_2 \odot m_1 (A) \quad (\text{C.9})$$

et donc  $\odot$  est commutative.

## C.2 Preuve de la proposition 2

L'écriture de cette preuve est proche de celle de la précédente mais les notations sont plus complexes.

Si  $\odot$  est associative, alors  $\forall A \subseteq \Omega, \forall m_1, m_2, m_3 \in \mathfrak{B}^\Omega$  :

$$\begin{aligned}
(m_1 \odot m_2) \odot m_3 (A) &= m_1 \odot (m_2 \odot m_3) (A) \\
\Leftrightarrow \sum_{B,C} \psi_A(B,C) (m_1 \odot m_2) (B) m_3 (C) &= \\
&= \sum_{B',C'} \psi_A(B',C') m_1 (B') (m_2 \odot m_3) (C') \\
\Leftrightarrow \sum_{B,C} \psi_A(B,C) m_3 (C) \sum_{D,E} \psi_B(D,E) m_1 (D) m_2 (E) &= \\
&= \sum_{B',C'} \psi_A(B',C') m_1 (B') \sum_{D',E'} \psi_{C'}(D',E') m_2 (D') m_3 (E') \\
\Leftrightarrow \sum_{B,C,D,E} \psi_A(B,C) \psi_B(D,E) m_1 (D) m_2 (E) m_3 (C) &= \\
&= \sum_{B',C',D',E'} \psi_A(B',C') \psi_{C'}(D',E') m_1 (B') m_2 (D') m_3 (E') \quad (C.10)
\end{aligned}$$

Il est alors possible de réécrire les indices de sorte à obtenir :

$$\sum_{B,C,D,E} (\psi_A(B,C) \psi_B(D,E) - \psi_A(D,B) \psi_B(E,C)) m_1 (D) m_2 (E) m_3 (C) = 0 \quad (C.11)$$

Comme ce résultat est valable  $\forall m_1, m_2, m_3 \in \mathfrak{B}^\Omega$ , il est toujours possible de choisir  $m_1 = D^0$ ,  $m_2 = E^0$  et  $m_3 = C^0$ , conséquemment  $\forall A, C, D, E \subseteq \Omega$  il vient :

$$\sum_{B \subseteq \Omega} \psi_A(B,C) \psi_B(D,E) - \psi_A(D,B) \psi_B(E,C) = 0 \quad (C.12)$$

$$\sum_{B \subseteq \Omega} \psi_A(B,C) \psi_B(D,E) = \sum_{B \subseteq \Omega} \psi_A(D,B) \psi_B(E,C) \quad (C.13)$$

Réciproquement,  $\forall A \subseteq \Omega$ , l'expression suivante peut se développer :

$$(m_1 \odot m_2) \odot m_3 (A) = \sum_{B,C} \psi_A(B,C) (m_1 \odot m_2) (B) m_3 (C) \quad (C.14)$$

$$= \sum_{B,C} \psi_A(B,C) m_3 (C) \sum_{D,E} \psi_B(D,E) m_1 (D) m_2 (E) \quad (C.15)$$

$$= \sum_{B,C,D,E} \psi_A(B,C) \psi_B(D,E) m_1 (D) m_2 (E) m_3 (C) \quad (C.16)$$

$$= \sum_{C,D,E} m_1 (D) m_2 (E) m_3 (C) \sum_B \psi_A(B,C) \psi_B(D,E) \quad (C.17)$$

Or, par hypothèse,  $\forall C, D, E \subseteq \Omega$  :

$$\sum_{B \subseteq \Omega} \psi_A(B,C) \psi_B(D,E) = \sum_{B \subseteq \Omega} \psi_A(D,B) \psi_B(E,C) \quad (C.18)$$

Alors, il vient

$$= \sum_{C,D,E} m_1(D) m_2(E) m_3(C) \sum_B \psi_A(D, B) \psi_B(E, C) \quad (\text{C.19})$$

$$= \sum_{B,C,D,E} \psi_A(D, B) \psi_B(E, C) m_1(D) m_2(E) m_3(C) \quad (\text{C.20})$$

$$= \sum_{B,D} \psi_A(B, C) m_1(D) \sum_{C,E} \psi_B(E, C) m_2(E) m_3(C) \quad (\text{C.21})$$

$$= \sum_{B,D} \psi_A(B, C) m_1(D) (m_2 \odot m_3)(B) \quad (\text{C.22})$$

$$= m_1 \odot (m_2 \odot m_3)(A) \quad (\text{C.23})$$

et donc  $\odot$  est associative.

### C.3 Preuve de la proposition 3

Si  $A$  est un sous-ensemble absorbant pour  $\odot$ , alors  $\forall m_1, m_2 \in \mathfrak{B}^\Omega$

$$\begin{aligned} & \Leftrightarrow m_1 \odot m_2(A) \geq m_1(A) \\ \Leftrightarrow & \sum_{B,C \subseteq \Omega} \psi_A(B, C) m_1(B) m_2(C) \geq m_1(A) \end{aligned}$$

Soit  $D$  un sous-ensemble de  $\Omega$ , alors l'expression peut être réécrite sous la forme :

$$\psi_A(A, D) m_1(A) m_2(D) + \sum_{B,C/(B,C) \neq (A,D)} \psi_A(B, C) m_1(B) m_2(C) \geq m_1(A)$$

En choisissant  $m_1 = A^0$  et  $m_2 = D^0$ , il vient :

$$\psi_A(A, D) \geq 1$$

ce qui implique  $\psi_A(A, D) = 1$ .

Réciproquement, si  $\forall D \subseteq \Omega, \psi_A(A, D) = 1$ , alors

$$\begin{aligned} m_1 \odot m_2(A) &= \sum_{B,C \subseteq \Omega} \psi_A(B, C) m_1(B) m_2(C) \\ m_1 \odot m_2(A) &= \sum_{C \subseteq \Omega} \psi_A(A, C) m_1(A) m_2(C) + \sum_{B,C/B \neq A} \psi_A(B, C) m_1(B) m_2(C) \\ m_1 \odot m_2(A) &= m_1(A) \sum_C m_2(C) + \sum_{B,C/B \neq A} \psi_A(B, C) m_1(B) m_2(C) \\ m_1 \odot m_2(A) &= m_1(A) + \sum_{B,C/B \neq A} \psi_A(B, C) m_1(B) m_2(C) \end{aligned} \quad (\text{C.25})$$

ce qui implique  $m_1 \odot m_2(A) \geq m_1(A)$ ,  $A$  est absorbant.

## C.4 Preuve de la proposition 4

Soit  $\odot$  une règle de combinaison directe commutative. Si  $A$  est un sous-ensemble de rejet, alors  $\forall m_1, m_2 \in \mathfrak{B}^\Omega$  :

$$m_1 \odot m_2 (A) \leq m_1 (A) \quad (\text{C.26})$$

$$\Leftrightarrow \sum_{B, C \subseteq \Omega} \psi_A(B, C) m_1(B) m_2(C) \leq m_1(A) \quad (\text{C.27})$$

$$\Leftrightarrow m_1(A) \sum_{C \subseteq \Omega} \psi_A(A, C) m_2(C) + \sum_{\substack{B, C \subseteq \Omega \\ B \neq A}} \psi_A(B, C) m_1(B) m_2(C) \leq m_1(A) \quad (\text{C.28})$$

Cette expression est vraie  $\forall m_1, m_2 \in \mathfrak{B}^\Omega$ . En choisissant un cas particulier pour  $m_1$  de sorte que  $m_1(A) = 0, \forall B \neq A, m_1(B) = \frac{1}{\#(2^\Omega) - 1} = \lambda_1 > 0$  et pour  $m_2$  de sorte que  $\forall C \subseteq \Omega, m_2(C) = \frac{1}{\#(2^\Omega)} = \lambda_2 > 0$ , il vient :

$$\lambda_1 \lambda_2 \sum_{\substack{B, C \subseteq \Omega \\ B \neq A}} \psi_A(B, C) \leq 0 \quad (\text{C.29})$$

Comme il s'agit d'une somme de termes positifs, il vient :  $\forall B, C \subseteq \Omega, B \neq A, \psi_A(B, C) = 0$ . En appliquant ce résultat pour  $C = A$ , et par commutativité, il vient de plus :

$$\forall B \neq A, \psi_A(B, A) = 0 \Leftrightarrow \psi_A(A, B) = 0 \quad (\text{C.30})$$

Soit au final :

$$\forall (B, C) \neq (A, A) \subseteq \Omega, \psi_A(B, C) = 0 \quad (\text{C.31})$$

Réciproquement, si  $A$  est tel que  $\forall (B, C) \neq (A, A), \psi(B, C) = 0$ , alors  $\forall m_1, m_2 \in \mathfrak{B}^\Omega$

$$m_1 \odot m_2 (A) = \sum_{B, C \subseteq \Omega} \psi_A(B, C) m_1(B) m_2(C) \quad (\text{C.32})$$

$$\Leftrightarrow m_1 \odot m_2 (A) = m_1(A) m_2(A) \quad (\text{C.33})$$

$$\Rightarrow m_1 \odot m_2 (A) \leq m_1(A) \quad (\text{C.34})$$

$A$  est de rejet pour  $\odot$ .

## C.5 Contraintes liées à l'accumulation de propriétés

### C.5.1 Associativité, commutativité, consistance faible et règle directe

Soit  $\odot$  une règle directe de combinaison associative commutative et faiblement consistante. Par associativité :  $\forall A, C, D, E \subseteq \Omega$

$$\sum_{B \subseteq \Omega} \psi_A(B, C) \psi_B(D, E) = \sum_{B \subseteq \Omega} \psi_A(D, B) \psi_B(E, C) \quad (\text{C.35})$$

En choisissant  $E = C$ , l'expression s'écrit :

$$\Leftrightarrow \sum_{B \subseteq \Omega} \psi_A(B, C) \psi_B(D, C) = \sum_{B \subseteq \Omega} \psi_A(D, B) \psi_B(C, C) \quad (\text{C.36})$$

Par hypothèse de consistance faible  $\forall B \neq C \psi_B(C, C) = 0$  et  $\psi_C(C, C) = 1$ , d'où :

$$\sum_{B \subseteq \Omega} \psi_A(B, C) \psi_B(D, C) = \psi_A(D, C) \quad (\text{C.37})$$

$$\Leftrightarrow \psi_A(A, C) \psi_A(D, C) + \sum_{B \neq A \subseteq \Omega} \psi_A(B, C) \psi_B(D, C) = \psi_A(D, C) \quad (\text{C.38})$$

En choisissant  $C = A$ , et en utilisant le fait que  $\psi_A(A, A) = 1$  l'expression s'écrit :

$$\Leftrightarrow \psi_A(D, A) + \sum_{B \neq A \subseteq \Omega} \psi_A(B, A) \psi_B(D, A) = \psi_A(D, A) \quad (\text{C.39})$$

$$\sum_{B \neq A \subseteq \Omega} \psi_A(B, A) \psi_B(D, A) = 0 \quad (\text{C.40})$$

Comme l'expression précédente est une somme de termes positifs, il vient :

$$\forall A, D, B \neq A \psi_A(B, A) \psi_B(D, A) = 0 \quad (\text{C.41})$$

Le résultat précédent exprime une contrainte sur les fonctions  $\psi_A$  si toutes les propriétés évoquées sont souhaitées pour la règle associée aux fonctions  $\psi_A$ .

En rajoutant la propriété d'impartialité, le sous-ensemble  $B$  ne peut être de rejet. Il existe donc un couple  $(D', A') \neq (B, B)$  tel que  $\psi_B(D', A') \neq 0$ . Par commutativité, il est possible de choisir  $A = A'$  et  $D = D'$  et spécifiant que  $A' \neq B$ . D'après la contrainte précédente cela impose :

$$\forall B \subseteq \Omega, \exists A' \neq B | \psi'_A(B, A') = 0 \quad (\text{C.42})$$

Cette contrainte renforce le fait  $A'$  ne puisse être absorbant. La non-absorbance impose  $\psi'_A(B, A') < 1$ .

### C.5.2 Idempotence, commutativité et règle directe

**Proposition :** *Il n'existe pas de règle de combinaison directe commutative idempotente.*

**Preuve :**

Soit une règle directe commutative et idempotente. L'idempotence impose que sa forme matricielle correspond à la matrice identité. En conséquence, pour un ensemble  $A$  donné,  $\forall B \subseteq \Omega, \psi_A(A, B) = 1$  et  $\forall C, D \subseteq \Omega, D \neq A, \psi_A(C, D) = 0$ . En utilisant la commutativité sur la première expression et en l'appliquant en particulier pour  $B = B' \neq A$ , il vient :  $\psi_A(B', A) = 1$ . En repartant du fait que  $\psi_A(C, D) = 0$  et en prenant les cas  $C = A$  et  $D = B'$ , il vient :  $\psi_A(B', A) = 0$ . Il y a contradiction, et donc aucune règle directe commutative et idempotente existe.

En conséquence de ce résultat, le fait que les règles de Murphy, conjonctive prudente, et disjonctive hardie ne soient pas directe est justifié.

## C.6 Unicité et liens entre sous-ensembles absorbants et de rejet

**Proposition :** *Si une règle directe commutative  $\odot$  possède un sous-ensemble absorbant  $A$ , alors ce sous-ensemble absorbant est unique.*

**Preuve :**

Comme  $A$  est absorbant,  $\forall B \subseteq \Omega$ ,  $\psi_A(A, B) = 1$ . D'après la définition des fonctions caractéristiques d'une combinaison directe :

$$\sum_{C \subseteq \Omega} \psi_C(A, B) = 1 \quad (\text{C.43})$$

$$\Leftrightarrow \sum_{C \subseteq \Omega, C \neq A} \psi_C(A, B) = 0 \quad (\text{C.44})$$

Or, les termes de cette somme étant tous positifs, il vient  $\forall C \neq A$ ,  $\psi_C(A, B) = 0$ . En particulier, pour  $B = C$ ,  $\psi_C(A, C) = 0$ . Par commutativité,  $\psi_C(C, A) = 0$ , et conséquemment  $C$  ne peut être absorbant.

Le résultat n'a pas d'équivalent pour les sous-ensembles de rejet. Voici un exemple de règle directe commutative possédant plusieurs sous-ensembles de rejet :

$$M_{\odot}^m = \begin{pmatrix} 1 & m(\emptyset) + m(b) + m(\Omega) & m(\emptyset) + m(a) + m(\Omega) & m(\emptyset) + m(a) + m(b) \\ 0 & m(a) & 0 & 0 \\ 0 & 0 & m(b) & 0 \\ 0 & 0 & 0 & m(\Omega) \end{pmatrix} \quad (\text{C.45})$$

De même, la présence d'un élément absorbant n'impose pas la présence d'un élément neutre.

Voici un exemple de règle directe, faiblement consistante et commutative possédant un seul élément absorbant et aucun élément neutre :

$$M_{\odot}^m = \begin{pmatrix} 1 & m(\emptyset) & m(\emptyset) & m(\emptyset) \\ 0 & m(a) + m(\Omega) & 0 & m(a) \\ 0 & m(b) & m(b) + m(a) & 0 \\ 0 & 0 & m(\Omega) & m(\Omega) + m(b) \end{pmatrix} \quad (\text{C.46})$$

Cette règle ne présente, cela dit, presque aucun intérêt tant elle est construite de manière extravagante.

# Bibliographie

- [Abellan 99] J. Abellan, S. Moral. – Completing a total uncertainty measure in Dempster-Shafer theory. *International Journal of General Systems*, 28 :299–314, 1999.
- [Adam 01] S. Adam, J.-M. Ogier, C. Cariou, R. Mullot, J. Gardes, Y. Lecourtier. – Utilisation de la transformée de Fourier-Mellin pour la reconnaissance de formes multi-orientées et multi-échelles : application à l'analyse automatique de documents techniques. *Revue Traitement du Signal*, 18(1) :17–33, 2001.
- [Aldon 98] M.-J. Aldon, S. Bouaziz, V. Cherfaoui, C. Durieu, D. Jolly, A.-M. Jolly-Desodt, T. Maurin, E. M. Mouaddib, J. Pontois, R. Reynaud, M. Rombaut, L. Trassoudaine. – *La perception et la fusion de données pour véhicules intelligents*. – Rapport de recherche, GdR Isis, thématique FUSION, opération thématique 6.3, 1998.
- [Appriou 91] A. Appriou. – Probabilités et incertitude en fusion de données multi-senseurs. *Revue scientifique de la défense*, 11 :27–40, 1991.
- [Appriou 97] A. Appriou. – Multiple signal tracking processes. *Aerospace Science and Technology*, 1(3) :165–178, 1997.
- [Arnaud 04a] E. Arnaud. – *Méthodes de filtrage pour du suivi dans des séquences d'images - Application au suivi de points caractéristiques*. – PhD. Thesis, University of Rennes I, 2004.
- [Arnaud 04b] E. Arnaud, E. Memin. – Optimal importance sampling for tracking in image sequences : application to point tracking. – *European Conf. on Computer Vision 2004*, pp. cd-rom, 2004.
- [Arumpalam 02] M.S. Arumpalam, S. Maskell, N. Gordon, T. Clapp. – A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Trans. on Signal Processing*, 50(2) :174–188, 2002.
- [Aycard 06] O. Aycard, A. Spalanzani, M. Yguel, J. Burlet, N. Du lac, A. De la Fortelle, T. Fraichard, H. Ghorayeb, M. Kais, C. Laugier, C. Laugeau, G. Michel, D. Raulo, B. Steux. – Puvame - new french approach for vulnerable road users safety. – *Intelligent Vehicles symposium*, Tokyo (Japan), june 13-15 2006.
- [Ayer 95] S. Ayer, H. Sawhney. – Layered representation of motion video using robust maximum-likelihood estimation of mixture models and mdl encoding. – *Int. Conf. on Computer Vision*, 1995.
- [Barbu 07] E. Barbu. – *Fouille et classification de graphes, application à la reconnaissance de symboles dans les documents techniques*. – PhD. Thesis, University of Rouen, june 2007.
- [BarShalom 88] Y. Bar-Shalom, T. Fortmann. – Tracking and data association. *Academic press*, 1988.

- [Bashir 06] F. Bashir, F. Poriki. – Performance evaluation of object detection and tracking systems. – *IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, pp. 7–14, New-York, NY (USA), june 2006.
- [Bensrhair 02] A. Bensrhair, M. Bertozzi, A. Broggi, A. Fascioli, S. Mousset, G. Toulminet. – Stereo-vision based feature extraction for vehicle detection. – *IEEE Intelligent Vehicles Symp.*, 2002.
- [Berretti 00] S. Berretti, A.D. Bimbo, P. Pala. – Retrieval by shape similarity with perceptual distance and effective indexing. *IEEE trans. on Multimedia*, 2(4) :225–239, 2000.
- [Besag 86] J.E. Besag. – On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society B*, 48(3) :259–302, 1986.
- [Betke 07] M. Betke, D.E. Hirsh, A. Bagchi, N.I. Hristov, N.C. Makris, , T.H. Kunz. – Tracking large variable numbers of objects in clutter. – *int. conf. on Computer Vision and Pattern Recognition*, 2007.
- [Black 96] M.J. Black, A. Jepson. – Eigentracking : robust matching and tracking of articulated objects using a view-based representation. – *European Conference on Computer Vision*, pp. 329–342, 1996.
- [Blinn 77] J. Blinn. – Models of light reflection for computer synthesized pictures. – *Int.l Conference on Computer Graphics and Interactive Techniques*, pp. 192–198, San Jose (USA), 1977.
- [Bloch 96] I. Bloch. – Information combination operators for data fusion : A comparative review with classification. *IEEE trans. on Systems, Man and Cybernetics*, 26(1) :52–67, 1996.
- [Bloch 08] I. Bloch. – Defining belief functions using mathematical morphology - application to image fusion under imprecision. *Int. Journal of Approximate Reasoning*, 48(2) :437–465, june 2008.
- [Bovik 91] A.C. Bovik. – Analysis of multichannel narrow-band filters for image texture segmentation. *IEEE trans. on Signal Processing*, 39(9) :2025–2043, 1991.
- [Braid 06] D. Braid, A. Broggi, G. Schmiedel. – The terramax autonomous vehicle. *Journal of Field Robotics*, 23(9) :693–708, 2006.
- [Brasnett 06] P. Brasnett, L. Mihaylova, D. Bull, N. Canagarajah. – Sequential monte carlo tracking by fusing multiple cues in video sequences :. *Image and Vision Computing*, 25 :1217–1227, 2006.
- [Bray 07] M. Bray, E. Koller-Meier, N.N. Schraudolph, L. Van Gool. – Fast stochastic optimization for articulated structures tracking. *Image and Vision Computing*, 25 :352–364, 2007.
- [Brochard 01] J. Brochard, M. Khoudeir, B. Augerau. – Invariant feature extraction for 3d texture analysis using the autocorrelation function. *Pattern Recognition Letters*, 22 :759–768, 2001.
- [Broggi 00] A. Broggi, M. Bertozzi, A. Flascioli, C. G. Lo Bianco, A. Piazzini. – Visual perception of obstacles and vehicles for platooning. *IEEE trans. on Intelligent Transportation Systems*, 1(3) :164–176, september 2000.

- 
- [Brèthes 06] L. Brèthes, P. Danès, F. Lerasle. – Stratégie de filtrage particulière pour le suivi visuel de personnes : descriptions et évaluation. – *RFIA 2006*, Tours, France, 2006.
- [Bugeau 07a] A. Bugeau. – *Détection et suivi d'objets en mouvement dans des scènes complexes, application à la surveillance des conducteurs*. – PhD. Thesis, University of Rennes I, 2007.
- [Bugeau 07b] A. Bugeau, P. Perez. – Detection and segmentation of moving objects in highly dynamic scenes. – *IEEE int. conf on Computer Vision and Pattern Recognition*, pp. 1–8, Minneapolis (USA), 2007.
- [Bullock 04] D.J. Bullock, J. S. Zelek. – Real-time tracking for visual interface applications in cluttered and occluding situations. *Image and Vision Computing*, 22 :1083–1091, 2004.
- [Campbell 68] F.W. Campbell, J.G. Robson. – Application of fourier analysis to the visibility of gratings. *Journal Physiol*, 197 :551–566, 1968.
- [Canny 86] J. Canny. – A computational approach to edge detection. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 8(6) :679–698, 1986.
- [Caron 07a] F. Caron, M. Davy, E. Duflos, P. Vanheeghe. – Particle filtering for multi-sensor data fusion with switching observation models : applications to land vehicle positioning. *IEEE Trans. on Signal Processing*, 55(6) :2703–2719, june 2007.
- [Caron 07b] F. Caron, B. Ristic, E. Duflos, P. Vanheeghe. – Least committed basic belief density induced by a multivariate gaussian : formulation with applications. *Int. Journal of Approximate Reasoning*, 48(2) :419–436, 2007.
- [Caselles 93] V. Caselles, F. Catte, T. Coll, F. Dibos. – A geometric model for active contours. *Numerische Mathematik*, 66 :1–31, 1993.
- [Cavallaro 02] A. Cavallaro, O. Steiger, T. Ebrahimi. – Multiple video object tracking in complex scenes. – *Multimedia'02*, Juan-les-Pins, France, 2002.
- [Cerny 85] V. Cerny. – Thermodynamical approach to the traveling salesman problem : An efficient simulation algorithm. *Journal of Opt. Theory Appl.*, 45 :41–51, 1985.
- [Cham 99] T.J. Cham, J. Regh. – A multiple hypotheses approach to figure tracking. – *int conf. on Computer Vision and Pattern Recognition*, pp. 239–245, 1999.
- [Chang 99] P. Chang, J. Krumm. – Object recognition with color cooccurrence histograms. – *IEEE int. conf. on Computer Vision and Pattern Recognition*, Fort Collins, 1999.
- [Cheng 06] H. Cheng, N. Zheng, C. Sun. – Boosted gabor features applied to vehicle detection. – *IEEE Int. Conf. on Pattern Recognition*, vol. 1, Hong Kong (China), 2006.
- [Choi 96] J.G. Choi, S.D. Kim. – Multi-stage segmentation of optical flow field. *Signal Processing*, 54 :109–118, 1996.
- [Chou 90] P. Chou, C. Brown. – The theory and practice of bayesian image labeling. *International Journal of Computer Vision*, 4 :184–210, 1990.

- [Cinque 99] L. Cinque, S. Levialdi, A. Pellicano, K. A. Olsen. – Color-based image retrieval using spatial chromatic histograms. *IEEE Multimedia Systems*, 2 :969–973, 1999.
- [Collins 03] R.T. Collins, Y. Liu. – On-line selection of discriminative tracking features. – *Int. Conf. on Computer Vision*, 2003.
- [Comaniciu 00] D. Comaniciu, V. Ramesh, P. Meer. – Real time tracking of non rigid objects using mean shift. – *int. conf. on Computer Vision and Pattern Recognition*, pp. 142–149, Hilton Head Island, 2000.
- [Couprie 07] M. Couprie, D. Coeurjolly, R. Zrou. – Discrete bisector function and euclidean skeleton in 2d and 3d. *Image and Vision Computing*, 25(10) :1519–1698, 2007.
- [Cox 93] I. Cox. – A review of statistical data association techniques for motion correspondence. *Int. Journal of Computer Vision*, 10(1) :53–65, 1993.
- [Crisan 99] D. Crisan, P. Del Moral, T.J. Lyons. – Discrete filtering using branching and interacting monte carlo methods. *Markov Process and related fields*, 5(3) :293–318, 1999.
- [Czyz 05] J. Czyz, B. Ristic, B. Macq. – A color-based particle filter for joint detection and tracking of multiple objects. – *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, 2005.
- [Dai 08] S. Dai, Y. Wu. – Motion from blur. – *int. conf. on Computer Vision and Pattern Recognition*, 2008.
- [Daniel 03] M. Daniel. – Associativity in combination of belief functions ; a derivation of minc combination. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, 7(5) :288–296, 2003.
- [Darrel 91] T. Darrel, A. Pentland. – Robust estimation of a multi-layered motion representation. – *IEEE workshop on Visual Motion*, 1991.
- [Davies 97] E.R. Davies. – Machine vision : Theory, algorithms practicalities. – Academic press, New York, 1997. pp. 171-191.
- [Debain 00] C. Debain, T. Chateau, M. Berducat, P. Martinet, P. Bonton. – A guidance-assistance system for agricultural vehicles. *Computer and electronics for agriculture*, 25 :29–51, 2000.
- [Delalandre 05] M. Delalandre. – *Analyse des Documents Graphiques : une Approche par Reconstruction d’Objets*. – France, Phd thesis, University of Rouen, 2005.
- [Delmotte 95] F. Delmotte, L. Dubois, A.-M. Desodt, P. Borne. – Using trust in uncertainty theories. *Information and Systems Engineering*, 1 :303–314, 1995.
- [Delmotte 96] F. Delmotte, L. Dubois, P. Borne. – Context-dependent trust in data fusion within the possibility theory. – *IEEE Int. Conf. on Systems, Man, and Cybernetics*, pp. 538–543, Beijing (China), 14-17 Oct 1996.
- [Demonceaux 04] C. Demonceaux, D. Kachi-Akkouche. – Robust obstacle detection with monocular vision based on motion analysis. – *Intelligent Vehicles Symposium*, pp. 527 – 532, Parma (Italy), 2004.
- [Dempster 67] A.P. Dempster. – Upper and lower probabilities induced by a multiple valued mapping. *Annals of Mathematical Statistics*, 38 :325–339, 1967.

- 
- [Dempster 68] A.P. Dempster. – A generalization of bayesian inference. *Journal of Royal Statistical Society*, B 30 :205–247, 1968.
- [Dempster 77] A.P. Dempster, N.M. Laird, D.B. Rubin. – Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society series B*, 39 :1–38, 1977.
- [Denis 07] P. Denis, P. Carre, C. Fernandez-Maloigne. – Spatial and spectral quaternionic approaches for colour images. *Computer Vision and Visual Understanding*, 107 :7487, 2007.
- [Denoeux 95] T. Denoeux. – A k-nearest neighbour classification rule based on dempster-shafer theory. *IEEE trans. on Systems Man and Cybernetics*, 25(5) :804–813, 1995.
- [Denoeux 99] T. Denoeux. – Reasoning with imprecise belief structures. *Int. Journal of Approximate Reasoning*, 20 :79–111, 1999.
- [Denoeux 08] T. Denoeux. – Conjunctive and disjunctive combination of belief functions induced by non-distinct bodies of evidence. *Artificial Intelligence*, 172(2-3) :234–264, february 2008.
- [Deriche 90] R. Deriche. – Fast algorithms for low-level vision. *Pattern Analysis and Machine Intelligence*, 12(1) :78–87, 1990.
- [Dezert 02] J. Dezert. – Foundations of a new theory of plausible and paradoxical reasoning. *Information Security*, 9 :13–57, 2002.
- [Diaz 06] J. Diaz, M. Rifqi, B. Bouchon-Meunier. – A similarity measure between basic belief assignments. – *int. conf. on Information Fusion*, pp. 1–6, Florence (Italy), july 2006.
- [Doucet 97] A. Doucet. – *Monte Carlo Methods for Bayesian Estimation of Hidden Markov Models, Application to Radiation Signals*. – PhD. Thesis, University of Paris-sud, Orsay, 1997.
- [Doucet 98] A. Doucet. – *On sequential Simulation-based Methods for Bayesian Filtering*. – Rapport de Recherche n CUED/F-INFENG/TR. 310, Cambridge Univsersity, 1998.
- [Doucet 00] A. Doucet, N. de Freitas, K. Murphy, S. Russell. – Rao-blackwellised particle filtering for dynamic bayesian networks. – *Conference on Uncertainty in Artificial Intelligence*, 2000.
- [Dubois 83] D. Dubois, H. Prade. – Unfair coins and necessity measure : toward a possibilistic interpretation of histograms. *Fuzzy Sets and Systems*, 10 :15–22, 1983.
- [Dubois 85] D. Dubois, H. Prade. – A note on measures of specificity for fuzzy sets. *International Journal of General Systems*, 10 :279283, 1985.
- [Dubois 86a] D. Dubois, H. Prade. – On the unicity of dempster rule of combination. *int. journal of Intelligent Systems*, 1 :133–142, 1986.
- [Dubois 86b] D. Dubois, H. Prade. – A set-theoretic view of belief functions : logical operations and approximatons by fuzzy sets. *Int. Journal of General Systems*, 12(3) :193–226, 1986.

- [Dubois 88] D. Dubois, H. Prade. – Representation and combination of uncertainty with belief functions and possibility measures. *Comput. Intell.*, 4 :244–264, 1988.
- [Dubois 92] D. Dubois, H. Prade. – *Reliability Data Collection and Analysis*, chap. On the combination of evidence in various mathematical frameworks, pp. 213–241. – 1992.
- [Dubois 04] D. Dubois, H. Prade. – On the use of aggregation operations in information fusion processes. *Fuzzy Sets and Systems*, 142 :143–161, 2004.
- [Elgammal 00] A. Elgammal, D. Harwood, L. Davis. – Non-parametric model for background subtraction. – *Euro. Conf. on Computer Vision*, vol. 2, pp. 661–667, 2000.
- [Elouedi 04] Z. Elouedi, K. Mellouli, P. Smets. – Assessing sensor reliability for multisensor data fusion within the transferable belief model. *IEEE trans. on Systems, Man, and Cybernetics Part B : Cybernetics*, 34(1) :782–787, 2004.
- [Fabre 98] S. Fabre, A. Appriou, X. Briottet. – Introduction of contextual information in pixel fusion. – *int. conf. on Information Fusion*, pp. 601–609, Las Vegas, NV (USA), July 69 1998.
- [Faux 06] F. Faux, F. Luthon. – Robust face tracking using color dempster-shafer fusion and particle filter. – *int. conf. Information Fusion, FUSION*, 2006.
- [Fefilatyev 06] S. Fefilatyev, V. Smarodzinava, L. O. Hall, D. B. Goldgof. – Horizon detection using machine learning techniques. – *International Conference on Machine Learning and Applications*, pp. 17–21, 2006.
- [Feris 04] R. S. Feris, V. Krueger, R. M. Cesar Jr. – A wavelet subspace method for real-time face tracking. *Real Time Imaging*, 10 :339–350, 2004.
- [Fleet 90] D. Fleet, A. Jepson. – Computation of component image velocity from local phase information. *Int. Journal of Computer Vision*, 5(1) :77–104, 1990.
- [Florea 06] M. C. Florea, J. Dezert, P. Valin, F. Smarandache, A.-L. Jouselme. – Adaptive combination rule and proportional conflict redistribution rule for information fusion. – *COGIS-06*, Paris (France), 2006.
- [Freeman 61] H. Freeman. – On the encoding of arbitrary geometric configurations. *IRE trans. on Electron. Comput.*, 10 :260–268, 1961.
- [Fu 74] K.S. Fu. – Syntactic methods in pattern recognition. *Academic press, New York*, 1974.
- [Fu 82] K. S. Fu. – A syntactic approach to texture analysis. *Syntactic Pattern Recognition and Applications*, chap. 12, Prentice Hall, pp. 445–480, 1982.
- [Gagalowicz 83] A. Gagalowicz. – *Vers un modèle de textures*. – PhD. Thesis, University of Paris VI, 1983.
- [Galloway 75] M. M. Galloway. – Texture analysis using gray lengths. *Computer Graphics and Image Processing*, 4 :172–179, 1975.
- [Geman 84] S. Geman, D. Geman. – Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 6(6) :721–741, 1984.

- [Georgescu 04] B. Georgescu, P. Meer. – Point matching under large image deformations and illumination changes. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 26(6) :674–688, 2004.
- [Gerard 00] P. Gerard, A. Gagalowicz. – Three dimensional model-based tracking using texture learning and matching. *Pattern recognition letters*, 21 :1095–1103, 2000.
- [Gidel 08] S. Gidel, C. Blanc, T. Chateau, P. Checchin, L. Trassoudaine. – Non-parametric data association for particle filter based multi-object tracking : application to multi-pedestrian tracking. – *Intelligent Vehicles Symposium*, Eindhoven (Netherlands), june 4-6 2008.
- [Gordon 93] N.J. Gordon, A.F.M. Smith, D.J. Salmond. – A novel approach to nonlinear/non-gaussian bayesian state estimation. *IEEE proceedings-F*, 40(2) :107–113, 1993.
- [Gorskey 90] W.I. Gorskey, R. Mehrotra. – Index-based object recognition in pictorial data management. *Computer Vision Graphics Image Process*, 52 :416–436, 1990.
- [Goshtasby 85] A. Goshtasby. – Description and discrimination of planar shapes using shape matrices. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 7 :738–743, 1985.
- [Graciano 07] A. B. V. Graciano, R. M. Cesar-Jr., I. Bloch. – Graph-based object tracking using structural pattern recognition. – *Computer Graphics and Image Processing, 2007. SIBGRAPI 2007. XX Brazilian Symposium on*, pp. 179–186, Minas Gerais, Brazil, 2007.
- [Grimson 98] Y. Grimson, C. Stauffer, R. Romano, L. Lee. – Using adaptive tracking to classify and monitor activities in a site. – *int. conf. Computer Vision and Pattern Recognition*, 1998.
- [Ha 04] J. Ha, C. Alvinot, G. Pryo, M. Niethammer, E. Johnson, A. Tannenbaum. – Active contours and optical flow for automatic tracking of flying vehicles. – *American Control Conference*, pp. 3441–3446, 2004.
- [Haenni 02] R. Haenni. – Are alternatives to dempster’s rule of combination real alternatives? *Information Fusion*, 3 :237–239, 2002.
- [Haenni 05] R. Haenni. – Shedding new light on zadeh’s criticism of dempster’s rule of combination. – *int. conf. on Information Fusion*, Philadelphia (USA), 2005.
- [Hammersle 54] J.M. Hammersle, K.W. Morton. – Poor man’s monte carlo. *Journal of the Royal Statistics Society*, 16 :23–38, 1954.
- [Hammersley 71] J. Hammersley, P. Clifford. – Markov fields on finite graphs and lattices. *Unpublished manuscript*, 1971.
- [Han 05] B. Han, Y. Zhu, D. Comaniciu, L. Davis. – Kernel-based bayesian filtering for object tracking. – *int. conf. on Computer Vision and Pattern Recognition*, San Diego (USA), 2005.
- [Han 07] B. Han, S.-W. Joo, L. S. Davis. – Probabilistic fusion tracking using mixture kernel-bsaed bayesian filtering. – *IEEE Int. Conf. on Computer Vision*, Rio, september 2007.

- [Handmann 00] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, W.V. Seelen. – An image processing system for driver assistance. *Image and Vision Computing*, 18 :367376, 2000.
- [Haralick 73] R. M. Haralick, K. Shanmugan, I. Dinstein. – Textural features for image classification. *IEEE Trans. on Systems, Man and Cybernetics*, 3(6) :610–621, 1973.
- [Harmanec 96] D. Harmanec, G. Resconi, G. J. Klir, Y Pan. – On the computation of uncertainty measure in dempster-shafer theory. *International Journal of General Systems*, 25 :153–163, 1996.
- [Harris 88] C. Harris, M. Stephens. – A combined corner and edge detector. – *Proceedings of The Fourth Alvey Vision Conference*, pp. 147–151, Manchester (UK), 1988.
- [He 07] L.-P. He, F.-Z. Qu. – A comparative study of compatibility and transformation between probability and possibility. – *IEEE int. conf. on Fuzzy Systems and Knowledge Discovery*, vol. 1, pp. 12–16, 2007.
- [Heckbert 82] P. Heckbert. – Color image quantization for frame buffer display. *Computer Graphics*, 16(3) :297–307, 1982.
- [Heeger 88] D. Heeger. – Optical flow using spatio-temporal filters. *Int. Journal of Computer Vision*, 1(4) :279–302, 1988.
- [Hii 06] A.J.H. Hii, C.E. Hann, J.G. Chase, E.E.W. Van Houten. – Fast normalized cross correlation for motion tracking using basis functions. *Computer Methods and Programs in Biomedicine*, 82(2) :144–156, 2006.
- [Hoffmann 04] C. Hoffmann, T. Dang, C. Stiller. – Vehicle detection fusing 2d visual features. – *IEEE Intelligent Vehicles symposium*, Parma (Italy), june 2004.
- [Horn 81] B. Horn, B. Schunck. – Determining optical flow. *Artificial Intelligence*, 17(1-3) :185–203, 1981.
- [Hu 62] M.K. Hu. – Visual pattern recognition by moment invariants. *IEE trans. on Information Theory*, 8 :179–187, 1962.
- [Huang 02] Y. Huang, T. S. Huang, H. Niemann. – A region-based method for model-free object tracking. – *IEEE Int Conf. on Pattern Recognition*, pp. 592–595, 2002.
- [Huber 64] P.J. Huber. – Robust estimation of a location parameter. *Annals of Mathematical Statistics*, 35 :73101, 1964.
- [Huttenlocher 93] D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge. – Comparing images using the hausdorff distance. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 15(9) :850–863, 1993.
- [Iivarinen 96] J. Iivarinen, A. Visa. – Shape recognition of irregular objects. – *Intelligent robots and computer vision XV : Algorithms, Techniques, Active vision and material handling*, pp. 25–32, 1996.
- [Inagaki 91] T. Inagaki. – Interdependence between safety-control policy and multiple-sensor schemes via dempster-shafer theory. *IEEE trans. on Reliability*, 40(2) :182–188, 1991.

- 
- [Isard 98a] M. Isard, A. Blake. – Condensation-conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1) :5–28, 1998.
- [Isard 98b] M. Isard, A. Blake. – Icondensation : Unifying low-level and high-level tracking in a stochastic framework. *Lecture Notes in Computer Science*, 1406 :893–908, 1998.
- [Isard 01] M. Isard, J. Mc Cormick. – Bramble : A bayesian multiple-blob tracker. – *IEEE Int. Conf on Computer Vision*, pp. 34–41, 2001.
- [Izquierdo 04] D. Izquierdo. – *Contribution au développement d'une architecture générique dédiée au suivi d'objets en télésurveillance : application au suivi de véhicules et de visages*. – PhD. Thesis, University of Bordeaux I, 2004.
- [Jacquot 05] A. Jacquot, P Sturm, O Ruch. – Adaptive tracking of non-rigid objects based on color histograms and automatic parameter selection. – *IEEE workshop on Motion and Video Computing*, pp. 103 – 109, 2005.
- [Jain 79] R. Jain, H.H. Nagel. – On the analysis of accumulative difference pictures from image sequence of real world scenes. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 1(2) :206–214, 1979.
- [Jepson 93] A. Jepson, M. Black. – *Mixture models for optical flow*. – Rapport de recherche, Tech. Report, Res. in Biol. and Comp. Vision, Dept. of Comp. Sci., Univ. of Toronto, RBCV-TR-93-44, 1993.
- [Jepson 03] A. D. Jepson, D. J. Fleet, T. F. El-Maraghi. – Robust on-line appearance models for visual tracking. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 25 :1296–1311, 2003.
- [Josang 03] A. Josang, M. Daniel, P. Vannoorenberghe. – Strategies for combining conflicting dogmatic beliefs. – *int. conf. on Information Fusion*, pp. 1133–1140, Cairns (Australia), 2003.
- [Josso 05] B. Josso, D. R. Burton, M. J. Lalor. – Texture orientation and anisotropy calculation by fourier transform and principal component analysis. *Mechanical systems and Signal processing*, 19 :1152–1161, 2005.
- [Jousselme 01] A.-L. Jousselme, D. Grenier, E. Bosse. – A new distance between two bodies of evidence. *Information Fusion*, 2(2) :91–101, 2001.
- [Julier 97] S. Julier, J. Uhlmann. – A new extension of the kalman filter to non-linear systems. – *int. symposium on Aerospace/Defense Sensing Simulation and Control*, 1997.
- [Jung 06] C.R. Jung. – Unsupervised multiscale segmentation of color images. *Pattern Recognition Letters*, 28(4) :523–533, 2006.
- [Kahn 05] Z. Kahn, T. Balch, F. Dellaert. – Mcmc-based particle filtering for tracking a variable number of interacting targets. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 27(11) :1805–1819, 2005.
- [Kalman 60] R. Kalman. – A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82 :33–45, 1960.
- [Kamijo 02] S. Kamijo, M. Sakauchi. – Illumination invariant and occlusion robust vehicle tracking by spatio-temporal mrf model. – *9th World Congress on ITS*, Chicago (USA), 2002.

- [Kang 04] J. Kang, I. Cohen, G. Medioni. – Object reacquisition using invariant appearance model. – *IEEE Int. Conf on Pattern Recognition (ICPR04)*, 2004.
- [Kass 88] M. Kass, A. Witkin, D. Terzopoulos. – Snakes : Active contour models. *International Journal of Computer Vision*, 1(4) :321–331, 1988.
- [Kenna 07] S.J. Mc Kenna, H. Hait-Charif. – Tracking human motion using auxiliary particle filters and iterated likelihood weighting. *Image and Vision Computing*, 25 :852–862, 2007.
- [Kharbouche 04] S. Kharbouche, P. Vannoorenberghe. – Un filtre crédal pour la segmentation d’images. – *Troisième Conférence Internationale Francophone d’Automatique, CIFA’2004*, Douz (Tunisia), 2004.
- [Kholas 95] J. Kholas, P.A. Monney. – A mathematical theory of hints : an approach to the dempster-shafer theory of evidence. *Lecture Notes in Economics and Mathematical Systems*, 425, 1995.
- [Kim 00] E.Y. Kim, S.H. Park, H.J. Kim. – A genetic algorithm-based segmentation of markov random field modeled images. *IEEE Signal Processing Letters*, 11(7) :301–303, 2000.
- [Kim 05] K. Kim, D. Harwood, L. Davis. – Background updating for visual surveillance. – *Int. Symposium on Visual Computing*, 2005.
- [Kirkpatrick 83] S. Kirkpatrick, C.D. Gellat, M.P. Vecchi. – Optimization by simulated annealing. *Science*, 220 :671–680, 1983.
- [Kitagawa 96] G. Kitagawa. – Monte carlo filter and smoother for non-gaussian nonlinear state space models. *Journal of Computational and Graphic Statistics*, 5 :1–25, 1996.
- [Kittler 98] J. Kittler, M. Hatef, R. Duin, J. Matas. – On combining classifiers. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 20(3) :226–238, 1998.
- [Klein 07a] J. Klein, C. Lecomte, P. Miché. – Fast color-texture discrimination : application to car-tracking. – *IEEE int. conf. on Intelligent Transportation Systems ( ITSC)*, pp. 541–546, 2007.
- [Klein 07b] J. Klein, C. Lecomte, P. Miché. – Tracking objects in videos with texture features. – *IEEE int. conf. on Circuits Electronisc Cybernetics and Systems*, Marrakech (Morroco), december 2007.
- [Klir 99] G.J. Klir, M.J. Wierman. – Uncertainty-based information. – *Physica-Verlag*, 1999.
- [Koller 93] D. Koller, K. Danilidis, H.-H. Nagel. – Model-based object tracking in monocular image sequences of road traffic scenes. *International Journal of Computer Vision*, 10(3) :257–281, 1993.
- [Konrad 00] J. Konrad. – *Handbook of image and video processing*. – Academic press., 2000.
- [Kuehnle 91] A. Kuehnle. – Symmetry-based recognition of vehicle rears. *Pattern Recognition Letters*, 12(4) :249–258, 1991.
- [Kuhn 55] H. Kuhn. – The hungarian method for solving the assignment problem. *Naval research logistics quart*, 2 :83–97, 1955.

- 
- [Kyrki 04] V. Kyrki, J.-K. Kämäräinen, H. Kälviäinen. – Simple gabor feature space for invariant object recognition. *Pattern Recognition Letters*, 25(3) :311–318, 2004.
- [Lamata 94] M. T. Lamata, S. Moral. – *Calculus with linguistic probabilities and beliefs*, chap. Advances in the Dempster-Shafer theory of evidence, pp. 133 – 152. – New York, NY, USA, John Wiley & Sons, Inc., 1994.
- [Laptev 05] I. Laptev, S. Belongie, P. Perez, J. Wills. – Periodic motion detection and segmentation via approximate sequence alignment. – *int. conf. on Computer Vision*, 2005.
- [Lecomte 99] C. Lecomte. – *Modélisation et reconnaissance de textures sous échelle en traitement d'images*. – PhD. Thesis, University of Poitier (France), 1999.
- [Lefaix 02] G. Lefaix, E. Marchand, P. Bouthemy. – Motion-based obstacle detection and tracking for car driving assistance. – *IEEE Int. Conf. on Pattern Recognition*, 2002.
- [Lefevre 00] E. Lefevre, O. Colot, P. Vannoorenberghe, D. de Bruçq. – Contribution des mesures d'information à la modélisation crédibiliste de connaissances. *Traitement du Signal*, 17 :87–97, 2000.
- [Lefevre 02] E. Lefevre, O. Colot, P. Vannoorenberghe. – Belief function combination and conflict management. *Information fusion*, 3 :149–162, 2002.
- [Lefèvre 01] S. Lefèvre, C. Fluck, B. Maillard, N. Vincent. – Un modèle de contour actif pour le suivi rapide d'objets en mouvement. application au suivi de joueurs de football. – *colloque GRETSI*, Toulouse (France), 2001.
- [Legland 03] F. Legland. – Filtrage particulière. – *GRETSI (invited paper)*, 2003.
- [Leung 05] A. P. Leung, S. Gong. – Online feature selection using mutual information for real-time multi-view object tracking. *LNCS*, 3723/2005 :184–197, 2005.
- [Li 02] Y. Li, A. Hilton, J. Illingworth. – A relaxation algorithm for real-time multiple view 3d-tracking. *Image and Vision Computing*, 20 :841859, 2002.
- [Li 04] P. Li, T. Zhang, B. Ma. – Unscented kalman filter for visual curve tracking. *Image and Vision Computing*, 22 :157–164, 2004.
- [Lin 01] T. Lin, C.-W. Ng, H.-J. Zhang, Q.-Y. Shi. – Integrating color and spatial features for content-based videoretrieval. – *Int. Conf. on Image Processing*, 2001.
- [Lisani 03] J. Lisani, J. Morel. – Detection of major changes in satellite images. – *IEEE Int Conf. on Image Processing*, pp. 941–944, 2003.
- [Liu 91] X. Liu. – *Analyse d'images couleur en composantes indépendantes par réseau de neurones*. – PhD. Thesis, University of Grenoble, 1991.
- [Liu 98] J. Liu, R. Chen. – Sequential monte carlo methods for dynamical systems. *Journal of the American Statistics Association*, 93 :1032–1044, 1998.
- [Lowe 99] D. Lowe. – Object recognition from local scale-invariant features. – *Int. Conf. on Computer Vision*, vol. 2, p. 11501157, Kerkyra (Greece), 1999.
- [Lowe 04] D. Lowe. – Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 20 :91–110, 2004.

- [Lu 99] G.J. Lu, A. Sajjandar. – Region-based shape representation and similarity measure suitable for content-based image retrieval. *Multimedia Systems*, 7(2) :165–174, 1999.
- [Lucas 81] B. D. Lucas, T. Kanade. – An iterative image registration technique with an application to stereo-vision. – *int. joint conf. on Artificial Intelligence*, 1981.
- [Malladi 95] R. Malladi, J.A. Sethian, B. C. Vemuri. – Shape modeling with front propagation : a level set approach. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 17(2) :158–175, 1995.
- [Marquardt 63] D. W. Marquardt. – An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11(2) :431–441, 1963.
- [Martin 07] A. Martin, C. Osswald. – Toward a combination rule to deal with partial conflict and specificity in belief functions theory. – *int. conf. on Information Fusion*, pp. 1–8, Quebec (Canada), 9-12 July 2007.
- [Megherbi 05a] N. Megherbi, S. Ambellouis, O. Colot, F. Cabestaing. – Data association in multi-target tracking using belief theory : handling target emergence and disappearance issue. – *IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS'2005*, pp. cd-rom, Como (Italy), september 2005.
- [Megherbi 05b] N. Megherbi, S. Ambellouis, O. Colot, F. Cabestaing. – Joint audio-video people tracking using belief theory. – *IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS'2005*, pp. cd-rom, Como (Italy), september 2005.
- [Megherbi 05c] N. Megherbi, S. Ambellouis, O. Colot, F. Cabestaing. – Multimodal data association based on the use of belief functions for multiple target tracking. – *Int. Conf. on Information Fusion*, pp. cd-rom, Philadelphia , PA (USA), july 2005.
- [Mehrotra 95] M. Mehrotra, J.E. Gary. – Similar shape retrieval in shape data management. *IEEE Computer*, 28(9) :57–62, 1995.
- [Memin 03] E. Memin. – *Estimation du flot-optique : contributions et panorama de différentes approches*. – Habilitation à diriger les recherches, INRIA - IRISA, 2003.
- [Mäenpää 04] T. Mäenpää, M. Pietikäinen. – Classification with color and texture : jointly or separately? *Pattern Recognition*, 37 :1629–1640, 2004.
- [Mercier 08] D. Mercier, B. Quost, T. Denoeux. – Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information Fusion*, 9(2) :246–258, april 2008.
- [Meribout 04] M. Meribout, L. Khriji, M. Nakanishi. – A robust hardware algorithm for real-time object tracking in video sequences. *Real time Imaging*, 10 :145–159, 2004.
- [Miche 95] P. Miche, R. Debrie. – Fast and self-adaptive image segmentation using extended declivity. *Annals of Telecommunication*, 50(3-4) :401–410, 1995.

- [Moravec 80] H. Moravec. – *Obstacle avoidance and navigation in the real world by a seeing robot rover*. – Rapport de Recherche n CMU-RI-TR-3, Carnegie-Mellon university, robotics institute, 1980.
- [Moreno 07] P. Moreno, M. J. Marin-Jimenez, A. Bernardino, J. Santos-Victor, N. Perez de la Blanca. – *A Comparative Study of Local Descriptors for Object Category Recognition : SIFT vs HMAX*. – LNCS, 515–522p., 2007.
- [Morrison 06] P. Morrison, J. J. Zou. – Skeletonization based on error reduction. *Pattern Recognition*, 39 :1099–1109, 2006.
- [Mourllion 05] B. Mourllion, D. Gruyer, C. Royere, S. Theroude. – Multi-hypotheses tracking algorithm based on the belief theory. – *int. conf. on Information Fusion , FUSION*, july 25-28 2005.
- [Moxey 03] C.E. Moxey, S.J. Sangwine, T.A. Ell. – Hypercomplex correlation techniques for vector images. *IEEE trans. on Signal Processing*, 51 :1941–1953, 2003.
- [Murphy 00] C.K. Murphy. – Combining belief functions with evidence conflicts. *Decision Support Systems*, 29 :1–9, 2000.
- [Muselet 05] D. Muselet. – *Reconnaissance automatique d'objets sous éclairage non contrôlé par analyse d'images couleur*. – PhD. Thesis, University of Lille I, 2005.
- [Nguyen 04] H. T. Nguyen, A. W. M. Smeulders. – Fast occluded object tracking by a robust appearance filter. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 26(8) :1099–1104, August 2004.
- [Nicolas 06] S. Nicolas. – *Segmentation par champ aléatoires pour l'indexation d'images de documents*. – PhD. Thesis, University of Rouen, 2006.
- [Odobez 95] J.M. Odobez, P. Bouthemy. – Robust multiresolution estimation of parametric motion models. *Journal of Visual Communication and Image Representation*, 6(4) :348–365, 1995.
- [Ohta 80] Y.I. Ohta, T. Kanade, T. Sakai. – Color information for region segmentation. *Computer Graphics and Image Processing*, 13 :222–241, 1980.
- [Ojala 96] T. Ojala, M. Pietikäinen, D. Harwood. – A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1) :51–59, 1996.
- [Okuma 04] K. Okuma, A. Taleghani, N. de Freitas, J. J. Little, D. G. Lowe. – A boosted particle filter : multitarget detection and tracking. – *European Conf. on Computer Vision*, 2004.
- [Oussalah 03] M. Oussalah. – On the use of hamacher's t-norm family for information aggregation. *Information Sciences*, 153 :107–154, 2003.
- [Palm 03] C. Palm. – Color texture classification by integrative co-occurrence matrices. *Pattern Recognition*, 37 :965–976, 2003.
- [Paragios 00] N. Paragios, R. Deriche. – Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 22(3) :266–280, 2000.

- [Parikh 04] D. Parikh, M. T. Kim, J. Oagaro, S. Mandayam, R. Polikar. – Combining classifiers for multisensor data fusion. – *IEEE Int. Conf. on Systems, Man and Cybernetics*, pp. 1232–1237, 2004.
- [Paschos 00] G. Paschos. – Fast color texture recognition using chromacity moments. *Pattern Recognition Letters*, 21 :837–841, 2000.
- [Pass 96] G. Pass, R. Zabih, J. Miller. – Comparing images using colour coherence vectors. – *ACM conf. on Multimedia*, pp. 65 – 73, Boston (USA), 1996.
- [Pentland 84] A.P. Pentland. – Fractal based description of natural scenes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6(6) :661–674, 1984.
- [Perez 02] P. Perez, C. Hue, J. Vermaak, M. Gangnet. – Color-based probabilistic tracking. – *European conf. on Computer Vision 2002*, pp. 661–675, 2002.
- [Perez 04] P. Perez, J. Vermaak, A. Blake. – Data fusion for visual tracking with particles. *Proceedings of the IEEE*, 92(3) :495–513, 2004.
- [Pham 01] D.-T. Pham. – Stochastic methods for sequential data assimilation in strongly non-linear systems. *Monthly Weather Review*, 129(5) :217–244, 2001.
- [Phong 75] B. Phong. – Illumination for computer generated pictures. *Communications of the ACM*, 18(16) :311–317, 1975.
- [Pietikäinen 02] M. Pietikäinen, T. Mäenpää, J. Viertola. – Color texture classification with color histograms and local binary patterns. – *International Workshop on Texture Analysis and Synthesis*, pp. 109–112, 2002.
- [Pitt 99] M. K. Pitt, N. Shephard. – Filtering via simulation : Auxiliary particle filters. *Journal of the American Statistical Association*, 94(446) :590–599, 1999.
- [Prewitt 70] J.M.S. Prewitt. – Object enhancement and extraction. *Picture Processing and Psychopictorics*, Academic Press, New York, 1970.
- [Quevedo 02] R. Quevedo, L. Carlos, J.M. Aguilera, L. Cadoche. – Description of food surfaces and micro structural changes using fractal image texture analysis. *Journal of Food Engineering*, 53 :361371, 2002.
- [Ramasso 07] E. Ramasso. – *Reconnaissance de séquences d'états par le modèle des croyances transférables. Application à l'analyse de video d'athlétisme.* – PhD. Thesis, University of Grenoble (France), 2007.
- [Rangarajan 91] K. Rangarajan, M. Shah. – Establishing motion correspondence. – *Conference Vision Graphics Image Process*, pp. 56–73, 1991.
- [Rasmussen 01] C. Rasmussen, G.D. Hager. – Probabilistic data association methods for tracking complex visual objects. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 23(6) :560–576, 2001.
- [Regh 94] J.M. Regh, T. Kanade. – Visual tracking of high dof articulated structures : an application to human hand tracking. – *IEEE European Conf. on Computer Vision*, pp. 35–46, 1994.
- [Reid 79] D.B. Reid. – An algorithm for tracking multiple target. *IEEE trans. on Autom. Control*, 24 :843–854, 1979.

- 
- [Ren 03] Y. Ren, C. Chua, Y. Ho. – Statistical background modeling for non-stationary camera. *Pattern Recognition Letters*, 24(1-3) :183–196, 2003.
- [Ren 08] X. Ren. – Local grouping for optical flow. – *int. conf. on Computer Vision and Pattern Recognition*, 2008.
- [Rogova 08] G. Rogova, E. Boisse. – *Information quality effects on information fusion*. – Rapport de Recherche nTR 2005-270, Defence R&D Canada - Valcartier, may 2008.
- [Rosales 99] R. Rosales, S. Scaloff. – 3d trajectory recovery for tracking multiple objects and trajectory guided recognition of actions. – *IEEE int conf. on Computer Vision and Pattern Recognition*, pp. 117–123, 1999.
- [Rosenbluth 55] M.N. Rosenbluth, A.W. Rosenbluth. – Monte carlo calculation of the average extension of molecular chains. *Journal of Chemical Physics*, 23 :356–359, 1955.
- [Ross 03] M. Ross. – Evaluation and improvement of region-based motion segmentation. – *8th Workshop on Vision, Modeling, and Visualization*, pp. 55–61, Berlin (Germany), 2003.
- [Royère 00] C. Royère, D. Gruyer, V. Berge-Cherfaoui. – Data association with believe theory. – *Int. Conf. on Information Fusion*, vol. 1, pp. cd-rom, Paris (France), 2000.
- [Salari 90] V. Salari, I.K. Sethi. – Feature point correspondence in the presence of occlusion. *IEEE trans. Pattern Analysis and Machine Intelligence*, 12 :87–91, 1990.
- [Scassellati 94] B. Scassellati, S. Slexopoulos, M. Flickner. – Retrieving images by 2d shape : a comparison of computation methods with human perceptual judgements. – *SPIE, Storage and Retrieval for Image and Video Databases II*, pp. 2–14, San Jose (USA), 1994.
- [Schulz 01] D. Schulz, W. Burgard, D. Fox, A.B. Cremers. – Tracking multiple moving targets with a mobile robot using particle filters and statistical data association. – *IEEE International Conference on Robotics and Automation*, vol. 2, pp. 1665 – 1670, 2001.
- [Sentz 02] K. Sentz, S. Ferson. – *Combination of evidence in Dempster-Shafer theory*. – Rapport de recherche, SANDIA tech. report, 2002.
- [Sethi 87] I. K. Sethi, R. Jain. – Finding trajectories of feature points in a monocular image sequence. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 9 :56–73, 1987.
- [Shafer 76] G. Shafer. – *A Mathematical Theory of Evidence*. – Princeton University Press, 1976.
- [Sheikh 05] Y. Sheikh, M. Sha. – Bayesian modeling of dynamic scenes for object detection. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 27(11) :603–619, 2005.
- [Shi 94] J. Shi, C. Tomasi. – Good features to track. – *IEEE int. conf. Computer Vision and Pattern Recognition*, pp. 593–600, 1994.

- [Shi 98] J. Shi, J. Malik. – Motion segmentation and tracking using normalized cuts. – *Int. Conf. on Computer Vision*, 1998.
- [Siddiqi 98] K. Siddiqi, A. Shokoufandeh, S. Dickinson, S. Zucker. – Shock graphs and shape matching. – *IEEE Int. Conf. on Computer Vision*, pp. 222–229, Bombay (India), 1998.
- [Silveira 07] G. Silveira, E. Malis. – Real-time visual tracking under arbitrary illumination changes. – *int. conf. Computer Vision and Pattern Recognition*, 2007.
- [Smarandache 04a] F. Smarandache. – An in-depth look at information fusion rules and the unification of fusion theories. – Univeristy of New Mexico, 2004.
- [Smarandache 04b] F. Smarandache, J. Dezert. – *Advances and Applications of DSMT for Information Fusion*,. – American Research Press, 2004.
- [Smarandache 05a] F. Smarandache, J. Dezert. – The combination of paradoxical, uncertain, and imprecise sources of information based on dsmt and neutro-fuzzy inference. – *10th International Conference on Fuzzy Theory and Technology*, Salt Lake City, Utah (USA), July 21-26 2005.
- [Smarandache 05b] F. Smarandache, J. Dezert. – Information fusion based on new proportional conflict redistribution rules. – *Int. conf. on Information Fusion*, Philadelphia (USA), 25-29 july 2005.
- [Smets 90] P. Smets. – The combination of evidence in the transferable belief model. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(5) :447–458, 1990.
- [Smets 93] P. Smets. – Belief functions : the disjunctive rule of combination and the generalized bayesian theorem. *Int. Journal of Approximate Reasoning*, 9 :1–35, 1993.
- [Smets 94] P. Smets, R. Kennes. – The transferable belief model. *Artificial Intelligence*, 66(2) :191–234, 1994.
- [Smets 95] P. Smets. – The canonical decomposition of weighted belief. – *int. joint. conf. on Artificial Intelligence*, pp. 1896–1901, 1995.
- [Smets 97] P. Smets. – The alpha-junctions : Combination operators applicable to belief functions. – *First International Joint Conference on Qualitative and Quantitative Practical Reasoning*, vol. 1244, pp. 131–153. Lecture Notes In Computer Science, 1997.
- [Smets 00] P. Smets. – Belief functions and the transferable belief model. *Philippe Smets and the Imprecise Probabilities Project*, 2000.
- [Smets 05] P. Smets. – Belief functions on real numbers. *International Journal of Approximate Reasoning*, 40 :181–223, 2005.
- [Smets 06] P. Smets. – Analyzing the combination of conflicting belief functions. *Information Fusion*, 8 :387–412, 2006.
- [Smets 07] P. Smets, B. Ristic. – Kalman filter and joint tracking and classification based on belief functions in the tbm framework. *Information Fusion*, 8 :16–27, 2007.
- [Sonka 93] M. Sonka, V. Hlavac, R. Boyle. – *Image processing, analysis and machine vision*. – Chapman & Hall, London (UK), 193–242p., 1993.

- 
- [Spinei 98] A. Spinei, D. Pellerin, J. Herault. – Spatiotemporal energy-based method for velocity estimation. *Signal Processing*, 65 :347–362, 1998.
- [Steux 02] B. Steux, C. Laugeau, L. Salesse, D. Wautier. – Fade : a vehicle detection and tracking system featuring monocular color vision and radar data fusion. – *Intelligent Vehicles symposium*, vol. 2, pp. 632– 639, june 17-21 2002.
- [Suard 06] F. Suard. – *Méthodes à noyaux pour la détection de piétons*. – PhD. Thesis, INSA of Rouen, 2006.
- [Sun 06] Z. Sun, G. Bebis, R. Miller. – On-road vehicle detection using optical sensors : A review. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 28(6) :694–711, 2006.
- [Swain 91] M.J. Swain, D.H. Ballard. – Color indexing. *Int. Journal of Computer Vision*, 11 :1132, 1991.
- [Tarel 06] J.-P. Tarel, P. Charbonnier, S.-S. Ieng. – Ajustement robuste et simultané de plusieurs courbes et application au suivi de marquages routiers. – *Journées des Sciences de l'Ingénieur*, Marne la vallée (France), 2006.
- [Taubin 92] G. Taubin, D.B. Cooper. – Object recognition based on moment. – Geometric invariance in computer vision, MIT press, 1992. pp. 375-397.
- [Teague 80] M.R. Teague. – Image analysis via the general theory of moments. *Journal of Opt. Soc. Am.*, 70(10) :920–930, 1980.
- [Tian 05] Y.L. Tian, A. Hampapur. – Robust salient motion detection with complex background for real-time video surveillance. – *Workshop on Motion and Video Computing*, 2005.
- [Tombre 00] K. Tombre, C. Ah Soon, P. Dosch, G. Masini, S. Tabbone. – Stable and robust vectorization : How to make the right choices. *In Workshop on Graphics Recognition (GREC), Lecture Notes in Computer Science (LNCS)*, 1941 :3–18, 2000.
- [Torma 04] P. Torma, C. Szepesvari. – Enhancing particle filters using local likelihood sampling. – *European Conf. on Computer Vision 2004*, pp. 16–27, 2004.
- [Toulminet 06] G. Toulminet, M. Bertozzi, S. Mousset, A. Benschair, A. Broggi. – Vehicle detection by means of stereo vision-based obstacles features extraction and monocular pattern analysis. *IEEE trans. on Image Processing*, 15(8) :2364–2375, 2006.
- [Toyama 99] K. Toyama, J. Krumm, B. Brummit, B. Meyers. – Wallflower : principles and practice of background maintenance. – *int. conf. on Computer Vision*, 1999.
- [Trémeau 04] A. Trémeau, C. Fernandez-Maloigne, P. Bonton. – *Image Numérique Couleur*. – Dunod, 2004.
- [Unser 95] M. Unser. – Texture classification and segmentation using wavelet frames. *IEEE trans. on Image Processing*, 4(11) :1549–1560, 1995.
- [Van der Merwe 00] R. Van der Merwe, A. Doucet, N. de Freitas, E. Wan. – *The Unscented Particle Filter*. – Rapport de Recherche n TR380, Cambridge University Engineering Department, 2000.

- [Van der Merwe 01] R. Van der Merwe, E. Wan. – The square-root unscented kalman filter for state and parameter-estimation. – *International Conference on Acoustics, Speech, and Signal Processing*, 2001.
- [Vandenbroucke 05] N. Vandenbroucke, L. Macaire. – Les espaces couleurs. – *Ecole d'hiver sur l'image numérique couleur*, pp. 140–162, Lille, 2005.
- [Veenman 03] C.J. Veenman, M.J.T. Reinders, E. Backer. – Establishing motion correspondence using extended temporal scope. *Artificial Intelligence*, 145 :227243, 2003.
- [Veit 04] T. Veit, F. Cao, P. Bouthemy. – *An a contrario decision framework for motion detection*. – Rapport de recherche, INRIA, 2004.
- [Velduis 99] J. H. Velduis, G. W. Brodland. – A deformable block-matching algorithm for tracking epithelial cells. *Image and Vision Computing*, 17 :905–911, 1999.
- [Vermaak 03] J. Vermaak, A. Doucet, P. Perez. – Maintaining multi-modality through mixture tracking. – *IEEE European Conf. on Computer Vision*, pp. 1110–1116, Nice (France), october 2003.
- [Vidal 04] R. Vidal, Y. Ma. – A unified algebraic approach to 2-d and 3-d motion segmentation. – *European Conf. on Computer Vision*, 2004.
- [Vidal 05] R. Vidal, D. Singaraju. – A closed form solution to direct motion segmentation. – *int. conf. on Computer Vision and Pattern Recognition*, 2005.
- [Viola 04] P. Viola, M. Jones. – Robust real-time face detection. *International Journal of Computer Vision*, 57(2) :137154, 2004.
- [Wan 00] E.A. Wan, R. van der Merwe. – The unscented kalman filter for non-linear estimation. – *IEEE Symposium on Adaptive Systems for Signal Processing, Communication and Control*, pp. 205–219, 2000.
- [Wang 94] J.Y.A Wang, E.H. Adelson. – Representing moving images with layers. *IEEE trans. on Image Processing Special Issue*, 3(5) :625–638, 1994.
- [Weber 95] J. Weber, J. Malik. – Robust computation of optical flow in a multi-scale differential framework. *int. journal of Computer Vision*, 14(1) :5–19, 1995.
- [Weiss 97] Y. Weiss. – Smoothness in layers : Motion segmentation using non-parametric mixture estimation. – *int. conf. on Computer Vision and Pattern Recognition*, 1997.
- [Weng 06] S.-K. Weng, C.-M. Kuo, S.-K. Tu. – Video object tracking using adaptive kalman filter. *Journal of Visual Communication and Image Representation*, 17 :11901208, 2006.
- [Whelan 01] P.F. Whelan, A. Drimbarean. – Experiments in colour textures analysis. *Pattern Recognition Letters*, 22 :1161–1167, 2001.
- [Wixson 00] L. Wixson. – Detecting salient motion by accumulating directionnaly consistent flow. *IEEE trans. on Pattern Analysis and Machine Intelligence*, 22(8) :774–780, 2000.
- [Wolf 93] D. Wolf, R. Husson. – Application des ondelettes à l'analyse de textures et à l'inspection de surfaces industrielles. *Journal de Physique III*, 3 :2133–2148, 1993.

- 
- [Wu 04] Y. Wu, T.S. Huang. – Robust visual tracking by integrating multiple cues based on co-inference learning. *Int. Journal of Computer Vision*, 58 :55–71, 2004.
- [Wu 05] Q. Wu, H. Cheng, B. Jeng. – Motion detection via change-point detection for cumulative histograms of ratio images. *Pattern Recognition Letters*, 26(5) :555–563, 2005.
- [Yager 87a] R. Yager. – On the dempster-shafer framework and new combination rules. *Information sciences*, 41 :93–138, 1987.
- [Yager 87b] R. Yager. – Quasi-associative operations in the combination of evidence. *Kybernetes*, 16 :37–41, 1987.
- [Yilmaz 06] A. Yilmaz, O. Javed, M. Shah. – Object tracking : A survey. *ACM Computing Surveys*, 38, 2006.
- [Yin 08] Z. Yin, R. T. Collins. – Object tracking and detection after occlusion via numerical hybrid local and mode-seeking. – *int. conf. Computer Vision and Pattern Recognition*, 2008.
- [Yoo 99] T-W Yoo, I-S. Oh. – A fast algorithm for trackingnext term human faces based on chromatic previous termhistogramsnext term. *Pattern Recognition Letters*, 20 :967–978, 1999.
- [Zadeh 65] L. A. Zadeh. – Fuzzy sets. *Journal of Information and Control*, 8 :338–353, 1965.
- [Zadeh 78] L.A. Zadeh. – Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1 :3–28, 1978.
- [Zadeh 86] L.A. Zadeh. – A simple view of the dempster-shafer theory of evidence and its implication for the rule of combination. *the Artificial Intelligence Magazine*, 7 :85–90, 1986.
- [Zass 08] R. Zass, A. Shashua. – Probabilistic graph and hypergraph matching. – *int conf. Computer Vision and Pattern Recognition*, 2008.
- [Zhang 94] L. Zhang. – *Advances in the Dempster-Shafer theory of evidence*, chap. Representation, Independence, and combination of evidence in the Dempster-Shafer theory, pp. 51–69. – John Wiley & Sons, 1994.
- [Zhang 02] D.S. Zhang. – *Image retrieval based on shape*. – PhD. Thesis, Monash University (Australia), 2002.
- [Zhang 04] D. Zhang, G. Lu. – Review of shape representation and description techniques. *Pattern Recognition*, 37 :1–19, 2004.
- [Zhang 06] Y. Zhang, S.T. Kiselewich, W.A. Bauson, R. Hammoud. – Robust moving object detection at distance in the visible spectrum and beyond using a moving camera. – *int. conf. on Computer Vision and Pattern Recognition Workshop*, 2006.
- [Zhang 07] P. Zhang, I. Gardin, P. Vannoorenberghe. – Information fusion using evidence theory for segmentation of medical images. – *Int. Col. on Information Fusion*, pp. 265–272, Xi'An, P.R.China, August 2007.

- 
- [Zhao 07] W. Zhao, T. Fang, Yan Jiang. – Data fusion using improved Dempster-Shafer evidence theory for vehicle detection. – *int. conf. on Fuzzy Systems and Knowledge Discovery*, 2007.
- [Zhong 06] X. Zhong, J. Xue, N. Zheng. – Graphical model based cue integration strategy for head tracking. – *British Machine Vision Conf.*, 2006.
- [Zhu 05] S. Zhu, Q. Avidan, K.-T. Cheng. – Learning a sparse, corner-based representation for time-varying background modeling. – *int. conf. on Computer Vision*, 2005.
- [Zucker 76] S. W. Zucker. – Toward a model of texture. – *CGIP*, vol. 5, pp. 686–693, 1976.