



REPUBLIQUE ALGÉRIENNE DÉMOCRATIQUE ET POPULAIRE
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université des Sciences et de la Technologie d'Oran "Mohamed Boudiaf"
Faculté de génie électrique
Département d'électronique



THÈSE

en vue de l'obtention du diplôme de Doctorat En Sciences

Spécialité : Électronique

Option : Technique de communication moderne

Présentée par:

M. BOUMEDIENE MOHAMMED

SUJET DE THÈSE

Contributions aux systèmes d'aide à la conduite automobile

Soutenue le : 09 juin 2015

Devant le jury composé de :

M. Belbachir Mohamed Fouazi	Professeur	USTO-MB, Algérie	Président
M. Ouamri Abdelaziz	Professeur	USTO-MB, Algérie	Rapporteur
M. Bereksi Reguig Fethi	Professeur	Univ. Tlemcen, Algérie	Examineur
M. Loukil Abdelhamid	Maître de conférences	USTO-MB, Algérie	Examineur
M. Ould Mammar Madani	Maître de conférences	Univ. Mostaganem, Algérie	Examineur
M. Cudel Christophe	Maître de conférences	Univ. Haute-Alsace, France	Examineur
M. Lauffenburger Jean-Philippe	Maître de conférences	Univ. Haute-Alsace, France	Invité

Remerciements

Les travaux présentés dans ce manuscrit de thèse ont été financés par une bourse PROFAS, un programme franco-algérien de formation supérieure, obtenue via l'université de Mostaganem. Ils ont été réalisés en étroite collaboration entre le Laboratoire Signaux et Images (LSI) de l'Université des Sciences et de la Technologie d'Oran (USTO) en Algérie et le laboratoire Modélisation, Intelligence, Processus et Systèmes (MIPS) de l'Université de Haute-Alsace (UHA) en France. Je suis très reconnaissant envers le professeur Abdelaziz Ouamri, directeur du LSI, et le professeur Olivier Haeberlé, directeur du MIPS, de m'avoir donné cette opportunité de recherche durant 18 mois.

Tout d'abord je tiens à remercier mes parents et mes soeurs de m'avoir soutenu et supporté durant toutes ces années de recherche.

Mes remerciements s'adressent tout naturellement à mon directeur de thèse, Abdelaziz Ouamri, qui a dirigé mon travail. Je remercie par ailleurs les membres du jury : Mohamed Faouazi Belbachir, Fethi Reguig Bereksi, Abdelhamid Loukil, et Madani Ould Mammar d'avoir accepté d'examiner ce modeste travail.

Mes sincères remerciements s'adressent également à Christophe Cudel et à Jean-Philippe Lauffenburger du laboratoire MIPS, qui ont beaucoup contribué à mon travail ; leur encadrement, leurs conseils et commentaires précieux m'ont permis de progresser dans ma thèse. Je voudrais aussi leur exprimer ma profonde reconnaissance d'avoir corrigé mon manuscrit et de m'avoir soutenu jusqu'à la finalisation de ma thèse.

Un grand merci à Jérémie Daniel pour sa collaboration et ses conseils pertinents qui ont permis d'améliorer la qualité de mon travail. Je le remercie également d'avoir corrigé une bonne partie du manuscrit ainsi que les publications.

Je tiens à remercier aussi Michel Basset pour son accueil au sein de l'équipe MIAM du MIPS.

Que l'ensemble des membres du MIAM soit remercié pour son accueil et nos échanges, en particulier Benjamin Mourllion pour avoir partagé avec moi son bureau pendant une année et Joël Lambert de m'avoir fait découvrir la région Alsace, que de beaux souvenirs des villages alsaciens visités. Je remercie de même Rodolfo Orjuela, Abderazik Birouche, Hossni Zebiri, Sabra Hamza, Rachid Attia, Oussama Derbel, Emmeric Thevenet, Dragos Dobre, Laura Alarcon, Thomas Brunner, et Guillaume Strub.

Ces remerciements seraient incomplets si je ne mentionnais pas l'ensemble de mes collègues du LSI ainsi que ceux du département d'informatique de l'université de Mostaganem dont certains sont devenus des amis.

Je dédie cette thèse à ma FAMILLE.

Table des matières

Remerciements	ii
Dédicace	iii
Table des matières	iv
Liste des figures	vii
Liste des tableaux	xi
Abréviations	xii
Liste des publications	xiv
Introduction générale	xv
1 Contexte de recherche	1
1.1 Introduction	1
1.2 Sécurité routière	2
1.3 Systèmes avancés d'aide à la conduite	4
1.3.1 Principe des <i>ADAS</i>	5
1.3.2 Niveaux de coopération des <i>ADAS</i>	6
1.3.2.1 <i>ADAS</i> autonome	6
1.3.2.2 <i>ADAS</i> coopératif	6
1.3.3 Exemples d' <i>ADAS</i> autonomes	6
1.4 Système de reconnaissance de panneaux routiers	8
1.4.1 Reconnaissance par système d'information géographique	8
1.4.2 Reconnaissance par vision	9
1.5 Contributions	11
1.6 Conclusion	13
2 Détection et suivi de panneaux routiers : état de l'art	15
2.1 Introduction	15
2.2 Détection de panneaux : état de l'art	16
2.2.1 Approches couleurs	22
2.2.2 Approches contours	24

2.2.3	Approches descripteurs	31
2.2.3.1	Descripteurs utilisés pour l'extraction	31
2.2.3.2	Classification des descripteurs	36
2.2.4	Conclusion	41
2.3	Fusion de données	42
2.3.1	Processus de fusion	42
2.3.2	Imperfections de l'information	43
2.4	Stratégies pour le suivi temporel	44
2.4.1	Étapes du suivi temporel	44
2.4.1.1	Filtrage	45
2.4.1.2	Association de données	46
2.4.2	Suivi temporel : état de l'art	48
2.4.2.1	Approches suivi- <i>après</i> -reconnaissance	48
2.4.2.2	Approches suivi- <i>après</i> -détection	50
2.4.2.3	Conclusion	52
2.4.3	Méthodes d'association de données	53
2.4.3.1	Méthode du plus proche voisin	53
2.4.3.2	Méthodes probabilistes	54
2.4.3.3	Le pistage à hypothèses multiples	56
2.4.3.4	Association crédale	57
2.5	Conclusion	58
3	Détection de panneaux par codage de contours	59
3.1	Introduction	59
3.2	Approche proposée : codage de contours	59
3.2.1	Principe	60
3.2.2	Détection de coins	60
3.2.3	Codage de contours	62
3.2.4	Classification de coins	62
3.2.5	Détection de panneaux triangulaires	65
3.2.5.1	Principe	65
3.2.5.2	L'algorithme RSLD	65
3.2.5.3	Estimation des gains de RSLD	66
3.2.5.4	Détection multi-échelle	69
3.2.5.5	Détection de la ligne de base	70
3.2.6	Détection de panneaux circulaires	70
3.2.6.1	Principe	70
3.2.6.2	Détection par <i>template matching</i>	71
3.2.7	Groupement des résultats de la détection	72
3.3	Évaluation expérimentale	74
3.3.1	Bases de données	74
3.3.2	Analyse des performances	74
3.3.3	Robustesse des méthodes proposées aux dégradations d'images	78
3.4	Conclusion	83
4	Suivi temporel de panneaux	84
4.1	Introduction	84

4.2	Association crédale multi-objets pour le suivi de ROIs	85
4.2.1	Principe du Suivi de Multi-ROIs	86
4.2.2	Filtrage	87
4.2.2.1	Modèle dynamique	87
4.2.2.2	Modèle d'observation	89
4.2.3	Association par fusion de données	89
4.2.4	Analyse des pistes	90
4.3	Théorie des croyances	90
4.3.1	Introduction	90
4.3.2	Modélisation du contexte	91
4.3.3	Estimation des croyances	92
4.3.4	Combinaison des sources	95
4.3.4.1	Combinaison conjonctive	95
4.3.4.2	Combinaison disjonctive	96
4.3.4.3	Combinaison mixte	97
4.3.5	Prise de décision	97
4.3.5.1	Maximum de crédibilité	97
4.3.5.2	Maximum de plausibilité	98
4.3.5.3	Maximum de probabilité pignistique	99
4.4	Association multi-objets par MCT	99
4.4.1	Niveau crédal	100
4.4.1.1	Modélisation	101
4.4.1.2	Estimation	102
4.4.1.3	Combinaison des bbms	104
4.4.2	Niveau pignistique	106
4.5	Algorithmes de prise de décision par MCT	108
4.5.1	Probabilité pignistique jointe	109
4.5.2	Probabilité pignistique jointe classée	109
4.5.3	Probabilité pignistique locale	111
4.5.4	Probabilité pignistique classée par gradient	112
4.5.5	Seuil	113
4.5.6	Décision duale	114
4.5.7	Conclusion	116
4.6	Evaluation des algorithmes d'association	117
4.6.1	Résultats de simulation	117
4.6.2	Suivi de panneaux par association crédale : résultats expérimentaux	120
4.6.2.1	Configuration du système	120
4.6.2.2	Performances de l'algorithme SMR	121
4.6.2.3	Impact de l'association crédale sur l'algorithme SMR	124
4.7	Conclusion	125
	Conclusion générale	127
	Bibliographie	130

Table des figures

1.1	Répartition des victimes tuées et du parc mondial de véhicules dans le monde [OMS, 2013].	2
1.2	Répartition des victimes tuées selon les catégories d'usagers [OMS, 2013].	3
1.3	L' <i>Avertisseur de Franchissement Involontaire de Ligne</i> (AFIL) de PSA Peugeot Citroën [Jean-Christophe RIAT, 2005].	7
1.4	Le système ILENA (<i>Intelligent LEarning NAVigation</i>) de BMW pour la navigation intelligente.	8
1.5	Schéma général d'un système de navigation par satellite.	8
1.6	Schéma général d'un système de reconnaissance de panneaux routiers.	10
1.7	Exemple de fonctionnement d'un système de reconnaissance de panneaux routiers.	10
1.8	Exemple de reconnaissance de panneaux au niveau d'une sortie d'autoroute.	11
2.1	Diagramme de l'étape de <i>Détection</i>	17
2.2	Représentation de l'espace HSI. (a) représentation conique de l'espace RGB et (b) les composantes des espaces HSI.	22
2.3	Exemple de segmentation dans l'espace HSI [Maldonado-bascón <i>et al.</i> , 2007]. (a) et (d) Images originales de la scène routière. (b) et (e) Respectivement segmentation de la couleur bleue et rouge. (c) et (f) Suppression des régions trop petites ou trop grandes.	23
2.4	Principe de la transformée de Hough.	26
2.5	Détection de cercles par (a) la transformée de Hough et (b) la Transformée en Symétrie Radiale (TrSR).	27
2.6	Principe du Détecteur de Polygones Réguliers (DPR) ([Houben, 2011]).	28
2.7	Illustration du principe de la Transformée en Sommet et Bissectrice (TSB).	28
2.8	Exemple de résultat de la Transformée en Sommet et Bissectrice [Belaroussi et Tarel, 2009]. (a) Image originale, (b) l'accumulateur S_{acc} , et (c) l'accumulateur B_{acc}	29
2.9	La symétrie centrale et la symétrie axiale entre les vecteurs. (a) Symétrie centrale et (b) symétrie axiale.	30
2.10	Illustration du principe de la Transformée Chinoise Bilatérale. (a) Cas des contours clairs/sombres, (b) cas des contours sombres/clairs.	30
2.11	Illustration du principe des méthodes de vote de Houben [Houben, 2011]. (a) la méthode STVUT, et (b) la méthode STVUE.	31
2.12	Les caractéristiques de Haar.	32
2.13	Représentation de l'image intégrale [Viola et Jones, 2001a] où la somme des pixels dans la zone D est calculée directement par $ii(4) - (ii(2) + ii(3)) + ii(1)$	33
2.14	Les caractéristiques dipôles dissociées [Balas et Sinha, 2003].	33
2.15	Le principe de décomposition pour le calcul des HOG [Dalal, 2005]. L'image est divisée en blocs qui peuvent se recouvrir. Chaque bloc est subdivisé en cellules où un histogramme d'orientations est calculé.	34

2.16	Principe des distances aux bords (DtB) représentées par $D1$, $D2$, $D3$, et $D4$ ([Lafuente-Arroyo <i>et al.</i> , 2005]).	35
2.17	Les vecteurs DtB normalisés des distances : $D1$, $D2$, $D3$, et $D4$ du panneau triangulaire de la figure 2.16 ([Lafuente-Arroyo <i>et al.</i> , 2005]).	35
2.18	Décomposition fréquentielle du bord d'un panneau triangulaire ([Gil Jiménez <i>et al.</i> , 2008]). (a) Représentation de la distance aux bords en fonction de l'orientation α . (b) Spectre du signal illustré dans (a).	36
2.19	Diagramme d'un détecteur en cascade attentionnelle en 3 étages [Viola et Jones, 2001b].	37
2.20	Séparation linéaire par les hyperplans H_{+1} , H_{-1} [Burges, 1998]. Les vecteurs supports sont entourés d'un cercle.	40
2.21	La fonction non linéaire transférant les données de l'espace \mathbb{R}^2 vers l'espace de Hilbert \mathcal{H} [Zaklouta, 2011].	41
2.22	Diagramme du processus de fusion [Martin, 2005].	43
2.23	Diagramme de construction d'un système de suivi multi-objets.	45
2.24	Exemple du suivi temporel de panneaux routiers [Garcia-Garrido <i>et al.</i> , 2006].	49
2.25	Exemple de fusion de panneaux identiques par l'approche Bayésienne [Meuter <i>et al.</i> , 2011].	50
2.26	Exploitation de l'information spatio-temporelle pour l'élimination de faux positifs [Šegvić <i>et al.</i> , 2014]. a) Le contexte spatio-temporel et les faux positifs. b) Exemple de trois trajectoires de panneaux dans un espace 3D (image 2D + temps).	51
2.27	Exemple d'association multiple. la cible z_4 est associée à deux pistes x_1 et x_2	53
2.28	Détection des formes triangulaires d'un panneau de cédez le passage.	54
3.1	Détection des points d'intérêt liés aux panneaux.	61
3.2	Diagramme de l'approche proposée.	61
3.3	Illustration de l'espace de direction des contours divisé selon les 4 principales classes $\{1, 2, 3, 4\}$	62
3.4	Distribution de contours utilisée pour le codage des coins. (a) La distribution de contours pour la détection d'un sommet supérieur (triangle). (b) La distribution de contours pour la détection d'une pointe (triangle inversé).	63
3.5	Illustration d'un exemple de codage de coins sur une image de synthèse.	64
3.6	Résultats de la classification de coins sur 32 images de la base de Grigorescu [Grigorescu et Petkov, 2003]. La courbe supérieure représente le nombre total des coins détectés par le détecteur de Harris. La courbe inférieure représente le nombre de positions retenues suite à la classification de coins pour la <i>Détection</i> de panneaux triangulaires	64
3.7	Illustration de l'inversion des coordonnées y autour des coins sélectionnés. (a) représente le résultat de l'inversion verticale des coordonnées y autour d'un sommet supérieur. (b) représente le résultat de l'inversion verticale des coordonnées y autour d'une pointe.	66
3.8	Images de synthèse. les points (o) représentent les inliers, les données liées aux côtés diagonaux du triangle et bruitées par un bruit Gaussien ; les points ($*$) représentent les outliers qui sont distribués uniformément. (a) représente les côtés diagonaux d'un triangle et (b) représente ces côtés après une inversion des coordonnées y	67
3.9	Comparaison entre la détection par RANSAC avec et sans l'inversion des coordonnées y	67

3.10	Illustration du temps de calcul pour la détection (RANSAC) avec et sans l'inversion des coordonnées y . Le nombre d'inliers est de 30.	69
3.11	Détection Multi-échelle.	69
3.12	Détection de la ligne de base. (a) La sous région horizontale et son histogramme pour la <i>Détection</i> d'un triangle. (b) La sous région horizontale et son histogramme pour la <i>Détection</i> d'un triangle inversé.	70
3.13	Exemples de <i>Détection</i> de panneaux triangulaires. (a) Détection des positions candidates par le détecteur de Harris. (b) La classification des coins. (c) Résultat de la détection de segments diagonaux par RANSAC. (d) Résultats de la détection de la base ligne.	71
3.14	Le modèle (<i>template</i>) proposé pour la <i>Détection</i> de panneaux circulaires.	72
3.15	Exemples de modèle à plusieurs échelles.	72
3.16	Exemple de groupement des détections. (a) <i>Détection</i> d'un panneau circulaire à partir d'image de scène routière, (b) zoom sur le résultat de la <i>Détection</i> , et (c) celui du processus de groupement.	73
3.17	Exemples de Faux Positifs (FP) lors de la <i>Détection</i> de panneaux triangulaires et circulaires sur la base de données <i>Stéréopolis</i>	75
3.18	Exemples de panneaux circulaires sans coin détecté au centre.	76
3.19	Exemples de <i>Détection</i> de panneaux de danger.	78
3.20	Exemples de <i>Détection</i> de panneaux cédez le passage.	79
3.21	Exemples de <i>Détection</i> de panneaux d'interdiction.	79
3.22	Exemples de Faux Positifs (FP) lors de la <i>Détection</i> de panneaux triangulaires et circulaires sur la base de données du MIPS.	80
3.23	Taux de détection de panneaux ((a) triangulaires et (b) circulaires) sur des images dégradées par l'ajout d'un bruit Gaussien.	81
3.24	Taux de détection de panneaux ((a) triangulaires et (b) circulaires) sur des images floutées à l'aide d'un filtre Gaussien.	82
3.25	Exemples de <i>Détection</i> de panneaux triangulaires sur les images de <i>Grigorescu</i> après application d'une rotation. (a) L'image subie une rotation de +10 degrés (<i>image de gauche</i>) et -10 degrés (<i>image de droite</i>). (b) L'image subie une rotation de +14 degrés (<i>image de gauche</i>) et -14 degrés (<i>image de gauche</i>).	82
3.26	Exemples de <i>Détection</i> de panneaux circulaires sur les images de <i>Grigorescu</i> après application d'une rotation. (a) L'image subie une rotation de +14 degrés (<i>image de gauche</i>) et -14 degrés (<i>image de droite</i>). (b) L'image subie une rotation de +20 degrés (<i>image de gauche</i>) et -20 degrés (<i>image de gauche</i>).	83
4.1	Diagramme du flux de données entre le SMR et les étapes de <i>Détection</i> et <i>Reconnaissance</i>	86
4.2	Diagramme du SMR.	87
4.3	Jeu de masses d'une source spécialisée selon [Rombaut, 1998].	94
4.4	Représentation des sous-ensembles. (a) de la crédibilité de H_1 ; (b) de la plausibilité de H_1 [Bernet, 2012].	98
4.5	Principe d'association de données par théorie des croyances [Daniel et Lauffenburger, 2012].	100
4.6	Exemple de scénario avec 3 objets perçus et 2 objets connus.	111
4.7	Exemple réel d'association avec 2 objets perçus et 2 objets connus.	116
4.8	Résultats de la détection sans SMR sur plusieurs images successives. (a) et (d) : détection de panneau correcte. (b) et (c) : présence de FP qui peuvent être supprimés par le SMR.	121

4.9	Comparaison entre les courbes ROC (Receiver Operating Characteristic) du système sans et avec SMR pour les panneaux circulaires (a) et triangulaires (b).	123
-----	---	-----

Liste des tableaux

1.1	Avantages et limites des deux techniques de reconnaissance : par vision et par SIG.	10
2.1	Avantages et limites des différentes approches de détection de panneaux.	17
3.1	Comparaison des taux de détection des panneaux triangulaires sur la base <i>Stéréopolis</i>	75
3.2	Comparaison des taux de détection des panneaux circulaires sur la base <i>Stéréopolis</i>	76
3.3	Séquences de test capturées sur le réseau routier français.	76
3.4	Performances pour les séquences 1 et 2	78
4.1	Masses initiales des associations $X \longrightarrow Y$	103
4.2	Masses initiales des associations $X \longleftarrow Y$	103
4.3	Matrice de croyance $M_{i..}^{cr} (X \longrightarrow Y)$ [Gruyer et Berge-Cherfaoui, 1999]	107
4.4	Matrice de croyance $M_{..j}^{cr} (X \longleftarrow Y)$ [Gruyer et Berge-Cherfaoui, 1999]	107
4.5	Matrice pignistique $BetP_{i..}$ (associations cible-à-piste) [Royère, 2002]	108
4.6	Matrice pignistique $BetP_{..j}$ (associations piste-à-cible) [Royère, 2002]	108
4.7	Matrice pignistique $BetP_{i..}$ du scénario illustré par la figure 4.6	111
4.8	Matrice pignistique duale	115
4.9	Matrice duale correspondante à la scène illustrée par la figure 4.7	116
4.10	Scénario simulé de [Mercier <i>et al.</i> , 2011] : 3 objets perçus vs. 4 objets connus	118
4.11	Probabilités pignistiques des associations $X \longrightarrow Y$	118
4.12	Probabilités pignistiques des associations $X \longleftarrow Y$	118
4.13	Résultats de la prise de décision	119
4.14	Probabilités pignistiques avec la classification du gradient	120
4.15	Principales propriétés des algorithmes suivi- <i>après</i> -détection sélectionnés [Møgelmoose <i>et al.</i> , 2012]	122
4.16	Comparaison entre le SMR et des algorithmes suivi- <i>après</i> -détection [Møgelmoose <i>et al.</i> , 2012]	122
4.17	Performances pour les séquences 1 et 2	123
4.18	Temps de calcul moyen	124
4.19	Evaluation de l'apport des algorithmes de décision : PPCJ, PPL, PPCG, et la décision duale ainsi que de la méthode GNN sur la détection de panneaux circulaires sur la séquence vidéo 1.	124
4.20	Temps de calcul moyen des algorithmes : PPCJ, PPL, PPCG, et la décision duale ainsi que de la méthode GNN.	125

Abréviations

ADAS	<i>Advanced Driver Assistance System</i> , Système Avancé d'Aide à la Conduite
AFIL	<i>Avertisseur de Franchissement Involontaire de Ligne</i>
AFU	<i>Assistance au Freinage d'Urgence</i>
BAS Plus	<i>Brake Assist System Plus</i>
BBM	<i>Basic Beliefs Mass</i> , masse de croyance élémentaire
CMBS	<i>Collision Mitigation Brake System</i>
DtB	<i>Distances to Borders</i>
DPR	<i>Détecteur de Polygones Réguliers</i>
FFT	<i>Fast Fourier Transform</i> , Transformée de Fourier Rapide
FK	<i>Filtre de Kalman</i>
FP	<i>Faux Positifs</i>
FPPI	<i>Faux Positifs Par Image</i>
GNN	<i>Global Nearest Neighbor</i> , Plus Proche Voisin Global
GPS	<i>Global Positioning System</i>
HOG	<i>Histogram of Oriented Gradients</i> , Histogramme de Gradients Orientés
HOMHT	<i>Hypothesis Oriented MHT</i> , MHT Orienté Hypothèse
HSI	<i>Hue-Saturation-Intensity</i> , Teinte-Saturation-Intensité
HSV	<i>Hue-Saturation-Value</i> , Teinte-Saturation-Valeur.
ILENA	<i>Intelligent Learning Navigation</i>
ITS	<i>Intelligent Transportation Systems</i> , Systèmes de Transport Intelligents
JPDA	<i>Joint Probabilistic Data Association</i> , Association Probabiliste Conjointe de Données
LIDAR	<i>LIght Detection And Ranging</i>
MCT	<i>Modèle des Croyances Transférables</i>
MHT	<i>Multiple Hypothesis Tracking</i> , Suivi à Hypothèses Multiples

NN	<i>Nearest Neighbor, Plus Proche Voisin</i>
OMS	<i>Organisation Mondiale de la Santé</i>
PDA	<i>Probabilistic Data Association, Association Probabiliste de Données</i>
PPCG	<i>Probabilité Pignistique Classée par Gradient</i>
PPJ	<i>Probabilité Pignistique Jointe</i>
PPJC	<i>Probabilité Pignistique Jointe Classée</i>
PPL	<i>Probabilité Pignistique Locale</i>
RFID	<i>Radio-Frequency IDentification, Identification Radio-Fréquence</i>
RGB	<i>Red-Green-Bleu, Rouge-Vert-Bleu</i>
ROC	<i>Receiver Operating Characteristic</i>
ROI	<i>Region Of Interest, Région d'Intérêt</i>
RSLD	<i>Ransac Symmetric Lines Detection</i>
SMR	<i>Suivi Multi-ROIs</i>
SIG	<i>Système d'Information Géographique</i>
STVUE	<i>Single Target Vote for Upright Ellipses</i>
STVUT	<i>Single Target Vote for Upright Triangles</i>
SVM	<i>Support Vector Machine, Séparateur à Vastes Marges</i>
TCB	<i>Transformée Chinoise Bilatérale</i>
TDPI	<i>Taux de Détection Par Image</i>
TDPP	<i>Taux de Détection Par Panneau</i>
TFP	<i>Taux de Faux Positifs</i>
TOMHT	<i>Track Oriented MHT, MHT Orienté Piste</i>
TSR	<i>Traffic Sign Recognition, Reconnaissance de Panneaux Routiers</i>
TrSR	<i>Transformation en Symétrie Radiale</i>
TSB	<i>Transformée en Sommet et Bissectrice</i>
VANET	<i>Vehicle Ad Hoc Network</i>
VP	<i>Vrais Positifs</i>
VPP	<i>Valeur Prédicative Positive</i>
V2I	<i>Vehicle-to-Infrastructure, Véhicule à Infrastructure</i>
V2V	<i>Vehicle-to-Vehicle, Véhicule à Véhicule</i>
WGS84	<i>World Geodetic System 1984</i>

Liste des publications

1. Journaux internationaux

- (a) Boumediene, M., Cudel, C., Basset, M. et Ouamri, A. (2013). Triangular traffic signs detection based on RSLD algorithm. *Machine Vision and Applications*, 24(8): 1721-1732.
- (b) Boumediene, M., Lauffenburger, J.-P., Daniel, J., Cudel, C. et Ouamri, A. (2014). Multi-ROI Association and Tracking With Belief Functions : Application to Traffic Sign Recognition. *IEEE Transaction on Intelligent Transportation Systems*, 15(6): 2470-2479.

2. Conférences internationales avec actes et comité de lecture

- (a) Boumediene, M., Lauffenburger, J.-P., Daniel, J. et Cudel, C. (2014). Coupled Detection, Association and Tracking for Traffic Sign Recognition. *In IEEE Intelligent Vehicle Symposium*, pages 1402-1407, Dearborn, Michigan, USA.
- (b) Lauffenburger, J.-P., Daniel, J. et Boumediene, M. (2014). Traffic Sign Recognition : Benchmark of Credal Object Association Algorithms. *In International Conference on Information Fusion*, Salamanca, Spain.

Introduction générale

De nos jours, le développement économique et social est tributaire de la mobilité des biens et des hommes. En effet, le transport joue un rôle important dans les échanges de marchandises ainsi que pour la circulation des personnes. Il existe cinq modes de transport : maritime, aérien, routier, ferroviaire, et fluvial. Cependant, le transport routier reste irremplaçable vu ses spécificités et son rôle vital dans notre quotidien. En plus, d'être accessible pour chacun et disponible quasiment partout, le transport routier est complémentaire à tous les autres modes. En effet, il intervient dans chaque opération de transport que ce soit au début et/ou à la fin. Par conséquent, le transport routier est le mode de transport le plus utilisé générant ainsi un trafic important qui ne cesse d'augmenter notamment dans les pays émergents. Malheureusement, cette augmentation du trafic routier a des conséquences sur la santé publique.

D'après l'Organisation Mondiale de la Santé (OMS), les accidents de la route représentaient la huitième cause de mortalité au monde en 2010 et devraient être la cinquième cause d'ici 2030. Les conséquences des accidents sont néfastes sur le plan humain, économique, et sociétal. En effet, en plus des souffrances physiques et psychologiques des victimes (décès, blessures et handicaps), les accidents occasionnent des pertes économiques. Les blessures et les lésions causées nécessitent souvent des soins coûteux (séjour à l'hôpital, transport des victimes, mobilisation des pompiers, etc.) et parfois de longues convalescences ce qui entraîne une baisse de productivité chez les victimes ainsi que leurs proches. Face aux dangers de la route, une multitude de recherches ont été menées dans le domaine automobile afin de réduire les accidents ainsi que leurs conséquences. Pour cela, la sécurité routière est organisée en trois niveaux : primaire, secondaire, et tertiaire. La sécurité primaire consiste à éviter au maximum les accidents. Dans le cas où ces derniers n'ont pas pu être évités, la sécurité secondaire intervient en minimisant leurs conséquences. Pour améliorer la prise en charge des victimes, la sécurité tertiaire travaille sur le développement des moyens de secours.

Beaucoup de pays ont réduit leur taux de mortalité sur les routes grâce à une politique de sécurité routière rigoureuse basée sur la sensibilisation et la répression. Cependant, avec tous les dispositifs mise en place le taux utopique de 0 morts n'est encore atteint dans aucun pays. Cela veut dire qu'une baisse significative des accidents est encore envisageable à travers le développement de systèmes d'aide à la conduite. En effet, 90% des accidents ont pour cause la défaillance humaine. C'est à partir de ce constat, qu'un grand nombre de chercheurs dans le domaine des Systèmes de Transport Intelligents (ITS - *Intelligent Transportation Systems*) travaillent sur l'assistance du conducteur afin de réduire le risque d'accident. C'est sur ce principe que les systèmes avancés d'aide à la conduite communément appelés "ADAS" (*Advanced Driver Assistance System*) sont développés. Ces derniers sont intégrés aux véhicules dits "*intelligents*" pour assister le conducteur durant sa tâche de conduite en lui assurant plus de sécurité et de confort.

Un conducteur peut être distrait de sa tâche principale, la conduite, ce qui occasionne un manque de vigilance vis à vis de la signalisation courante. Cette situation augmente le risque d'accident. En effet, manquer un panneau de limitation de vitesse ou d'interdiction de doubler génère une situation à risque en plus du fait d'exposer le conducteur à des sanctions. Afin de réduire ce risque, le système de reconnaissance de panneaux routiers (*TSR-Traffic Signs Recognition*) assiste le conducteur en le tenant informé de la signalisation liée à la route parcourue. Cet ADAS est usuellement basé sur une caméra embarquée sur l'avant du véhicule. Son principe consiste à analyser le flux vidéo fourni par la caméra afin de reconnaître les panneaux routiers présents dans la scène. Le système est basé sur deux principales étapes : la *Détection* et la *Reconnaissance* (aussi appelée classification). La première étape localise les panneaux dans les images en se basant sur leurs caractéristiques visuelles, comme la couleur et la forme. La deuxième étape reconnaît le type de chaque panneau détecté grâce à son pictogramme. Au final, les panneaux reconnus sont signalés au conducteur ou transmis à un autre ADAS. Il est possible d'ajouter une étape de *Suivi temporel* afin de ne pas re-signaliser au conducteur les mêmes panneaux reconnus. En effet, les panneaux apparaissent sur plusieurs images avant leur disparition du champs de vision de la caméra ce qui permet leur suivi. Ce processus de *Suivi temporel* est assuré par une étape d'association de données qui consiste à mettre en correspondance les panneaux connus (reconnus précédemment) avec les panneaux perçus (à l'instant courant).

Les travaux de recherche de cette thèse s'intéressent à l'aspect temporel qui peut être mieux exploité pour l'amélioration de la *Détection*. En effet, comme la *Détection* est imparfaite, des fausses détections peuvent être générées ce qui dégrade les performances d'un système TSR. L'idéal étant d'avoir un système ayant un fort taux de bonnes détections avec le plus faible nombre possible de fausses détections. C'est pourquoi, les objectifs de cette thèse sont multiples : l'amélioration de la *Détection* et l'exploitation de l'information temporelle afin d'éliminer les fausses détections. Les travaux de recherche menés ont abouti à deux contributions :

- une nouvelle méthode robuste pour la *Détection* de panneaux triangulaires et circulaires. La méthode proposée se base sur une classification des pixels selon les orientations de contours. Son principe consiste, dans un premier temps, à déterminer les positions candidates des panneaux à l'aide d'un détecteur de Harris. Ensuite, la présence d'un panneau est vérifiée à chaque position par la recherche d'une forme géométrique particulière. Une étape de classification des positions permet de distinguer les sommets des triangles des centres des cercles, ce qui permet de spécifier la forme géométrique à rechercher (triangle ou cercle).
- une nouvelle approche de *Suivi temporel* pour réduire le nombre de fausses détections. Ce suivi est basé sur la fusion des résultats de la *Détection* qui sont imprécis et incertains. Il est important d'en tenir compte de ces imperfections lors de ce processus. C'est pourquoi, l'approche proposée est basée sur la théorie des croyances. Cette dernière offre les mécanismes nécessaires pour la modélisation des données ainsi que de ses imperfections.

Le plan de ce manuscrit est divisé en deux parties principales : traitement d'images pour la *Détection* (chapitres 2 et 3) et la fusion de données pour le *Suivi temporel* (chapitres 2 et 4).

Le premier chapitre discute quelques statistiques mondiales et nationales sur les accidents de la route ainsi que les solutions proposées pour réduire leur nombre. Ces solutions peuvent intervenir à trois niveaux : primaire, secondaire, et tertiaire. Ensuite, une description plus détaillée est consacrée aux dispositifs sécuritaires du premier niveau, principalement ceux intégrés dans le véhicule afin d'aider le conducteur dans sa tâche de conduite. Tout d'abord, une brève description des ADAS est présentée ainsi que leur principe. Par la suite, le système TSR est présenté ainsi que ses étapes : *Détection*, *Reconnaissance*, et *Suivi temporel*.

Le chapitre 2 présente, en premier lieu, un état de l'art des récentes méthodes de *Détection* structuré selon trois méthodologies : approches couleurs, approches contours, et approches descripteurs. La deuxième partie du chapitre présente un état de l'art du *Suivi temporel* dans les systèmes TSR. Tout d'abord, les étapes majeurs du *Suivi temporel* sont décrites : le filtrage et l'association de données. Ensuite, une classification des récents systèmes TSR intégrant le *Suivi temporel* est proposée. Deux classes d'approches sont définies selon la séquentialité des étapes : suivi-*après*-reconnaissance et suivi-*après*-détection. Finalement, un résumé des trois approches majeures d'association de données est présenté : méthode du plus proche voisin, la théorie des probabilités, et la théorie des croyances.

Le chapitre 3 est consacré à la première contribution de cette thèse, modélisée dans une nouvelle méthode de *Détection* de panneaux triangulaires et circulaires. La première partie du chapitre décrit le principe de la méthode qui a été modélisé dans un algorithme dit "*Ransac Symmetric Lines Detection*" (RSLD). Ensuite, la deuxième partie du chapitre discute les performances de l'algorithme en les comparant avec celles des récentes méthodes.

Le chapitre 4 présente la deuxième contribution de cette thèse, formulée dans une approche originale de suivi-*après*-détection basée sur la théorie des croyances. La première partie du chapitre présente la modélisation de cette idée dans un algorithme dit "*Suivi Multi-ROIs*" (SMR) basé sur une association crédale. Cette dernière est issue de la théories des croyances présentée dans la deuxième partie du chapitre. Quant à la dernière partie, elle évalue les performances de l'algorithme SMR en le comparant avec les récents TSR intégrant un *Suivi temporel*.

Finalement, une conclusion générale est présentée, ainsi que quelques perspectives des travaux de recherche discutés dans ce manuscrit.

Chapitre 1

Contexte de recherche

1.1 Introduction

La mobilité est un facteur important dans le développement des économies modernes. En plus d'assurer les échanges économiques, elle facilite également la circulation des personnes et des services. Dans notre quotidien, l'automobile est le moyen le plus utilisé (pour se rendre au travail, accompagner ses enfants, faire ses courses, etc.). La vie moderne de transport contraint les personnes à se déplacer plus fréquemment, en parcourant de plus grandes distances (notamment pour des raisons professionnelles) en moins de temps, augmentant ainsi le risque d'accident de la route. En 2010, la route a tué près de 1,24 million de personnes dans le monde [OMS, 2013]. Ce constat alarmant attire l'attention de la communauté internationale sur la question de la sécurité routière. Par conséquent, des améliorations se développent à tous les niveaux (véhicule, infrastructure, formation des conducteurs, répression, etc.). Les acteurs de la filière automobile, constructeurs et équipementiers en tête, sont eux aussi interpellés par ce problème et intègrent de plus en plus de dispositifs sécuritaires lors de la conception des véhicules. Dans cette thèse, un de ces dispositifs est étudié : la reconnaissance de panneaux routiers dont l'objectif est d'assister le conducteur tout au long de son trajet en le tenant informé de la signalisation.

Ce chapitre propose, dans un premier temps, un bref constat sur la sécurité routière en présentant quelques statistiques mondiales et nationales ainsi que les solutions proposées pour réduire le nombre et les conséquences des accidents. Ces solutions peuvent intervenir à différents niveaux : primaire pour éviter les accidents ; secondaire afin de minimiser leurs conséquences ; et tertiaire pour une meilleure prise en charge des victimes. La deuxième partie du chapitre s'intéresse aux dispositifs sécuritaires du premier niveau, principalement ceux intégrés dans le véhicule afin d'aider le conducteur dans sa tâche de conduite. En premier lieu nous décrivons brièvement ces systèmes avancés d'aide à la conduite en présentant leur principe et en citant quelques exemples. Ensuite, nous nous intéressons aux systèmes de reconnaissance de panneaux routiers principalement ceux

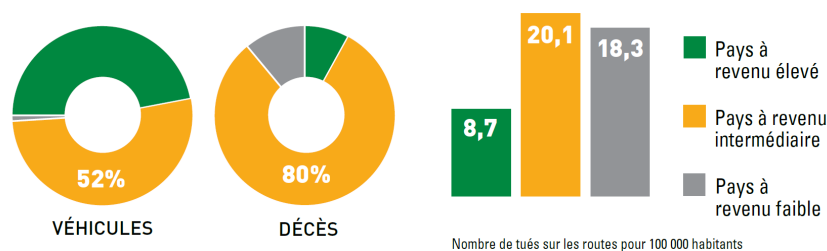


FIGURE 1.1 – Répartition des victimes tuées et du parc mondial de véhicules dans le monde [OMS, 2013].

basés sur la vision. Ces derniers utilisent habituellement une caméra pour la perception de l'environnement, l'objectif étant de détecter et de reconnaître les panneaux afin, soit d'informer le conducteur, soit de réguler la vitesse d'un véhicule autonome par exemple.

1.2 Sécurité routière

Selon le rapport de l'Organisation Mondiale de la Santé (OMS) sur la sécurité routière, les accidents de la route représentaient en 2010 la huitième cause de mortalité au monde et devraient être la cinquième cause d'ici 2030 [OMS, 2013] (l'Algérie n'a pas participé à cette étude). En plus de ce constat alarmant, le rapport révèle que les taux de mortalité les plus élevés sont observés dans les pays à revenu intermédiaire (cf. Figure 1.1). En effet, 80% des accidents mortels surviennent dans ces pays alors que ces derniers possèdent la moitié du parc automobile mondial. Ces chiffres sont principalement dûs au non-respect du code de la route. L'Algérie n'échappe pas à ce constat, avec un parc automobile de plus de 7 millions de véhicules et 4447 tués enregistrés en 2012 [CNPSR, 2012]. Les routes algériennes tuent 5 fois plus qu'en France et 8 fois plus qu'en Allemagne. Ces chiffres effrayants ont obligé les responsables à mettre en œuvre une politique de lutte par l'augmentation des sanctions et prochainement par l'instauration du permis à point.

Les statistiques de l'OMS démontrent aussi une différence entre les pays (à haut revenu et à revenu intermédiaire) par rapport aux catégories d'usagers tués. La figure 1.2 illustre la répartition des usagers décédés dans un accident de la route dans les pays suivants : la France, l'Allemagne (deux pays à haut revenu), le Maroc et la Tunisie (deux pays à revenu intermédiaire)¹. Nous observons que le taux des passagers (occupants du véhicule) décédés lors d'un accident ne dépasse pas les 15% en France et en Allemagne alors qu'il atteint 24% et 21% respectivement au Maroc et en Tunisie. En effet, dans ces pays, la législation sur l'utilisation des dispositifs tels que la ceinture de sécurité, reste très peu appliquée et la plupart du temps concerne uniquement les occupants avant. Un constat identique peut être fait concernant le pourcentage des piétons tués, il est de plus de 25% au Maroc et en Tunisie et moins de 14% en France et en Allemagne. Ces taux peuvent être expliqués par le fait que dans les pays à haut revenu, les aménagements pour piétons comme les

1. ce type de statistiques concernant l'Algérie n'est pas disponible



FIGURE 1.2 – Répartition des victimes tuées selon les catégories d'usagers [OMS, 2013].

trottoirs sont plus fréquents. Par ailleurs, les règles du code de la route sont davantage respectées par l'ensemble des usagers (piétons, cyclistes, et automobilistes).

En plus, des préjudices physiques et psychologiques que subissent les victimes (décès, blessures et handicaps), les accidents occasionnent des pertes économiques considérables [OMS, 2013]. Les blessures et les séquelles causées nécessitent souvent des soins coûteux (séjour à l'hôpital, transport des victimes, mobilisation des pompiers, etc.), et entraînent une baisse de productivité pour les victimes et leurs entourages. Face à ce constat, une multitude de recherches ont été menées dans le domaine automobile afin de réduire les accidents ainsi que leurs conséquences humaines, économiques, et sociétales. La sécurité routière est organisée en trois niveaux :

- **Sécurité primaire (active)** : permet d'éviter au maximum les accidents. Pour cela, la sécurité active intervient sur les trois composantes de la conduite automobile :
 - conducteur : par des formations visant à améliorer ses performances de conduite, la répression, une application plus rigoureuse du code la route, la sensibilisation aux dangers de la route, etc.
 - véhicule : par l'introduction de dispositifs de sécurité (les systèmes d'aide à la conduite) et la mise en place du contrôle technique.

- infrastructure : revêtement et marquage des routes, adaptation de la signalisation (panneaux) aux situations, éclairage du réseau routier, etc.
- **Sécurité secondaire (passive)** : lorsque l'accident ne peut être évité, elle consiste à minimiser ses conséquences. La sécurité passive intervient au niveau du véhicule et de l'infrastructure :
 - véhicule : la ceinture de sécurité (réduit le risque d'accident mortel de 50% pour les passagers assis à l'avant et de 75% pour les autres) et les airbags pour la protection des passagers, pare-chocs et airbag capot pour la protection des piétons, utilisation de matériaux déformables lors d'un choc pour absorber l'énergie cinétique, etc.
 - infrastructure : élimination des obstacles dangereux aux bords des routes, mise en place de barrières de sécurité, etc.
- **Sécurité tertiaire** : permet une meilleure prise en charge des victimes. L'objectif est de développer les services de secours et d'améliorer leur qualité afin de réduire les risques d'aggravation des lésions au cours du transport vers l'hôpital.

Beaucoup de pays ont pu réduire le taux de mortalité routière grâce à une politique de sécurité rigoureuse basée principalement sur une meilleure formation des jeunes conducteurs et l'augmentation des sanctions envers les conducteurs ne respectant pas le code de la route. Néanmoins, malgré toutes les campagnes de sensibilisation et les dispositifs de répression possibles, le taux utopique de 0 morts n'est encore atteint dans aucun pays. Le facteur humain étant de loin la première cause d'accidents de la circulation, de nombreux chercheurs dans le domaine des Systèmes de Transports Intelligents (ITS - *Intelligent Transportation Systems*) travaillent sur le développement d'aides à la conduite.

1.3 Systèmes avancés d'aide à la conduite

La majorité des constructeurs intègrent de plus en plus la notion d'anticipation des situations à risque lors de la conception d'un nouveau véhicule. En termes de sécurité routière, il vaut mieux détecter les situations potentiellement dangereuses afin d'éviter l'accident (sécurité primaire) que de minimiser sa gravité (sécurité secondaire). C'est sur ce principe que les systèmes avancés d'aide à la conduite communément appelés "ADAS" (*Advanced Driver Assistance System*) sont développés. Ces systèmes sont intégrés aux véhicules dits "intelligents" afin d'assister le conducteur durant sa tâche de conduite en lui assurant plus de sécurité et de confort. L'assistance apportée par les ADAS peut être *passive* en se limitant à de l'information du conducteur ou *active* en contrôlant partiellement ou totalement les commandes du véhicule afin d'éviter l'accident ou d'en limiter les conséquences dans le cas où ce dernier est inévitable.

1.3.1 Principe des ADAS

Un ADAS est considéré comme un dispositif de sécurité primaire basé sur quatre fonctions pour éviter l'accident [Salameh, 2011] :

- Percevoir le contexte de conduite : consiste à récolter un maximum d'informations sur les objets statiques et mobiles (autres véhicules, piétons, panneaux, etc.) de la scène routière ; ainsi que les informations sur le véhicule lui-même (comme sa vitesse, l'état de ses freins, et sa position 3D sur l'infrastructure routière). L'acquisition de ces informations nécessite l'intégration de capteurs extéroceptifs et proprioceptifs. La première catégorie perçoit l'environnement du véhicule sans aucune interaction physique avec l'infrastructure ou avec les autres véhicules. Ces capteurs de perception sont communément représentés par les caméras, les télémètres laser à balayage ou LIDAR (*Light Detection And Ranging*), et les radars. En plus de ces capteurs standards, il existe d'autres capteurs comme le récepteur GPS (*Global Positioning System*). Ce dernier détermine la position 3D du véhicule dans le repère terrestre (WGS84 - *World Geodetic System 1984*) par la réception des signaux multi satellitaires. La position déterminée peut alors servir à la localisation du véhicule sur une carte numérique du réseau routier. Les capteurs proprioceptifs renvoient des informations sur les paramètres (vitesse, accélération, etc.) du véhicule lui-même sans tenir compte de son environnement mais nécessaires à l'évaluation du risque. Parmi ces capteurs il y a l'odomètre et le gyromètre pour mesurer respectivement la distance parcourue et la vitesse angulaire.
- L'analyser et l'interpréter : à partir des informations perçues, il faut interpréter la scène routière afin de détecter les situations potentiellement dangereuses. L'analyse nécessite des algorithmes de détection et de reconnaissance des différents usagers (véhicules et piétons) présents sur l'espace de conduite ainsi que de la signalisation (panneaux routiers et marquages au sol) mise en place pour faciliter le partage de la route [Alshaqqi, 2014]. En plus de l'interprétation de l'environnement, le comportement du conducteur lui aussi est analysé à partir de certains paramètres comme le temps entre le lâché de l'accélérateur et l'enfoncement de la pédale de frein.
- Prendre une décision : à partir de l'évolution de l'environnement, un ou plusieurs critères de sécurité comme la vitesse limite et la distance de sécurité permettent une estimation du risque d'accident. Dans le cas où ce dernier dépasse un niveau acceptable, l'ADAS doit réagir afin de revenir à un état sûr.
- Agir afin d'éviter l'accident : une fois le risque identifié, le système entreprend l'action adéquate pour éviter l'accident et ramener le véhicule dans une situation normale. Trois niveaux d'action sont possibles [Ammoun, 2007] :
 - *Mode informatif* : le système informe uniquement le conducteur de la nature du risque (assistance passive).
 - *Mode suggestif* : en plus de signaler le risque au conducteur, l'ADAS lui propose une manoeuvre à entreprendre (comme freiner, ralentir, corriger sa trajectoire, etc.) afin de revenir à une situation sûre.

- *Mode actif* : le système contrôle partiellement ou totalement les commandes du véhicule contrairement aux deux premiers modes, dits passifs, où le conducteur reste le seul maître du véhicule.

1.3.2 Niveaux de coopération des ADAS

L'évolution des moyens de communication sans fil et leur pénétration dans notre quotidien a accéléré leur intégration dans le domaine de l'automobile [Ammoun, 2007]. Cette technologie a permis de repenser la façon de percevoir l'environnement. Actuellement, on distingue deux types d'ADAS : autonomes et coopératifs.

1.3.2.1 ADAS autonome

Ce type de systèmes perçoit l'environnement uniquement à l'aide des capteurs embarqués dans le véhicule. Le système n'interagit pas avec les autres véhicules ni avec l'infrastructure, ce qui lui procure une certaine autonomie. Néanmoins, sa perception de l'environnement est limitée par la portée et la précision des capteurs extéroceptifs (pas plus de 150 m [Ehrlich, 2009]).

1.3.2.2 ADAS coopératif

Pour repousser ces limites, une nouvelle approche émerge actuellement : la conduite coopérative. Grâce à l'intégration des moyens de communication sans fil dans l'infrastructure (panneaux et autres équipements de bord de route) et les véhicules, ces derniers contribuent à leur propre détection en transmettant des informations sur leur état (i.e identification et position) [Ehrlich, 2009]. En effet, cette intégration a permis le développement d'un nouveau type de réseau dit "réseau véhiculaire" (VANET - *Vehicle Ad Hoc Network*) où les communications peuvent être de type véhicule à infrastructure (V2I - *Vehicle-to-Infrastructure*) [Garcia-Garrido *et al.*, 2011] et/ou véhicule à véhicule (V2V - *Vehicle-to-Vehicle*) [Sepulcre *et al.*, 2013]. Deux types de technologies fonction de la portée des signaux sont employées : communications hertziennes locales ou RFID (*Radio-Frequency IDentification*) pour la portée moyenne et les communications hertziennes étendues pour la longue portée. En plus de repousser les limites de la perception, la conduite coopérative permet d'apporter (par transmission) des informations sur le comportement et les intentions des autres conducteurs permettant ainsi une meilleure interprétation de l'environnement.

1.3.3 Exemples d'ADAS autonomes

Dans cette thèse nous nous intéressons uniquement aux ADAS autonomes dont voici quelques exemples en cours de déploiement sur certains véhicules :



FIGURE 1.3 – L’Avertisseur de Franchissement Involontaire de Ligne (AFIL) de PSA Peugeot Citroën [Jean-Christophe RIAT, 2005].

- **Assistance au Freinage d’Urgence (AFU)** : des études ont montré que dans des situations d’urgence le conducteur ne freine pas de manière optimale. L’AFU a pour rôle de renforcer l’action du conducteur sur la pédale en cas de freinage d’urgence. Cette situation est détectée grâce à deux paramètres : l’intervalle de temps entre le lâcher de l’accélérateur et l’appui sur la pédale de frein ainsi que sa vitesse d’enfoncement. De plus, le “*Brake Assist System Plus*” (BAS Plus) de “*Mercedes-Benz*” mesure en continu, à l’aide de radars, la distance par rapport au véhicule précédent afin d’estimer l’assistance au freinage optimal [Daimler AG, 2013].
- **Avertisseur d’obstacles et de collisions** : ce système aide le conducteur à éviter les collisions en détectant les obstacles qui se trouvent devant lui sur la route, et en l’informant du danger estimé. La perception de l’environnement est basée sur des radars à longue et moyenne portée et/ou un LIDAR. Le système de prévention des collisions avec anticipation de freinage (CMBS - “*Collision Mitigation Brake System*”) de “*Honda*” [American Honda Motor, 2006] transmet un signal sonore au conducteur si la distance de sécurité par rapport au véhicule précédent est en-dessous d’un seuil acceptable. Dans le cas où la probabilité de collision ne diminue pas, le système entreprend un freinage modéré tout en rétractant les ceintures de sécurité grâce aux prétensionneurs électroniques afin de bien maintenir les passagers. En dernier recours le système déclenche le freinage d’urgence pour éviter la collision.
- **Système de surveillance de trajectoire latérale** : ce dispositif surveille la position latérale du véhicule sur sa voie. Si le conducteur s’écarte de sa trajectoire idéale, le système l’informe par un signal sonore et peut agir sur la direction pour corriger la trajectoire (cf. Figure 1.3). “L’Avertisseur de Franchissement Involontaire de Ligne” (AFIL) de “*PSA Peugeot Citroën*” utilise des capteurs infrarouges, implantés sous le bouclier avant, pour détecter les lignes blanches (marquages au sol continus ou discontinus) [Jean-Christophe RIAT, 2005]. Dans le cas où le conducteur s’approche trop des lignes (limites de la voie de roulement) sans avoir actionné son clignotant, le système provoque alors une vibration dans le siège (mode informatif) afin d’alerter le conducteur si jamais le franchissement est involontaire.
- **Système de reconnaissance de panneaux routiers** : manquer un panneau de limitation de vitesse ou d’interdiction de doubler génère une situation à risque en plus du fait d’exposer le conducteur à des sanctions. Le système de reconnaissance de panneaux assiste le conducteur en le tenant informé de la signalisation courante (cf. Figure 1.4). Le système peut être relié à un



FIGURE 1.4 – Le système ILENA (*Intelligent LEarning Navigation*) de BMW pour la navigation intelligente.

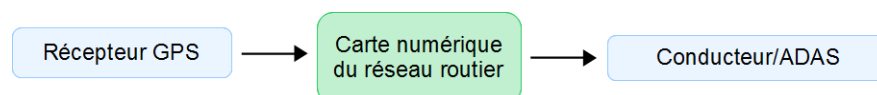


FIGURE 1.5 – Schéma général d'un système de navigation par satellite.

régulateur de vitesse pour maintenir une vitesse de croisière adaptée aux limitations et à une éco-conduite [Winckler, 2011].

1.4 Système de reconnaissance de panneaux routiers

La perception de l'environnement routier est primordiale pour un ADAS. Avoir des informations sur le contexte routier comme la signalisation permet d'informer le conducteur sur les limitations de vitesses par exemple. Le moyen le plus simple pour obtenir ce type d'information est l'utilisation d'un système de navigation par satellite [Ehrlich, 2009]. La figure 1.5 illustre son principe de fonctionnement. Le système est basé sur une carte numérique du réseau routier et une localisation GPS dans le repère terrestre. Dans un premier temps, le système détermine la position du véhicule à l'aide du récepteur GPS. Ensuite cette position est reportée sur la carte numérique (Système d'Information Géographique (SIG)) afin d'extraire les attributs statiques (type de route, nombre de voies, courbure, signalisation, etc.) de la route parcourue et de la route à venir. Finalement, le système informe le conducteur (mode informatif) de la signalisation courante (i.e vitesse limite) et peut aussi agir directement sur le véhicule (mode actif) via par exemple le régulateur de vitesse [Ehrlich, 2009].

1.4.1 Reconnaissance par système d'information géographique

Bien qu'il soit facile à mettre en œuvre pour fournir des informations sur la signalisation courante, le système de navigation a deux principaux inconvénients : la perte temporaire du signal GPS et l'aspect statique de sa carte numérique. En effet, dans les tunnels ou les ruelles étroites le signal GPS peut être temporairement perdu, et de ce fait la position du véhicule devient inconnue. Ce problème

peut être compensé par l'utilisation d'un odomètre et d'un gyromètre qui mesurent la distance et la vitesse angulaire de lacet du véhicule. En effet, ces deux mesures permettent d'estimer l'évolution de la position du véhicule durant une courte perte du signal GPS ce qui permet au système de rester quasi-opérationnel. Le deuxième problème est lié à la non-mise à jour de la carte numérique. En effet, la signalisation routière est modifiée en permanence par les autorités responsables, pour des raisons de sécurité et/ou d'écologie. La mise à jour des cartes numériques ne permet pas de suivre tous ces changements, de ce fait les informations fournies par la navigation (attributs liées à la signalisation) risquent de ne pas correspondre à la vérité terrain.

1.4.2 Reconnaissance par vision

Une solution consiste en l'utilisation d'un système de reconnaissance visuel de panneaux routiers (TSR-*Traffic Signs Recognition*). Ce système est usuellement basé sur un capteur de perception optique, i.e caméra, embarquée sur l'avant du véhicule. Son principe consiste à analyser le flux vidéo fourni par la caméra afin de détecter et de reconnaître les panneaux présents dans la scène.

L'approche par vision est basée sur deux principales étapes : la *Détection* et la *Reconnaissance* (aussi appelée classification) (cf. Figure 1.6). La première étape localise les panneaux dans les images. Cette étape exploite les caractéristiques visuelles des panneaux, comme la couleur et la forme, afin de les distinguer par rapport aux autres objets présents dans la scène. Cependant, en situations réelles la *Détection* n'est jamais parfaite. En plus des panneaux (*Vrais Positifs*) (VP), des fausses détections (*Faux Positifs*) (FP) peuvent être générées. Ces dernières seront rejetées normalement par le système lors de la *Reconnaissance* (cf. Figure 1.7). Cette deuxième étape reconnaît le type de chaque panneau détecté grâce à son pictogramme. Généralement, la *Reconnaissance* se base sur l'utilisation de classifieurs comme les Séparateurs à Vastes Marges (SVMs) où une phase d'apprentissage est nécessaire afin que les classifieurs apprennent les différents pictogrammes. A l'issue de l'étape de *Reconnaissance*, les panneaux non-reconnus sont ignorés par le système. Les panneaux reconnus sont quant à eux signalés au conducteur par un simple affichage sur son tableau de bord, ou transmis à un autre ADAS (cf. Figure 1.7). Puisque les panneaux apparaissent sur plusieurs images avant leur disparition, il est possible d'ajouter une étape de *Suivi temporel* au système (intégré souvent après la reconnaissance)(cf. Figure 1.6). Ce *Suivi temporel* identifie les panneaux déjà reconnus afin de ne pas les re-signaliser au conducteur ou tout simplement pour ne pas les re-classifier ce qui permet de réduire le temps de calcul. Ce processus est assuré par une étape d'association de données qui consiste à mettre en correspondance les panneaux connus (reconnus précédemment) avec les panneaux perçus (à l'instant courant). Il est aussi possible d'introduire un *a priori* (à l'aide d'une étape de filtrage) sur les positions et les résolutions des panneaux dans les images suivantes facilitant ainsi leur *Détection*.

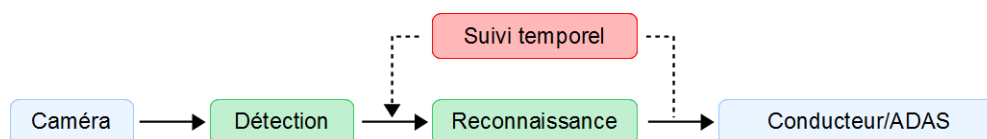


FIGURE 1.6 – Schéma général d'un système de reconnaissance de panneaux routiers.

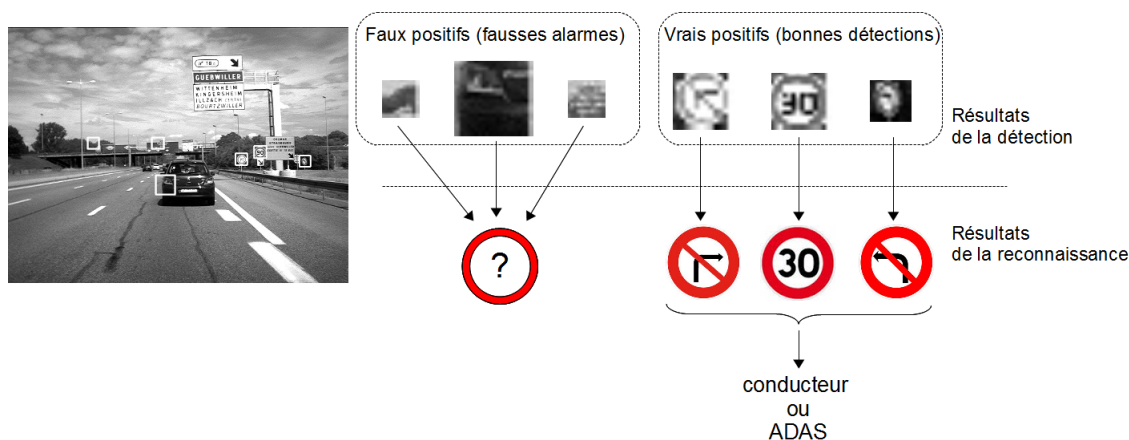


FIGURE 1.7 – Exemple de fonctionnement d'un système de reconnaissance de panneaux routiers.

TABLE 1.1 – Avantages et limites des deux techniques de reconnaissance : par vision et par SIG.

Approche	Avantages	Limites
Vision (caméra)	Reconnaît les panneaux perçus dans la scène (contexte locale), y compris les panneaux temporaires, s'adapte aux panneaux spécifiques de certains pays	Sensible aux mauvaises conditions de visibilité, difficulté de reconnaissance durant la nuit, ne reconnaît pas la totalité des panneaux (dangers, interdiction, etc.)
SIG (carte numérique)	Procure des informations globales sur le contexte routier, indépendant des conditions de visibilité, fonctionne de jour comme de nuit	La carte numérique ne correspond pas toujours à la signalisation courante, peut concerner que certains pays, son fonctionnement dépend du signal GPS

L'approche par vision dépend des conditions de visibilité. En effet, les performances de la *Détection* comme de la *Reconnaissance* diminuent lorsque la visibilité est dégradée. À partir de ce constat, des dispositifs s'appuyant sur la fusion des deux approches (par SIG et vision) ont été proposés afin d'obtenir des systèmes robustes [Daniel et Lauffenburger, 2013].

En effet, les deux approches par vision et par navigation sont complémentaires (cf. Table 1.1). Le système TSR reconnaît seulement les panneaux perçus dans la scène ce qui représente une information locale ou temporaire, contrairement à la navigation qui elle fournit des informations globales sur le contexte de conduite [Puthon, 2013]. La situation de conduite représentée dans la figure 1.8 illustre très bien cette complémentarité. Au niveau de la sortie d'autoroute, le système de navigation peut extraire à partir de sa cartographie plusieurs panneaux de limitation de vitesse, e.g. un panneau de 50 *km/h* destiné à la bretelle de sortie et un autre de 110 *km/h* lié à la

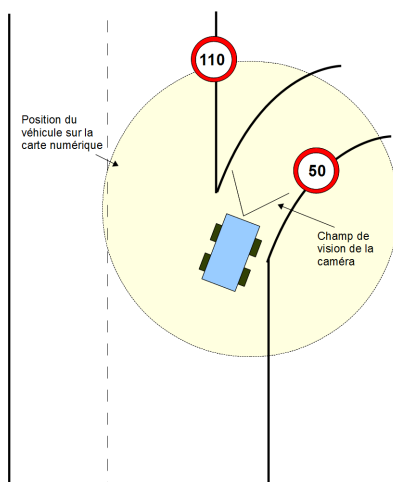


FIGURE 1.8 – Exemple de reconnaissance de panneaux au niveau d’une sortie d’autoroute.

limite sur l’autoroute, alors que le système par vision reconnaît uniquement le panneau perçu à ce moment là (50 km/h). En plus de la complémentarité, la fusion permet aussi d’obtenir une redondance d’information utile au renforcement de la prise de décision (i.e. la signalisation qui faut transmettre au conducteur) et à la robustesse du système. [Daniel et Lauffenburger, 2013] se basent sur la théorie des croyances pour la fusion des deux approches par navigation et par vision. L’avantage de cette théorie est qu’elle offre des mécanismes pour la représentation de l’incertitude et de l’imprécision des informations liées à la perception de l’environnement.

1.5 Contributions

Les travaux de recherche décrits dans ce manuscrit concernent les systèmes de reconnaissance de panneaux routiers par vision (perception locale du contexte de conduite) puisque nous estimons que l’aspect temporel peut être mieux exploité pour améliorer la *Détection*. En effet, plusieurs objets ayant des caractéristiques visuelles similaires aux panneaux peuvent être présents, la couleur rouge et la forme circulaire par exemple ce qui génère un certain nombre de FP. La densité de la scène ajoute aussi sa part de difficulté puisque les panneaux peuvent être temporairement occultés ou très proches des autres objets, ce qui ne facilite pas leur distinction. Dans ce contexte, la *Détection* n’est jamais parfaite et un nombre important de FP dégrade les performances d’un système TSR souvent mesurées par le Taux de Détection Par Image (TDPI) et le nombre de FP Par Image (FPPI). La première mesure, TDPI, évalue les VP par rapport à la vérité terrain. C’est-à-dire un TDPI de 100% signifie que les panneaux sont détectés dans toutes les images où ils apparaissent (1.1). Quant à la deuxième mesure, FPPI, donne une idée sur la véracité des résultats de la *Détection* obtenus sur une image (1.2). Un faible FPPI signifie que le détecteur est précis et ne génère que très peu de FP. Par conséquent, l’idéal étant d’avoir un système ayant un fort TDPI avec un faible FPPI. Outre ces mesures standards, il y a le Taux de Détection Par Panneau (TDPP) (1.3). Cette mesure évalue

le taux de panneaux détectés tout au long du trajet parcouru (séquence vidéo). Une détection est considérée comme correcte si un panneau est transmis à l'étape de *Reconnaissance* plus de $N = 3$ fois successives.

$$TDPI = \frac{\text{Nombre de VP}}{\text{Vérité terrain}} \cdot 100 \quad (1.1)$$

$$FPPI = \frac{\text{Nombre de FP}}{\text{Nombre d'images}} \quad (1.2)$$

$$TDPP = \frac{\text{Nombre de détections correctes}}{\text{Nombre total de panneaux}} \cdot 100 \quad (1.3)$$

Les contributions de cette thèse concernent d'une part la *Détection* des panneaux et d'autre part leur *Suivi temporel*. La première contribution se résume en une nouvelle méthode robuste pour la *Détection* de panneaux triangulaires et circulaires (cf. Figure 1.6). La méthode, présentée dans le chapitre 3, se base sur un codage de contours. Afin de réduire le temps de traitement, les positions candidates des panneaux sont définies à l'aide d'un détecteur de Harris. Ces positions sont classées en deux classes : sommets des triangles et centres des cercles. Ensuite, la *Détection* d'un panneau consiste à vérifier la présence de la forme géométrique appropriée (triangle ou cercle) à chaque position candidate (sommet de triangle ou centre de cercle). La reconnaissance des formes géométriques est basée sur la classification des pixels. Le chapitre 3 présente en détails cette approche et discute ses performances.

La deuxième contribution concerne le problème des FP. Nous considérons les résultats de la *Détection* comme des informations imprécises et incertaines qui peuvent générer de mauvaises interprétations de la scène routière et influent ainsi les performances du système TSR. Pour cela, nous proposons de fusionner ces informations afin de réduire le nombre de FP. Cette fusion est assurée par l'ajout d'une étape de *Suivi temporel* après la *Détection* ce qui n'est pas très fréquent pour les système TSR (cf. Figure 1.6). L'idée est d'exploiter la redondance de la signalisation. En effet, les panneaux, après leur apparition dans la scène, restent visibles sur plusieurs images avant leur disparition du champ de vision de la caméra ce qui permet de les distinguer par rapport aux FP. Notre approche de *Suivi temporel* est basée sur la théorie des croyances qui prend en considération en plus de la connaissance, la méconnaissance de l'environnement (les imperfections de la *Détection*). Le chapitre 4 décrit l'étape du *Suivi temporel* proposée et discute son impact sur le système TSR.

Ces contributions ont permis la rédaction de deux articles dans des revues internationales à comité de sélection :

- Boumediene, M., Cudel, C., Basset, M. et Ouamri, A. (2013). Triangular traffic signs detection based on RSLD algorithm. *Machine Vision and Applications*, 24(8):1721-1732.

- Boumediene, M., Lauffenburger, J.-P., Daniel, J., Cudel, C. et Ouamri, A. (2014). Multi-ROI Association and Tracking With Belief Functions : Application to Traffic Sign Recognition. *IEEE Transaction on Intelligent Transportation Systems*, 15(6):2470–2479.

1.6 Conclusion

Le taux alarmant de mortalité enregistré sur les routes suscite l'inquiétude de la communauté internationale. Tous les gouvernements sont appelés à redoubler d'efforts en matière de sécurité routière (primaire, secondaire, et tertiaire) afin de mettre un frein à ce phénomène qui engendre des conséquences néfastes sur le plan personnel, économique, et sociétal. Ce constat a obligé les industriels automobiles à intégrer de plus en plus les dispositifs de sécurité lors de la conception de nouveaux véhicules. Parmi ces dispositifs il y a les systèmes avancés d'aide à la conduite (ADAS) qui suscitent un intérêt particulier. L'objectif de ces derniers est d'assister le conducteur durant sa tâche de conduite afin d'éviter les situations potentiellement dangereuses, réduisant ainsi le risque d'accident. Dans cette thèse nous nous intéressons au système de reconnaissance de panneaux routiers (TSR) par vision qui assiste le conducteur en le tenant informé de la signalisation liée à la route sur laquelle il circule. Le système perçoit l'environnement via une caméra embarquée et reconnaît les panneaux dans la scène de conduite selon un processus composé de deux principales étapes : *Détection* et *Reconnaissance*. Dans cette thèse nous améliorons la qualité de la *Détection* en exploitant davantage l'aspect temporel. En effet, les panneaux restent visibles un certain temps avant de disparaître du champ de la caméra. Par conséquent, il est utile d'utiliser cette redondance d'information lors de l'étape de *Détection*.

La première partie des travaux de recherche concerne l'étape de *Détection*. Cette étape est importante au système puisqu'elle localise les panneaux, avant de les reconnaître, dans le flux d'images fourni par la caméra. La *Détection* se base sur les caractéristiques visuelles : couleur et forme afin de distinguer les panneaux des autres objets de la scène routière. Cette dernière peut être assez encombrée et les panneaux peuvent être occultés ou très proches des autres objets, ce qui ne facilite pas la tâche de *Détection* et génère des FP. Plusieurs approches combinent les deux aspects visuels : couleur et forme pour améliorer les performances du système au détriment parfois du temps de traitement. Or, l'approche adoptée dans cette thèse se base uniquement sur la forme des panneaux tout en procurant des performances similaires à ceux des récents algorithmes.

La deuxième partie décrit l'intégration de l'information temporelle dans le système afin de réduire les FP. Cela consiste à fusionner les résultats de la *Détection* par l'ajout d'une étape de *Suivi temporel*. Dans la plupart des systèmes TSR, le *Suivi temporel* est intégré après l'étape de *Reconnaissance* ce qui permet de confirmer la présence d'un panneau reconnu. Par contre, la démarche faite dans cette thèse est d'intégrer l'étape de *Suivi temporel* après la *Détection* afin de supprimer les FP et de ne pas les transmettre à l'étape de *Reconnaissance*. L'approche adoptée pour le *Suivi temporel*

se base sur la théorie des croyances. Cette dernière offre les mécanismes nécessaires pour modéliser les imperfections des informations à fusionner comme les résultats de la *Détection* qui peuvent être imprécis et incertains.

Chapitre 2

Détection et suivi de panneaux routiers : état de l'art

2.1 Introduction

On pourrait penser que la *Détection* de panneaux est une tâche simple, se basant sur l'apparence visuelle, constituée d'une couleur et d'une forme particulière. Néanmoins, la résolution, les problèmes d'occultation, ainsi que les variations climatiques et d'intensité de lumière rendent cette opération non-triviale. Sans l'information de profondeur de la scène, il est impossible de déterminer la résolution des panneaux à partir des images 2D, ce qui ne facilite pas leur détection. La scène routière peut être assez encombrée et les panneaux risquent d'être occultés ou très proches des autres objets de la scène, ce qui influe sur les performances de la *Détection* souvent mesurées avec le taux de détection (TDPI et/ou TDPP) et le nombre de FP. La *Détection* doit prendre en considération différents facteurs comme la diversité des catégories de panneaux (panneaux d'interdiction, panneaux de danger, panneaux d'obligation, panneaux d'indication, panneaux de cédez le passage et d'intersection) ; le changement d'éclairage et le faible contraste qui rendent la distinction difficile entre les panneaux et l'arrière plan ; et le changement d'apparence dû à une altération de la couleur par exemple au fil du temps. Avec ces facteurs, la *Détection* n'est jamais parfaite et sans erreur. Les récents algorithmes tentent d'augmenter le taux de détection tout en diminuant le nombre de FP. En effet, ces derniers sont considérés comme des hallucinations nuisibles au système TSR. Traiter un nombre important de fausses détections (données inutiles), durant la *Reconnaissance*, peut dégrader les performances du système et augmente le temps de traitement. C'est pourquoi, le développement des systèmes TSR plus robustes reste encore d'actualité.

Outre l'aspect visuel, les panneaux se caractérisent par une redondance dans la séquence d'images. En effet, les panneaux routiers apparaissent sur plusieurs images avant leur disparition du champ de la caméra et du conducteur. Dans ce contexte, il est intéressant d'exploiter cette redondance,

via un processus de fusion, afin d'améliorer les taux de détection et/ou de classification tout en réduisant le FP et/ou le taux d'erreur de classification. Ce processus de fusion se traduit par l'ajout d'une étape de *Suivi temporel* au système TSR. Intégré souvent après l'étape de *Reconnaissance*, le *Suivi temporel* identifie les panneaux déjà reconnus afin de ne pas les re-classifier ce qui réduit le temps de traitement. Cette identification est assurée par une étape d'association de données qui met en relation les panneaux connus (reconnus précédemment) avec les panneaux perçus (détecté à l'instant courant). L'association permet également de ne pas re-signaler les mêmes panneaux au conducteur afin de ne pas le perturber. À l'aide d'une étape de filtrage, le *Suivi temporel* peut également procurer un *a priori* sur les positions et les résolutions des panneaux dans les images suivantes facilitant ainsi leur *Détection*. Néanmoins, malgré tous ces avantages l'aspect temporel reste très peu exploité dans les systèmes TSR. Dans la littérature, peu de travaux se sont focalisés sur le *Suivi temporel* de panneaux afin d'améliorer les performances du système TSR.

La première partie de ce chapitre présente un état de l'art de la *Détection* de panneaux basé sur les récentes avancées du domaine. Les méthodes proposées peuvent être classées dans trois catégories selon la principale caractéristique visuelle (couleur, forme, apparence) utilisée : approches couleurs, approches contours, et approches descripteurs. La deuxième partie du chapitre est consacré au *Suivi temporel* dans les systèmes TSR. Tout d'abord le principe de la fusion de données est présenté ainsi que les imperfections de l'information observées dans un contexte réel et dont il faudra tenir compte lors du processus de fusion. Ensuite, les étapes du *Suivi temporel* sont décrites : le filtrage et l'association de données. Par la suite, nous présentons un état de l'art des récents systèmes TSR intégrant une phase de *Suivi temporel*. Ces systèmes peuvent être classés dans deux catégories selon la séquentialité des étapes : suivi-*après*-reconnaissance et suivi-*après*-détection. Finalement, un résumé des trois approches majeures d'association de données est présenté : méthode du plus proche voisin, la théorie des probabilités, et la théorie des croyances.

2.2 Détection de panneaux : état de l'art

Dans un état de l'art publié récemment, [Møgelmoose *et al.*, 2012] résument les stratégies de *Détection* de panneaux en trois étapes : *Segmentation*, *Extraction de caractéristiques* et *Détection*. Cette description est illustrée par la figure 2.1. L'étape de *Segmentation* exploite l'information couleur afin de délimiter la zone de recherche des panneaux ce qui permet de diminuer le temps de traitement. Les zones ayant une couleur similaire aux catégories de panneaux recherchés sont marquées comme des zones candidates. La deuxième étape consiste à extraire les caractéristiques discriminantes permettant de distinguer les panneaux des autres objets de la scène routière. La dernière étape vérifie la présence de panneau dans la zone restreinte de recherche en utilisant les caractéristiques extraites. Cependant, ce schéma général ne s'applique pas à toutes les approches décrites dans la littérature. Par conséquent, dans cette thèse nous avons choisi de présenter l'état

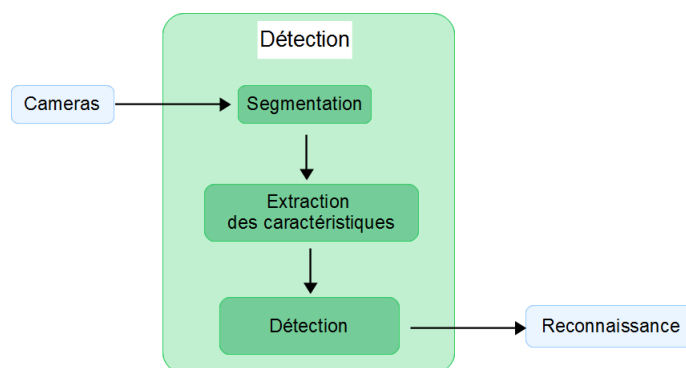
FIGURE 2.1 – Diagramme de l'étape de *Détection*.

TABLE 2.1 – Avantages et limites des différentes approches de détection de panneaux.

Approches	Intérêts	Limites
Couleurs	La couleur est très discriminante et facile à exploiter. Un simple seuillage permet d'extraire les régions dans l'image ayant une couleur similaire à celle des panneaux recherchés.	Les conditions de visibilité et la dégradation des surfaces réfléchissantes des panneaux influent la perception de la couleur et par conséquent les performances de la détection.
Contours	Plus robustes que les approches couleurs aux changements d'éclairage et peuvent être utilisées durant la nuit.	Les occultations partielles rendent la reconnaissance des formes difficile voir impossible dans certains cas.
Descripteurs	Décrivent l'apparence globale du panneau par des descripteurs plus discriminants que la couleur et/ou les contours ce qui procure souvent de meilleurs résultats.	Nécessitent une étape d'extraction de caractéristiques ce qui génère des calculs supplémentaires non négligeables pour une application temps réel tel que le TSR. De plus, le développement d'un tel détecteur nécessite une phase d'apprentissage du classifieur sur une base d'exemples positifs et négatifs.

de l'art de façon différente, en se basant sur les approches visuelles utilisées : couleur, forme, ou apparence globale.

Un panneau est caractérisé par sa couleur, sa forme et son pictogramme. Dans la littérature, l'information couleur est largement exploitée et tient son avantage du fait qu'un panneau se distingue facilement via sa couleur des autres objets de la scène routière. Le principe des approches couleurs est de rechercher des régions dans l'image ayant une similarité de teinte avec une ou plusieurs catégories de panneaux, par exemple la couleur rouge des panneaux d'interdiction et de danger. Une autre caractéristique discriminante des panneaux est la forme, qui est exploitée dans les approches contours. Le principe est de reconnaître les objets ayant une forme géométrique particulière, par exemple la forme triangulaire pour les panneaux de danger. Autre aspect exploité pour la *Détection* est l'apparence du panneau. Ceux sont des approches basées sur des descripteurs, caractérisant de façon directe ou indirecte l'apparence d'un panneau. La *Détection* dans ce cas revient à une mesure de similarité entre des descripteurs types et les descripteurs des potentiels panneaux. Chacune des

approches citées précédemment possèdent des avantages et des inconvénients (cf. Table 2.1). L'information couleur qui paraît assez discriminante sur le plan théorique peut être très sensible dans certaines situations réelles. En effet, les conditions d'illumination (jour, soleil, pluie, nuit, etc.) influent sur la perception de la couleur et rendent la détection des panneaux difficile. L'érosion climatique dégrade également la surface réfléchissante des panneaux, ce qui modifie légèrement leurs couleurs et influe donc sur les performances des algorithmes. D'un autre côté, la forme et l'apparence des panneaux sont plus robustes aux changements d'éclairage mais sont plus sensibles aux occultations partielles.

La table 2.2 regroupe les principales méthodes de détection de panneaux ainsi que leurs performances. Ces méthodes détectent un ou plusieurs types de panneaux (circulaire (cir.), triangulaire (tri.), octogonal (oct.), et rectangulaire (rec.)) et sont évaluées sur des données (Images d'évaluation) non communes ce qui rend leur comparaison très difficile. De plus, la plupart des articles traitant la même approche présentent les performances avec des mesures différentes. C'est pourquoi, dans la suite de cette section, nous présentons le principe de ces méthodes et nous discutons leurs performances sans effectuer de comparaison directe entre celles-ci.

TABLE 2.2 – Description et performances des principales méthodes de détection de panneaux.

Référence	Description de la méthode	Panneaux détectés	Images d'évaluation	Performances
		Nombre	Résolution	
Approches Couleurs				
[Piccioli <i>et al.</i> , 1996]	Seuillage sur l'espace HSV	600	512 × 512	TDPI= 95%, FPPI= 0.02
[Broggi <i>et al.</i> , 2007]	Seuillage sur l'espace RGB combiné avec une méthode de <i>template-matching</i>	-	-	TDPI= 92% avec FPPI= 0.13
[Nguwi et Kouzani, 2008]	Perceptron multi-couches pour la classification des résultats de la segmentation	164	-	TDPI= 9.45%
[Ruta <i>et al.</i> , 2011]	L'algorithme Quad-tree pour l'extraction des ROIs rouges	> 8000	-	-
[Qingsong <i>et al.</i> , 2010]	Seuillage sur l'espace HSI combiné avec analyse morphologique	85	640 × 480	TDPI= 68.42% avec FPPI= 0.09
[Gómez-moreno <i>et al.</i> , 2010]	Segmentation basée sur les SVMs	-	-	-
[Timofte <i>et al.</i> , 2014]	Seuillage adaptatif et transformée de Hough	> 15000	-	TDPI= 95%
Approches Contours				
[Loy et Barnes, 2004]	Détecteur de Polygones Réguliers (DPR)	45	-	TDPI= 97%, FPPI= 0.55
[Barnes et Zelinsky, 2004]	La Transformation en Symétrie Radiale (TrSR)	1107	-	TDPI= 90%

TABLE 2.2 – Suite

Référence	Description de la méthode	Panneaux détectés		Images d'évaluation		Performances
		Nombre	Résolution	Nombre	Résolution	
[Ruta <i>et al.</i> , 2011]						
[Belaroussi et Tarel, 2009]	La Transformée en Sommet et Bissectrice (TSB)	Tri.		48	320 × 270	TDPI= 95%, FPPI= 0.04
[Belaroussi et Tarel, 2010]	La Transformée Chinoise Bilatérale (TCB)	Cir.		89	640 × 480	TDPI= 86%, FPPI= 0.28
[Houben, 2011]	Méthodes de vote cible unique (STVUT et STVUE)	Tri. et cir.		847	1920 × 1080	TDPI= 81% (Tri.), TDPI= 91% (Cir.)
						[iTownns, 2009]
Approches Descripteurs						
[Lafuente-Arroyo <i>et al.</i> , 2005]	Classification des distances aux bords (DtB) par SVM	Tri. et cir.		300	-	TDPI= 97%, FPPI > 2
[Gil Jiménez <i>et al.</i> , 2008]	Analyse de la périodicité des bords par la FFT	-		-	-	-
[Baró <i>et al.</i> , 2009]	Cascade attentionnelle combinée aux dipôles dissociés	-		-	-	-
[Creusen <i>et al.</i> , 2010]	Classification des HOG couleurs par SVM	Tri. et cir.		3000	-	TDPI= 72%
[Landesa-Vzquez <i>et al.</i> , 2010]	Classification des descripteurs par cascade attentionnelle	Tri. et cir.		-	-	TDPP= 97.15%
[Deguchi <i>et al.</i> , 2011]	Cascade attentionnelle basée sur un tissage adaptatif	Cir.		3000	640 × 480	TDPI= 90%

TABLE 2.2 – Suite

Référence	Description de la méthode	Panneaux détectés		Images d'évaluation		Performances
				Nombre	Résolution	
[Šegvić <i>et al.</i> , 2014]	Cascade attentionnelle pour classifier les caractéristiques de Haar	Tri.	-	-	-	TDPI= 96%
[Overett <i>et al.</i> , 2014]	Une analyse discriminante linéaire pour classifier les HOG	Tri. et cir.	3300	(exemples positifs)		Taux de reconnaissance de 99%
[Zaklouta et Stanciu-lescu, 2014]	Classification des HOG couleurs par SVM	Tri. et cir.	1584	752 × 480		TDPI= 90%

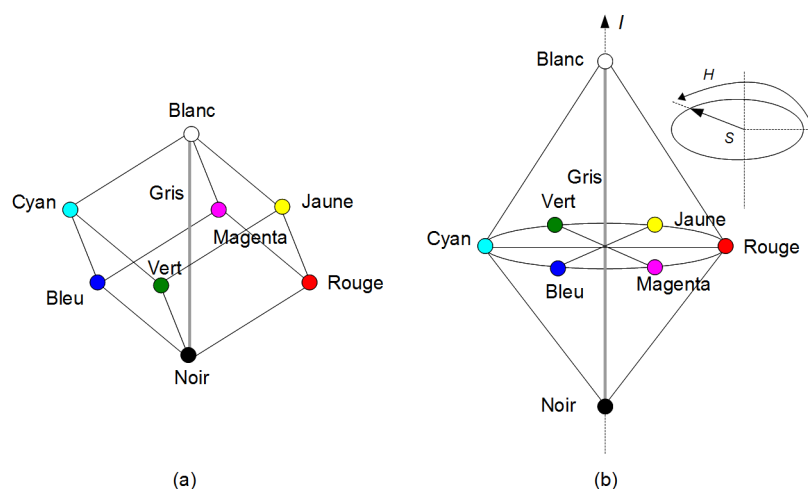


FIGURE 2.2 – Représentation de l'espace HSI. (a) représentation conique de l'espace RGB et (b) les composantes des espaces HSI.

2.2.1 Approches couleurs

L'aspect visuel est important lors de la conception d'un panneau. Chaque catégorie de panneaux est caractérisée par sa forme et sa couleur, facilitant ainsi sa distinction. L'information couleur est souvent exploitée afin de localiser la position des panneaux dans l'image. Dans [Broggi *et al.*, 2007], et [Timofte *et al.*, 2014] les régions d'intérêt (*Region Of Interest* - ROI), censées contenir un panneau, sont extraites par un simple seuillage de l'image couleur. Ce procédé permet l'extraction de plusieurs ROIs correspondant aux différentes couleurs de panneaux, par exemple le rouge, le bleu, et le jaune. Différents espaces de couleurs ont été utilisés. [Broggi *et al.*, 2007] utilisent l'espace de couleur RGB (*Red-Green-Blue* - Rouge-Vert-Bleu) pour l'extraction des ROIs suivie d'une reconnaissance de forme afin d'identifier la catégorie de panneaux. [Ruta *et al.*, 2011] proposent l'algorithme *Quad-tree* pour l'extraction des ROIs. L'image est divisée en 4 blocs de manière récursive jusqu'à une taille minimum. A chaque itération, les pixels ayant une similarité à la couleur de la catégorie de panneaux recherchés sont comptabilisés dans les 4 blocs. Si le nombre de pixels retenus est supérieur à un seuil, le bloc définit une ROI. Cependant, l'espace RGB est considéré comme un espace sensible aux variations de lumière et donc déconseillé pour l'environnement à l'extérieur. [Timofte *et al.*, 2014] proposent un seuillage adaptatif aux conditions d'éclairage pour remédier à la sensibilité de RGB. Dans d'autres travaux, l'espace RGB a été remplacé par des espaces plus robustes comme HSI (*Hue-Saturation-Intensity* - Teinte-Saturation-Intensité) ou HSV (*Hue-Saturation-Value* - Teinte-Saturation-Valeur).

Les espaces HSI et HSV sont dérivés de l'espace RGB et représentés par deux cônes accolés par la base (cf. Figure 2.2). Ces espaces décrivent mieux les composantes naturelles de la couleur en faisant intervenir des critères psychophysiques :

- la teinte H caractérise la couleur elle même ;
- la saturation S représente le degré de mélange de la couleur pure prédominante avec le blanc ;



FIGURE 2.3 – Exemple de segmentation dans l’espace HSI [Maldonado-bascón *et al.*, 2007]. (a) et (d) Images originales de la scène routière. (b) et (e) Respectivement segmentation de la couleur bleue et rouge. (c) et (f) Suppression des régions trop petites ou trop grandes.

– la luminance I ou la valeur V décrit l’aspect clair ou foncé de la couleur.

Les espaces HSI et HSV sont assez proches et constituent une alternative à l’espace RGB pour modéliser la perception humaine. Ces deux espaces sont souvent utilisés dans les systèmes TSR [Nguwi et Kouzani, 2008], [Qingsong *et al.*, 2010], [Paclíka *et al.*, 2000], et [Piccioli *et al.*, 1996]. Le seuillage peut être appliqué sur l’ensemble des composantes de l’espace comme dans [Nguwi et Kouzani, 2008] ou seulement sur les composantes chromatiques, teinte et saturation. [Nguwi et Kouzani, 2008] proposent un réseau de neurones constitué d’un perceptron multi-couches pour la classification des régions segmentées. Le perceptron classe ces régions en deux classes possibles : la classe panneau et la classe non-panneau. Les panneaux détectés sont par la suite reconnus par les SVMs. L’évaluation de la *Détection* n’est pas basée sur une base de données publique et n’est pas détaillée, les auteurs donnent uniquement le nombre de panneaux détectés et aucune information sur le nombre de FP n’est apportée. [Maldonado-bascón *et al.*, 2007] proposent un seuillage sur les composantes chromatiques pour localiser les groupes de pixels, composantes connexes ou blobs, rouges et bleus. Les blobs ne correspondant pas à l’échelle des panneaux recherchés sont supprimés. La figure 2.3 illustre les résultats de ce principe où les blobs trop petit ou trop grand sont supprimés. Cependant, l’espace HSI souffre d’une instabilité avec les couleurs achromatiques, comme le blanc [Gómez-moreno *et al.*, 2010]. Par conséquent, il ne peut être utilisé pour segmenter le fond blanc caractérisant la majorité des panneaux, panneaux de danger et d’interdiction. Face à cet inconvénient [Maldonado-bascón *et al.*, 2007] utilisent une décomposition achromatique [Liu *et al.*, 2002] pour détecter le fond de ces panneaux.

[Piccioli *et al.*, 1996] définissent des zones de recherches en regroupant des blocs de 16×16 pixels

dans l'espace HSV. Une étape de reconnaissance de forme appliquée sur ces zones permet l'extraction des ROIs représentant les panneaux triangulaires et circulaires. Seulement les composantes H et S sont exploitées pour la sélection des blocs qui est basée sur une mesure de similarité entre les pixels et la couleur rouge des panneaux. L'évaluation de l'approche a été effectuée sur une banque d'images des auteurs d'une résolution de 512×512 pixels. La méthode proposée a permis la détection de 92% des panneaux triangulaires sur 600 images, contenant chacune au moins un panneau, avec seulement 11 FP. Sur 400 images, 95% des panneaux circulaires ont été détectés avec 52 FP. [Paclíka *et al.*, 2000] se basent uniquement sur la composante H pour la détection des couleurs chromatiques, le rouge et le bleu, des panneaux. Les couleurs achromatiques sont détectées par un seuillage sur les composantes S et V . Les auteurs n'ont pas abordé les performances de la *Détection*.

La sélection des blobs peut être définie comme un processus de classification où chaque pixel est classifié sur plusieurs groupes ou classes. Par conséquent, les techniques de classification ont été appliquées à la segmentation. [Gómez-moreno *et al.*, 2010] proposent une segmentation couleur basée sur les SVMs. Contrairement à la technique de seuillage, la classification ne nécessite pas une adaptation des seuils manuellement mais via une phase d'apprentissage qui permet au classifieur d'apprendre la couleur recherchée. Ensuite, la forme des blobs est identifiée afin de préciser le type de panneau.

Bien que la couleur demeure un des aspects pertinents pour la *Détection* de panneau, elle reste sensible aux conditions d'éclairage, notamment durant la nuit. La surface réfléchissante des panneaux influe elle aussi sur la perception des couleurs par le système. Le niveau d'érosion et de détérioration de la surface des panneaux modifie leur perception ce qui nuit à la *Détection*. Outre ces points, l'information couleur n'est pas toujours utile pour les autres *ADAS* comme le détecteur de piétons et le détecteur des voies de roulement. C'est l'un des arguments justifiant des approches purement monochromes, comme les approches contours.

2.2.2 Approches contours

Les approches contours sont utilisées pour modéliser la forme géométrique des panneaux dans une scène. Le contour est considéré comme étant la caractéristique de bas niveau la plus utilisée après la couleur dans les systèmes TSR. Plusieurs travaux se basent uniquement sur les contours pour la *Détection* de panneaux [Loy et Barnes, 2004], [Barnes et Zelinsky, 2004], [Barnes *et al.*, 2008], [Belaroussi et Tarel, 2009], et [Belaroussi et Tarel, 2010], et d'autres fusionnent l'approche contours à l'information couleur [Piccioli *et al.*, 1996], [Ruta *et al.*, 2010], et [Timofte *et al.*, 2014].

Extraction de contours

L'extraction de contours représente la première étape dans une approche contours et consiste à détecter les discontinuités d'intensité dans l'image avant leur modélisation. Plusieurs détecteurs de contours ont été utilisés dans les systèmes TSR. Dans [Loy et Barnes, 2004], [Barnes et Zelinsky, 2004], et [Barnes *et al.*, 2008] les contours sont extraits par le filtre de Sobel, ensuite utilisés par la transformée symétrique radiale pour la détection de formes géométriques. Le filtre de Canny a été aussi utilisé dans plusieurs travaux [Piccioli *et al.*, 1996], [Belaroussi et Tarel, 2009], et [Garcia-Garrido *et al.*, 2011]. Pour ne sélectionner que les contours pertinents d'autres informations peuvent être prises en considération. [Garcia-Garrido *et al.*, 2011] sélectionnent uniquement les contours fermés afin de détecter la forme circulaire des panneaux. [Houben, 2011] utilise l'information couleur pour sélectionner les contours des panneaux. Pour cela, il propose une série de pre-traitements afin d'éliminer les contours n'ayant pas les principales couleurs des panneaux recherchés. Parmi cette série de pre-traitements il y a la méthode "Colour Gradient" (CG) et la méthode "Learned Colour Gradient" (LCG). Le CG consiste à pondérer chaque vecteur gradient obtenu à partir de l'image niveau de gris par la disparité couleur de son voisinage dans le plan UV de l'espace YUV (*Y luminance - UV chrominance*). Le LCG calcule une probabilité pour chaque pixel d'être similaire à la couleur principale d'une catégorie de panneaux. La distribution de probabilité est représentée par une mixture de Gaussiennes centrées sur la couleur moyenne d'une catégorie de panneaux et d'une déviation standard prédéfinie. Ensuite, l'information gradient est obtenue à partir de l'image des probabilités (*likelihood*). Un seuillage peut être appliqué sur cette image ce qui donne lieu à la méthode "Learned Colour Threshold" (CT). Un filtrage peut être appliqué sur l'orientation du gradient. La méthode "Expected Orientations" (EO) consiste à ignorer les gradients, obtenus par la méthode LCG, dont l'orientation ne correspond pas aux orientations des contours d'un panneau.

Les approches les plus connues pour la reconnaissance de forme géométrique sont les méthodes de votes. Ces méthodes se basent sur un modèle géométrique, définissant la forme recherchée, dont il faudra déterminer ses paramètres par un processus de *vote*. Le principe est que chaque point contour vote pour un ensemble de paramètres selon certains critères. Au final, les paramètres ayant reçus le plus de votes représentent la forme recherchée. Parmi ces méthodes, la transformée de Hough.

Transformée de Hough

Proposée par Hough en 1962 la transformée de Hough est devenue un standard pour la reconnaissance de formes géométriques. Le principe est de projeter l'image des contours en un espace représentatif de la forme recherchée. Par exemple, pour la détection de droites l'espace (r, θ) est l'espace représentatif où l'équation d'une droite est définie en coordonnées polaires :

$$r = x \cos(\theta) + y \sin(\theta) \quad (2.1)$$

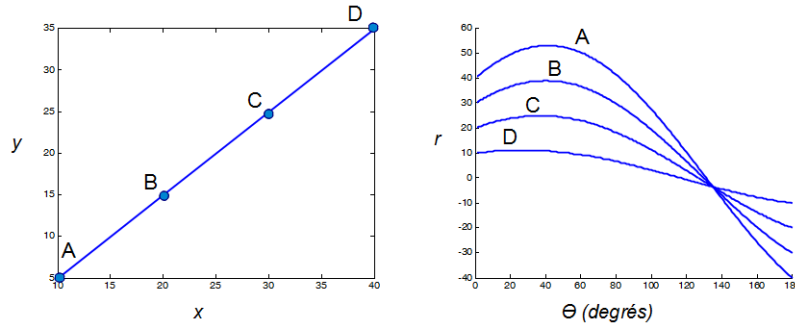


FIGURE 2.4 – Principe de la transformée de Hough.

Si pour un point (x, y) on calcule r en fonction de θ , on obtient une sinusoïde appelée *espace de Hough* (cf. Figure 2.4). Cet espace représente l'ensemble des droites passant par le point (x, y) . Appliquer ce principe à tous les points contours permet d'avoir un ensemble de sinusoïdes où leurs croisements représentent les paramètres (r, θ) des droites présentes dans l'image. L'espace de Hough est un accumulateur 2D associé aux paramètres (r, θ) où les cellules ayant reçu le plus de votes représentent les droites détectées.

La transformée s'étend à la détection des autres formes géométriques comme les cercles. Le principe reste identique, just l'espace qui doit être adapté au modèle. Dans le cas d'un cercle, un espace 3D est utilisé (x_c, y_c, r) où un cercle est représenté par son centre (x_c, y_c) et son rayon r :

$$\begin{cases} x = x_c + r \cos(\theta), \\ y = y_c + r \sin(\theta). \end{cases} \quad (2.2)$$

où l'angle θ varie de 0 à 2π permettant ainsi à chaque point contour (x, y) de tracer un cercle ayant un rayon r . L'intersection des cercles tracés permet de déterminer le centre recherché (x_c, y_c) (cf. Figure 2.5 (a)).

Plusieurs travaux se basent sur la transformée de Hough afin de reconnaître la forme triangulaire et/ou circulaire des panneaux. [Piccioli *et al.*, 1994], [Piccioli *et al.*, 1996] détectent la forme triangulaire en recherchant trois droites qui se croisent mutuellement en formant un angle de 60 degrés. [Garcia-Garrido *et al.*, 2006] proposent d'adapter la détection de panneaux triangulaires en fonction des conditions de luminosité. Cette approche permet de détecter les panneaux même en situations de faible éclairage comme la nuit. Afin d'améliorer les résultats de la détection, [Garcia-Garrido *et al.*, 2006] sélectionnent seulement les contours fermés correspondant à un ratio d'aspect prédéfini. Cette approche a été appliquée aussi pour la détection de panneaux circulaires [Garcia-Garrido *et al.*, 2011]. Évaluée sur une séquence de 30000 images, la méthode détecte 197 des 207 panneaux de la séquence soit 95% de TDPP. [Ruta *et al.*, 2011] utilisent la transformée de Hough pour reconnaître la forme circulaire des ROIs extraits par la méthode *Quad-Tree*. D'autres méthodes de détection ont été proposées pour la reconnaissance de la forme des panneaux.

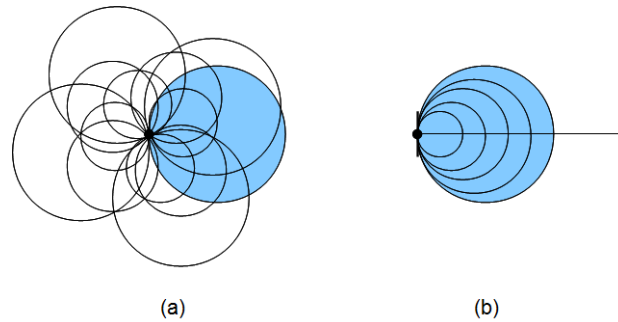


FIGURE 2.5 – Détection de cercles par (a) la transformée de Hough et (b) la Transformée en Symétrie Radiale (TrSR).

Bien qu'elle est simple à implémenter, la transformée de Hough reste lourde en calcul au delà d'un espace à dimension 2 d'où l'apparition de variantes plus rapide et exploitant d'autres informations comme l'orientation des gradients.

Transformée en Symétrie Radiale

Initialement proposée pour la détection des points d'intérêt dans [Loy et Zelinsky, 2003], la Transformée en Symétrie Radiale (TrSR) a été utilisée pour la détection de panneaux circulaires par [Barnes et Zelinsky, 2004], et [Barnes et al., 2008]. La TrSR est une variante de la transformée de Hough pour la détection de cercles. Son avantage est sa faible complexité qui est de " kp " au lieu de " kbp " pour la transformée de Hough, où p représente le nombre de points contours, k la taille de l'ensemble fini des rayons, et b le nombre d'orientations θ considérées. Dans la TrSR, pour chaque point contour (x, y) la valeur de θ est égale à l'orientation du gradient $\vec{g}(x, y)$ (cf. Figure 2.5 (b)). C'est-à-dire pour un rayon donné chaque point contour vote seulement pour un point dans l'espace (x_c, y_c) , diminuant ainsi la complexité d'un facteur b . [Barnes et al., 2008] utilisent un espace de vote pour chaque valeur du rayon.

Détecteur de Polygones Réguliers

[Loy et Barnes, 2004], [Barnes et al., 2005] proposent le Détecteur de Polygones Réguliers (DPR), une adaptation de la TrSR pour la détection de polygones réguliers à n coins comme les triangles, les rectangles, et les octogones. Chaque point de contour (x, y) vote pour l'ensemble des points d'un segment $l(r)$ défini à une distance r et perpendiculaire au gradient $\vec{g}(x, y)$. La longueur w du segment est en fonction de la résolution et du type du polygone recherché, $w = 2r \tan \frac{\pi}{n}$ (cf. Figure 2.6). Lors du vote les incréments sont pondérés par l'amplitude du gradient $\|g(x, y)\|$. Ce processus est réalisé pour un ensemble fini de rayons, ce qui permet d'obtenir pour chaque valeur de r un accumulateur O_r dont le maximum représente le centre du polygone recherché. Ce processus génère un certain nombre de FP vu sa simplicité. Afin d'affiner la détection les auteurs exploitent la principale caractéristique d'un polygone régulier qui est d'avoir des angles égaux. Évalué sur 45 images, pour la détection des panneaux de danger, de priorité, et de stop, le DPR procure un TDPI de 97% avec 25 FP, ce qui représente un FPPI de 0.55.

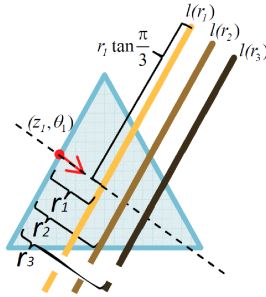


FIGURE 2.6 – Principe du Détecteur de Polygones Réguliers (DPR) ([Houben, 2011]).

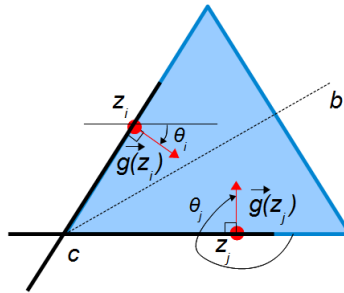


FIGURE 2.7 – Illustration du principe de la Transformée en Sommet et Bissectrice (TSB).

Transformée en Sommet et Bissectrice

[Belaroussi et Tarel, 2009] proposent la Transformée en Sommet et Bissectrice (TSB) pour la détection de panneaux triangulaires. Cette transformée permet la détection des sommets des triangles ainsi que de leur centre. Deux points contours (z_i, z_j) ayant des orientations non-parallèles $\theta_i \neq \theta_j$, votent pour le sommet c formé par l'intersection des deux droites perpendiculaires à leur gradient (cf. Figure 2.7). Pour détecter uniquement les triangles équilatéraux (panneaux triangulaires), une contrainte supplémentaire est ajoutée sur les orientations θ_i et θ_j , $|\theta_i - \theta_j| - \pi \pm \frac{\pi}{3} < \delta$ avec δ une précision angulaire prédéfinie. Ce processus de vote construit l'accumulateur des sommets S_{acc} permettant ainsi la détection des différents sommets d'un triangle (cf. Figure 2.8 (a et b)). Afin de déterminer le centre du triangle, un autre accumulateur B_{acc} est construit où la paire de points contours (z_i, z_j) vote pour le segment $[cb]$ de la bissectrice de l'angle $\angle(z_i c z_j)$ (cf. Figure 2.7). La longueur du segment $[cb]$ dépend de la taille maximale du panneau à détecter dans l'image. La figure 2.8 illustre un exemple des accumulateurs S_{acc} et B_{acc} . La détection des triangles consiste, dans un premier temps, à localiser leur centre à partir des maxima locaux de B_{acc} , ensuite à rechercher autour de chaque centre un triplet de sommets à partir des maxima locaux de S_{acc} . Sur une base de 48 images (base de Grigorescu [Grigorescu et Petkov, 2003]), contenant chacune au moins un panneau, la TSB détecte 38 panneaux triangulaires sur 40 avec seulement 2 FP, soit un TDPI de 95% avec un faible FPPI égal à 0.04.

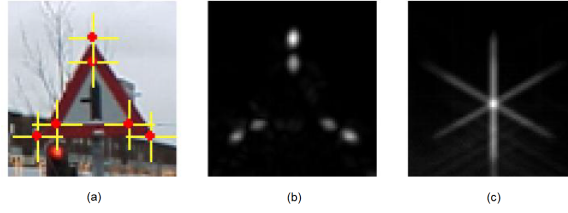


FIGURE 2.8 – Exemple de résultat de la Transformée en Sommet et Bissectrice [Belaroussi et Tarel, 2009]. (a) Image originale, (b) l'accumulateur S_{acc} , et (c) l'accumulateur B_{acc} .

Transformée Chinoise Bilatérale

Pour la détection de panneaux circulaires et rectangulaires, [Belaroussi et Tarel, 2010] proposent la Transformée Chinoise Bilatérale (TCB). Cette transformée se base sur une symétrie centrale et une symétrie axiale (cf. Figure 2.9). Deux points contours (z_i, z_j) ayant des gradients opposés, $\vec{g}(z_i)$ et $\vec{g}(z_j)$, et colinéaires avec le vecteur $\overrightarrow{z_i z_j}$ votent pour le centre $z = \frac{z_i + z_j}{2}$. La figure 2.10 illustre ce principe où les orientations θ_i et θ_j respectivement des gradients $\vec{g}(z_i)$ et $\vec{g}(z_j)$ vérifient :

- la symétrie centrale $|\theta_i - \theta_j| - \pi < \delta$, avec δ une précision angulaire prédéfinie,
- et la symétrie axiale :
 - $|\theta_i - \alpha_{ij}| < \beta$ pour les contours clairs/sombres (cf. Figure 2.10 (a))
 - et $|\theta_i - \alpha_{ij} - \pi| < \beta$ pour les contours sombres/clairs (cf. Figure 2.10 (b))

avec α_{ij} l'angle entre $\overrightarrow{z_i z_j}$ et l'axe horizontal, et β une précision angulaire prédéfinie. Ce processus de vote permet d'obtenir l'accumulateur Acc_{TCB} dont les maxima locaux correspondent aux centres des panneaux circulaires présents dans l'image. Cet accumulateur Acc_{TCB} est défini par :

$$Acc_{TCB}(z) = \sum_{(z_i, z_j) \in \Gamma(z)} D(z_i, z_j) \Theta(z_i, z_j) \log(\|\vec{g}(z_i)\|) \log(\|\vec{g}(z_j)\|) \quad (2.3)$$

avec :

- D une pondération basée sur la distance $\|z_i - z_j\|$, entre z_i et z_j , afin de limiter la résolution des panneaux recherchés ($[R_{max}, R_{min}]$),

$$D(z_i, z_j) = W_{R_{max}}(\|z_i - z_j\|) - W_{R_{min}}(\|z_i - z_j\|) \quad (2.4)$$

où

$$W_R(x) = \begin{cases} 1 & , si |x| < R, \\ 0 & , sinon \end{cases} \quad (2.5)$$

- Θ une transformation de phase qui reflète un indice de symétrie axiale, noté $W_\beta(\theta_i - \alpha_{ij})$ (clair/sombre) et $W_\beta(\theta_i - \alpha_{ij} - \pi)$ (sombre/clair), et un indice de symétrie centrale noté $W_\delta(|\theta_i - \theta_j| - \pi)$.

$$\Theta(z_i, z_j) = (W_\beta(\theta_i - \alpha_{ij}) + W_\beta(\theta_i - \alpha_{ij} - \pi)) \times W_\delta(|\theta_i - \theta_j| - \pi) \quad (2.6)$$

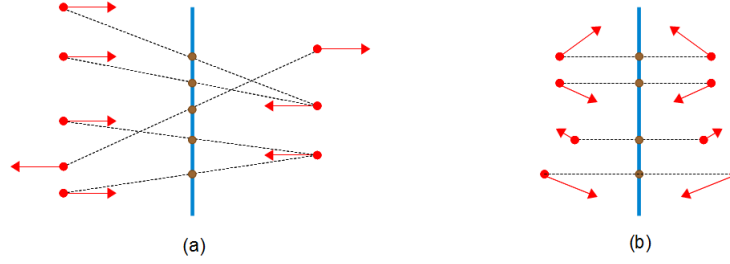


FIGURE 2.9 – La symétrie centrale et la symétrie axiale entre les vecteurs. (a) Symétrie centrale et (b) symétrie axiale.

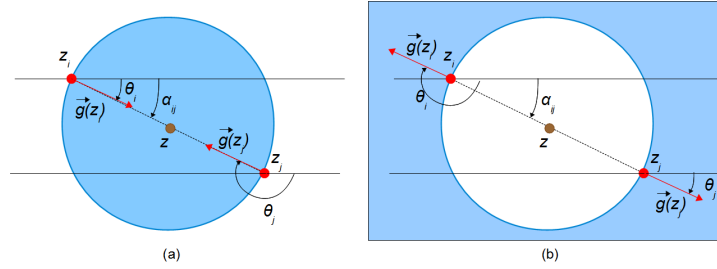


FIGURE 2.10 – Illustration du principe de la Transformée Chinoise Bilatérale. (a) Cas des contours clairs/sombres, (b) cas des contours sombres/clairs.

Pour déterminer la résolution des panneaux, un autre accumulateur *Rayon* est utilisé pour sauvegarder la somme des distances entre les paires de points votant (z_i, z_j) . L'équation 2.7 définit le rayon du panneau détecté avec $N = \text{Card}(\Gamma(z))$ et $\Gamma(z)$ représente l'ensemble des points votant défini par l'équation 2.8.

$$\text{Rayon}(z) = \frac{1}{N} \sum_{(i,j) \in \Gamma(z)} \|z_i - z_j\| \quad (2.7)$$

$$\Gamma(z) = \{(z_i, z_j) \mid \frac{z_i + z_j}{2} = z\} \quad (2.8)$$

Dans [Belaroussi et Tarel, 2010] la TCB est évaluée sur une base de données de 89 images contenant 92 panneaux dont 81 circulaires et 11 rectangulaires. La TCB procure un TDPI de 86% avec 0.28 de FPPI.

Méthodes de vote cible unique

Deux variantes des méthodes TSB et TCB sont présentées dans [Houben, 2011]. Dans un premier temps Houben propose la méthode STVUT (*Single Target Vote for Upright Triangles* - Méthode du Vote Cible Unique pour les Triangles). Contrairement à la TSB, STVUT définit un triangle par son centre et le rayon de son cercle circonscrit. Le processus de vote nécessite la sélection de trois points contours (z_i, z_j, z_k) qui permettent de déterminer les sommets (c_i, c_j, c_k) de la même façon que dans la TSB. Ces sommets contribuent au vote d'un centre z et d'un rayon r définis par :

$$\begin{cases} z &= \frac{1}{3}(c_i + c_j + c_k) \\ r &= \|z - c_i\| \cos(\frac{\pi}{3}) \end{cases} \quad (2.9)$$

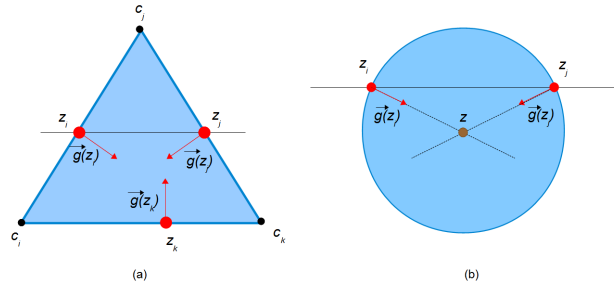


FIGURE 2.11 – Illustration du principe des méthodes de vote de Houben [Houben, 2011]. (a) la méthode STVUT, et (b) la méthode STVUE.

La sélection des points contours votants est basée sur une méthode heuristique. Pour chaque point contour z_i dont l'orientation de gradient θ_i est proche de $-\frac{\pi}{6}$, un point contour z_j dont le gradient vérifie une symétrie axiale avec $\vec{g}(z_i)$ est recherché sur la même ligne que z_i . La recherche du troisième point z_k consiste à sélectionner une colonne entre z_i et z_j sur laquelle le point contour z_k est recherché où son orientation de gradient θ_k doit être proche de $\frac{\pi}{2}$ (cf. Figure 2.11 (a)).

La méthode est étendue à la détection de cercles avec l'algorithme STVUE (*Single Target Vote for Upright Ellipses*) [Houben, 2011]. Cette méthode est proche de la TCB. Ici les deux points votants z_i et z_j doivent être positionnés sur la même ligne dans l'image et ayant des orientations de gradient symétriques. Le centre z du cercle est défini par l'intersection des deux lignes colinéaires aux gradients $\vec{g}(z_i)$ et $\vec{g}(z_j)$ et passant par z_i et z_j (cf. Figure 2.11 (b)).

$$r = \|z_1 - z\| \quad (2.10)$$

Le rayon r du cercle est défini par l'équations 2.10. Sur les 847 images de la base *Stéréopolis* [iTowns, 2009], les méthodes de Houben procurent en moyenne un TDPI de 86%.

2.2.3 Approches descripteurs

La dernière catégorie d'approches répertoriées se base sur l'apparence des panneaux. Ici l'apparence est calculée au moyen des descripteurs locaux ou globaux comme les caractéristiques de Haar, l'Histogramme de Gradients Orientés (HOG) (*Histogram of Oriented Gradients*), les distances aux bords, etc. Ces approches descripteurs demandent une phase d'extraction de caractéristiques et une phase de classification.

2.2.3.1 Descripteurs utilisés pour l'extraction

Caractéristiques de Haar

Initialement proposées par [Viola et Jones, 2001a] pour la détection de visages, ces caractéristiques sont devenues un standard en traitement d'images pour la détection d'objets et notamment pour

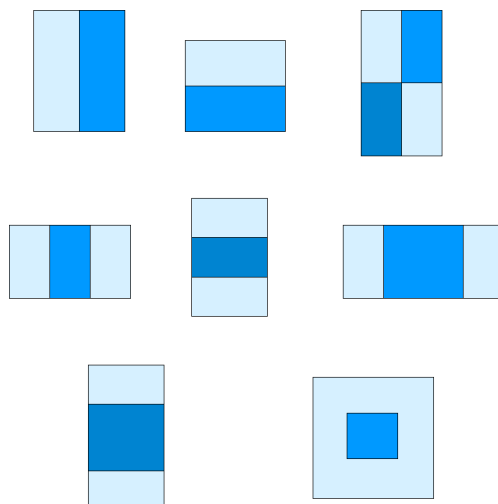


FIGURE 2.12 – Les caractéristiques de Haar.

certaines applications d'aide à la conduite telles que la détection de véhicules [Broggi *et al.*, 2009], la détection de piétons [Sivaraman et Trivedi, 2010], et s'étend bien sûr à la détection de panneaux [Šegvić *et al.*, 2014]. Le principe consiste à calculer autour d'une position donnée dans l'image la différence des sommes de pixels dans les régions adjacentes : *regions claires* et *regions sombres*. Plusieurs exemples de distribution de régions ont été proposées dans [Viola et Jones, 2001a] et étendues par la suite dans [Lienhart et Maydt, 2002] (cf. Figure 2.12). Un exemple est positif si la différence calculée est supérieure à un seuil prédéfini θ , sinon il est négatif.

Le calcul des exemples à différentes positions et échelles dans l'image est optimisé par l'utilisation de "l'image intégrale". La valeur de chaque point (x, y) de l'image intégrale (ii) est la somme des pixels situés au-dessus et à gauche de ce point dans l'image source (i).

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (2.11)$$

La figure 2.13 illustre la représentation de l'image intégrale où la somme des pixels aux points 1, 2, 3 et 4 sont respectivement ceux des régions A , $(A + B)$, $(A + C)$, et $(A + B + C + D)$. Ce principe d'image intégrale permet de calculer la somme des pixels dans la zone D directement par $ii(4) - (ii(2) + ii(3)) + ii(1)$. Cette image permet par simples additions et soustractions de calculer toutes les formes des descripteurs de Haar.

Parmi les variantes des descripteurs de Haar utilisés pour la détection de panneaux il y a les dipôles dissociées [Baró *et al.*, 2009]. Ces dernières sont des régions, *sombres et claires*, qui peuvent être distantes (non contiguës) (cf. Figure 2.14). Par conséquent, le nombre de caractéristiques possibles augmente ce qui peut améliorer la détection des panneaux [Lienhart et Maydt, 2002].

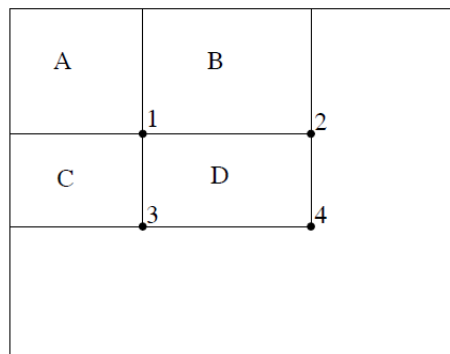


FIGURE 2.13 – Représentation de l’image intégrale [Viola et Jones, 2001a] où la somme des pixels dans la zone D est calculée directement par $ii(4) - (ii(2) + ii(3)) + ii(1)$.

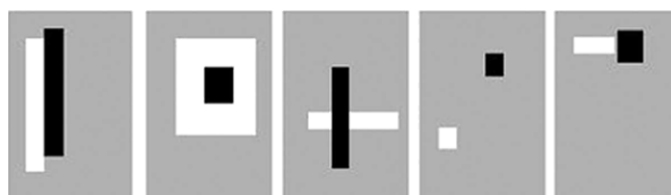


FIGURE 2.14 – Les caractéristiques dipôles dissociées [Balas et Sinha, 2003].

Histogramme de Gradients Orientés

Proposés par [Dalal et Triggs, 2005], les HOGs sont basés principalement sur l’algorithme SIFT (*Scale Invariant Feature Transform*) et connus pour leurs performances dans les systèmes de détection de piétons. Ces histogrammes ont également été utilisés pour la détection de panneaux dans [Creusen *et al.*, 2010], [Overett *et al.*, 2014], et [Zaklouta et Stanciulescu, 2014]. Le principe des HOGs consiste à décrire l’apparence de l’objet par des histogrammes d’orientation. Le calcul de ces histogrammes nécessite en premier la division de l’image ou d’une région en blocs, eux-même subdivisés en cellules. Cette décomposition est illustrée par la figure 2.15. Ensuite, l’amplitude et l’orientation du gradient de chaque pixel sont calculées. Dans chaque cellules, un histogramme d’orientations est formé en se basant sur l’amplitude du gradient pour pondérer les votes. Ensuite, au niveau de chaque bloc les histogrammes de ces cellules sont concaténés pour former l’histogramme du bloc. Afin d’être robuste aux changements de luminosité, ces histogrammes sont normalisés. Au final, la concaténation des histogrammes de tous les blocs, forme le vecteur de “HOG”. La décomposition en blocs et cellules conserve l’information spatiale utile à la reconnaissance d’objets.

Ce principe peut s’étendre aux images couleurs, comme l’ont fait [Creusen *et al.*, 2010] qui exploitent l’information couleur en déterminant les HOG sur l’ensemble des composantes couleurs. La méthode proposée permet de détecter les panneaux circulaires et triangulaires avec un taux moyen de 72% sur un ensemble de 3000 exemples positifs. [Zaklouta et Stanciulescu, 2014] proposent également l’utilisation des HOGs sur l’espace couleur RGB et démontrent une amélioration de la détection par

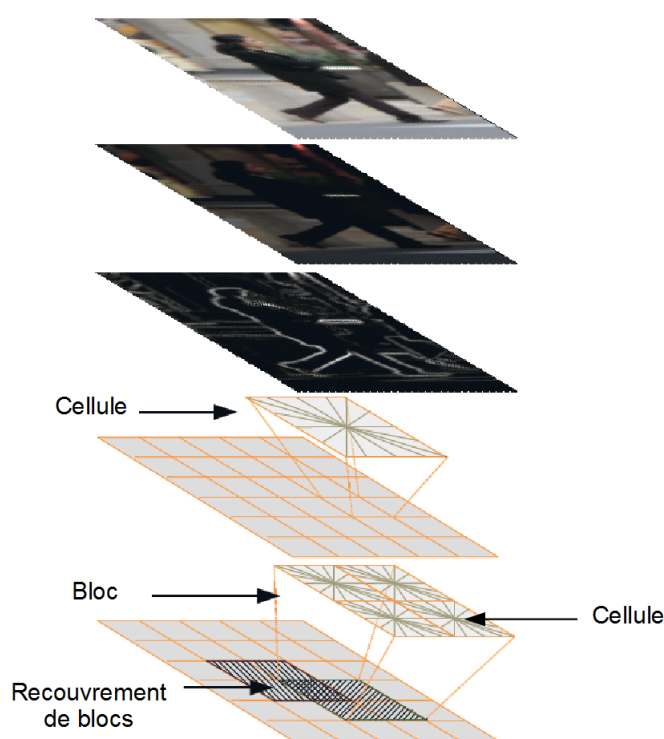


FIGURE 2.15 – Le principe de décomposition pour le calcul des HOG [Dalal, 2005]. L'image est divisée en blocs qui peuvent se recouvrir. Chaque bloc est subdivisé en cellules où un histogramme d'orientations est calculé.

rapport aux HOGs calculés à partir des images niveau de gris. Évalués sur 1584 images, l'utilisation des HOGs couleurs procurent un TDPI de 90% contre 80% uniquement pour les HOGs niveau de gris. Une analyse discriminante linéaire est utilisée par [Overett *et al.*, 2014] pour classifier les HOGs. Sur une base de 3300 exemples positifs, un taux de détection 99% est obtenu. En intégrant l'information couleur, les auteurs diminuent le taux de faux positifs à $10^{-8}\%$ sur leur base de test tout en augmentant le taux de détection mais sans donner de détails sur le FPPI.

Distances aux Bords

[Lafuente-Arroyo *et al.*, 2005] exploitent les distances aux bords (DtB) (*Distances to Borders*) d'une ROI censée englober un panneau (cf. Figure 2.16). Chaque distance est représentée par un vecteur de 20 composantes calculées tout au long de chaque bord. La figure 2.17 montre un exemple des 4 distances D_1 , D_2 , D_3 , et D_4 pour un panneau triangulaire. Ces vecteurs sont normalisés afin d'être invariants au changement d'échelle. La reconnaissance de la forme d'un panneau nécessite l'utilisation de 4 classifieurs SVMs, un pour chaque distance ($D_i, i = 1, \dots, 4$). Une forme est reconnue si au moins 2 des 4 distances sont identifiées comme celles de la forme recherchée. La méthode proposée procure un TDPI de plus de 97% des panneaux sur une base de données constituée de plus de 300 images. L'inconvénient de la méthode est son FPPI qui est supérieur à 2.

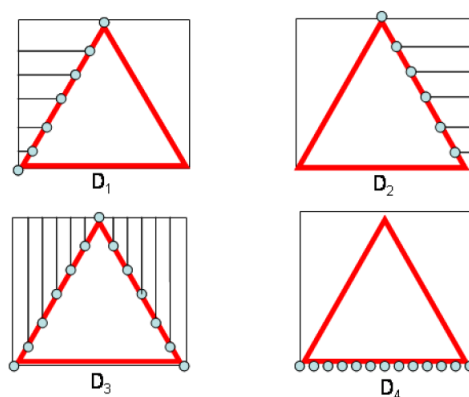


FIGURE 2.16 – Principe des distances aux bords (DtB) représentées par D_1 , D_2 , D_3 , et D_4 ([Lafuente-Arroyo *et al.*, 2005]).

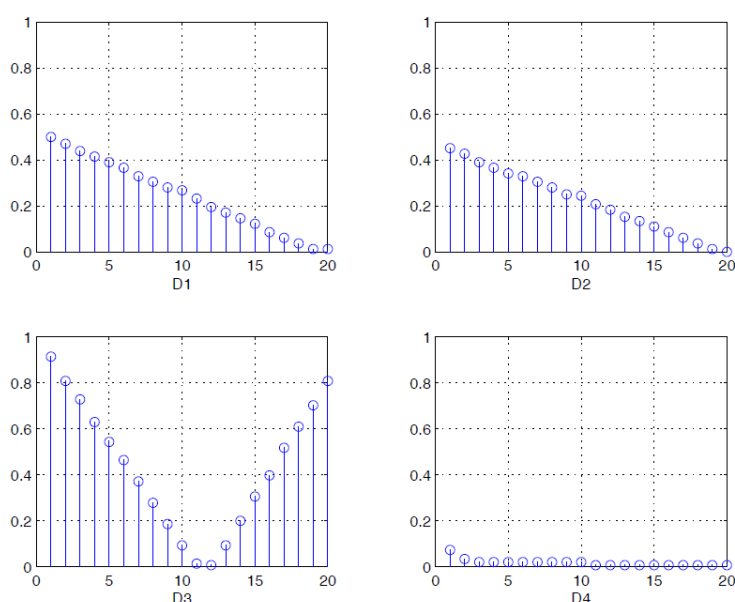


FIGURE 2.17 – Les vecteurs DtB normalisés des distances : D_1 , D_2 , D_3 , et D_4 du panneau triangulaire de la figure 2.16 ([Lafuente-Arroyo *et al.*, 2005]).

[Gil Jiménez *et al.*, 2008] reconnaissent la forme du panneau grâce à la distance entre le centre du blob et son contour selon différents angles (cf. Figure 2.18 (a)). Le vecteur distance obtenu est considéré comme un signal auquel est appliquée la transformée de Fourier rapide (FFT) (*Fast Fourier Transform*). L'intérêt est de reconnaître une périodicité du signal qui caractérise une forme géométrique particulière, comme cela est illustré sur la figure 2.18 (b). Les auteurs présentent uniquement des performances sur des images de synthèses, ce qui ne garantit pas une bonne évaluation de la méthode.

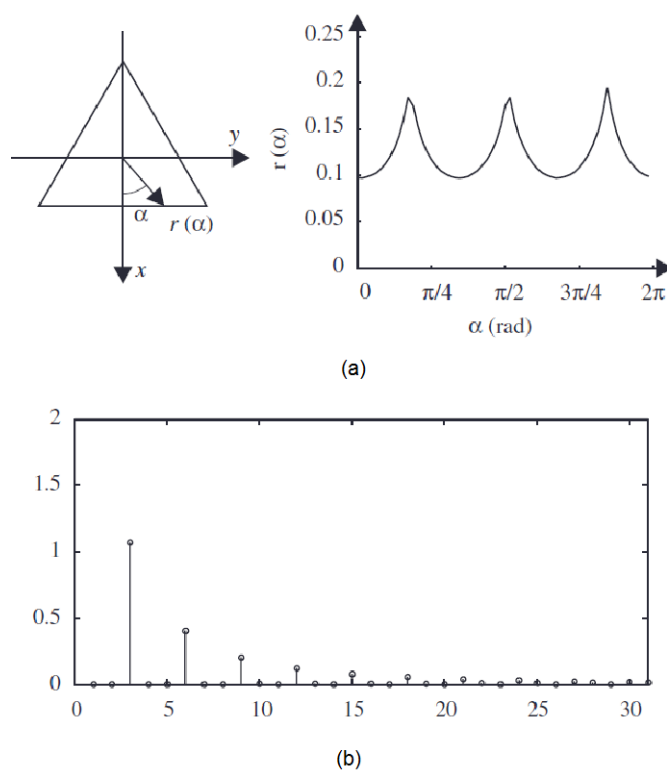


FIGURE 2.18 – Décomposition fréquentielle du bord d’un panneau triangulaire ([Gil Jiménez *et al.*, 2008]). (a) Représentation de la distance aux bords en fonction de l’orientation α . (b) Spectre du signal illustré dans (a).

2.2.3.2 Classification des descripteurs

L’utilisation des descripteurs décrivant “l’apparence” demande une classification en vue d’identifier des objets comme dans une scène routière. Dans le cadre de la détection de panneaux, deux classes sont utilisées pour la classification : la classe “panneau” (*objet*) et la classe “non-panneau” (*non-objet*). Dans la littérature, deux principaux classifieurs sont souvent décrits pour la détection de panneaux : le détecteur en cascade attentionnelle [Deguchi *et al.*, 2011], [Šegvić *et al.*, 2014] et les SVMs [Creusen *et al.*, 2010].

Détecteur en cascade attentionnelle

[Viola et Jones, 2001b] proposent un détecteur en cascade attentionnelle utilisant les descripteurs de Haar pour la détection de visages. Cet algorithme est basé sur une cascade attentionnelle de K classifieurs dits “forts”, performants, de complexité croissante. Les classifieurs “faibles”, ayant un résultat légèrement meilleur que le hasard, sont appliqués en premier pour éliminer la plupart des imageries (*sous-fenêtres*) ne contenant pas l’objet recherché (non-objet), ensuite des classifieurs plus complexes sont cascades afin de rejeter les FP les plus difficiles, suivant le schéma de la figure 2.19. La construction et l’apprentissage des K classifieurs sont basés sur l’algorithme AdaBoost (*Adaptive Boosting*).

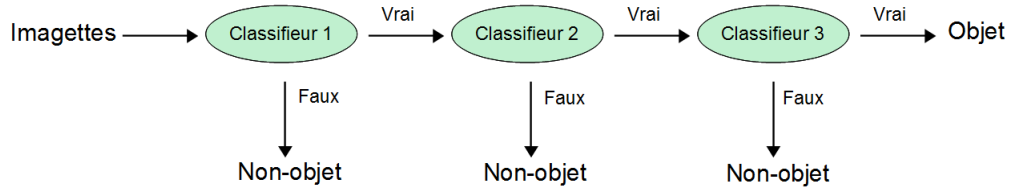


FIGURE 2.19 – Diagramme d’un détecteur en cascade attentionnelle en 3 étapes [Viola et Jones, 2001b].

Dans [Freund et Schapire, 1997] l’algorithme AdaBoost convertit un algorithme faible d’apprentissage en un autre plus efficace en minimisant l’erreur de classification. Adaboost est utilisé dans [Viola et Jones, 2001b] afin de sélectionner un nombre limité de descripteurs de Haar pour décrire l’apparence des visages. La sélection est en fonction de l’erreur commise par les fonctions de classification associées à chaque descripteur.

[Viola et Jones, 2001b] définissent les fonctions de classification h_j par la relation 2.12 avec f_j la valeur de sortie du descripteur de Haar (la différence des sommes de pixels dans les régions claires et sombres), θ_j un seuil et p_j la parité indiquant la direction de l’inégalité ($p_j \in \{-1, +1\}$).

$$h_j = \begin{cases} 1 & \text{si } p_j f_j < p_j \theta_j \\ 0 & \text{sinon} \end{cases} \quad (2.12)$$

La combinaison de ces fonctions permet la construction d’un classifieur fort caractérisé par sa fonction de classification H (cf. 2.13). Cette fonction H est la somme pondérée de T fonctions de classification faible h_t où α_t est le coefficient de pondération et S le seuil qui minimise l’erreur de classification.

$$H = \begin{cases} 1 & \text{si } \sum_{t=1}^T \alpha_t h_t \geq \frac{1}{2} \sum_{t=1}^T \alpha_t = S \\ 0 & \text{sinon} \end{cases} \quad (2.13)$$

L’algorithme AdaBoost est résumé dans l’algorithme 1. A chaque itération t , le classifieur h_t ayant la plus faible erreur ϵ_t est choisi avec une pondération de α_t qui est en fonction de β_t . Le paramètre β_t dépend de l’erreur de classification ϵ_t et il est utilisé lors de la mise à jour des poids $w_{t,i}$ des exemples x_i . Ce processus de mise à jour réduit la probabilité assignée aux exemples bien classés et incrémente la probabilité des exemples mal classés. Par conséquent, ces derniers auront une incidence plus forte lors de la prochaine itération. Au final, le classifieur fort H composé de T classifieurs faibles est obtenu.

Dans l’algorithme de la cascade attentionnelle (cf. Algorithme 2) le nombre de classifieurs faibles pour chaque classifieur fort H_i n’est pas fixé. Deux paramètres de performance sont utilisés afin de déterminer la valeur de T : le taux minimum de vrais positifs VP_{min} et le taux maximum de faux positifs TFP_{max} . Durant l’apprentissage, le seuil S_i du classifieur H_i est décrémenté jusqu’à ce que le classifieur H_i obtient au moins un taux de TVP_{min} de détections correctes sur la base

Algorithme 1 AdaBoost**Donnée :** Ensemble de N exemples x_i étiquetés $l_i \in \{1, 0\}$ **Résultat :** Classifieur fort H Initialisation des poids $w_{1,i} = \frac{1}{N}, i = 1, \dots, N$ **Pour** $t = 1, \dots, T$ **faire**Normaliser les poids $w_{t,i} = \frac{w_{t,i}}{\sum_{i=1}^N w_{t,i}}$ **Pour** chaque descripteur j **faire**Entraîner un classifieur h_j dont l'erreur est définie par $\epsilon_j = \sum_{i=1}^N w_{t,i} |h_j(x_i) - l_i|$ **Fin Pour**Choisir le classifieur h_t avec l'erreur ϵ_t la plus faibleMise à jour des poids $w_{t+1,i} = w_{t,i} \beta_t^{1-e_i}$ où $e_i = 0$ si $h_t(x_i) = l_i, e_i = 1$ sinon, avec $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$ **Fin Pour**Le classifieur fort est $H = \begin{cases} 1 & \text{si } \sum_{t=1}^T \alpha_t g_t \geq \frac{1}{2} \sum_{t=1}^T \alpha_t = S \\ 0 & \text{sinon} \end{cases}$ avec $\alpha_t = \log \frac{1}{\beta_t}$

de validation. De cette manière, le seuil S_i est adapté en fonction de la complexité souhaitée pour l'étage i de la cascade. L'apprentissage de H_i s'arrête quand se dernier atteint le TFP_{max} sur la base des exemples négatifs.

Algorithme 2 Cascade Attentionnelle**Donnée :** TVP_{min}, TFP_{max} ,et l'ensemble des exemples positifs et négatifs respectivement P et N_0 **Résultat :** Détecteur en cascade de K étages $i = 0$ **Tant que** $i < K - 1$ **faire** $f = 1$ $n_i = 0$ $i = i + 1$ **Tant que** $f > TFP_{max}$ **faire** $n_i = n_i + 1$ Entraîner le classifieur H_i où $T = n_i$ avec les exemples de P et N_i Décrémenter le seuil S_i jusqu'à obtenir un taux de TVP_{min} de détections correctes sur P Évaluer H_i sur la base N_i pour déterminer f **Fin Tant que** H_i est rajouté au détecteur en cascadeÉvaluer le détecteur en cascade sur les exemples de N_i et placer les exemples mal classés dans N_{i+1} **Fin Tant que**

[Šegvić *et al.*, 2014] utilisent un détecteur en cascade attentionnelle constitué de 14 étages avec $TVP_{min} = 0.99$ et $TFP_{max} = 0.40$ pour la détection de panneaux triangulaires, permettent de

détecter 96% des panneaux avec une précision de 54% ($precision = \frac{FP}{VP+FP}$). Bien que le détecteur est évalué sur une séquence vidéo le FPPI n'a pas été présenté.

[Deguchi *et al.*, 2011] proposent une méthode d'apprentissage adaptatif basée sur une collecte d'exemples automatique. L'approche consiste à améliorer la détection par un apprentissage continu du détecteur. En premier, un ensemble d'exemples sélectionnés manuellement est utilisé pour l'apprentissage du détecteur en cascade. Ensuite, un nouvel ensemble d'exemples est construit à partir des résultats de la détection. Les exemples positifs sont définis par les panneaux détectés (régions contenant un panneau) et les exemples négatifs sont extraits à partir de l'arrière plan des images. À partir de ce nouvel ensemble d'exemples un nouveau détecteur est construit. Ce processus est répété 4 fois et permet d'obtenir au final un détecteur plus robuste. Lors d'une évaluation sur 3000 images, le détecteur a présenté un TDPI de 90% avec 80% de précision.

Les travaux de [Landesa-Vzquez *et al.*, 2010] se basent sur trois détecteurs en cascade pour la détection de trois types de panneaux : circulaires, triangulaires, et triangulaires inversés. La méthode exploite la forme et l'apparence. Dans un premier temps, les panneaux sont détectés en recherchant une distribution particulière des contours. Ensuite, les détecteurs en cascade confirment la détection de ces panneaux. Les auteurs remplacent les descripteurs de Haar par des descripteurs appelés "Quantums" pour décrire l'apparence des panneaux. Similaires aux descripteurs de Haar, ces descripteurs *Quantums* comparent l'intensité de pixels entre les régions claires et sombres de même surface mais pouvant être distantes. Testée sur trois vidéos expérimentales, la méthode permet de détecter 171 des 176 panneaux présents soit un TDPP de 97,15%. On peut noter que ces auteurs ne donnent aucune information sur le nombre de FP.

Dans [Baró *et al.*, 2009] une variante de AdaBoost est proposée nommée *AdaBoost Évolutif* afin de rendre possible l'apprentissage sur de très grand ensemble de descripteurs. [Baró *et al.*, 2009] utilisent les dipôles dissociés qui sont une version générale des descripteurs de Haar. L'utilisation des dipôles dissociés génère plus de 2^{30} descripteurs différents sur une imagerie de 30×30 pixels au lieu de 600000 descripteurs pour Haar, ce qui rend la tâche d'apprentissage par AdaBoost lourde et complexe [Baró *et al.*, 2009]. Dans l'algorithme *AdaBoost Évolutif*, un algorithme génétique est utilisé pour minimiser l'erreur des fonctions de classification ϵ_t sur l'ensemble des descripteurs et ainsi sélectionner le meilleur classifieur faible h_t . Le détecteur obtenu grâce à *AdaBoost Évolutif* a été utilisé pour la détection des panneaux de danger et d'interdiction. Son évaluation sur une séquence de 9000 images permet un TDPI de 85% et un TDPP de 95% avec un faible FPPI (< 0.05).

Séparateur à Vastes Marges

Un Séparateur à Vastes Marges (SVM) [Burges, 1998] est un classifieur binaire. Considérons N exemples linéairement séparables $x_i \in \mathbb{R}^d$ et étiquetés $l_i \in \{-1, +1\}$. L'équation 2.14 définit l'hyperplan H séparant les deux classes et dont la distance orthogonale avec l'origine vaut $\frac{|b|}{\|w\|}$ et w

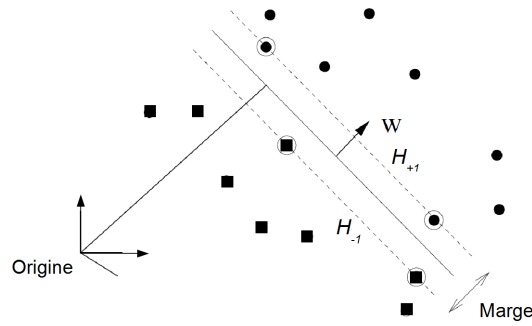


FIGURE 2.20 – Séparation linéaire par les hyperplans H_{+1} , H_{-1} [Burges, 1998]. Les vecteurs supports sont entourés d'un cercle.

le vecteur normal. Les exemples les plus proches de l'hyperplan H appelés vecteurs supports sont situés à une distance d_- et d_+ respectivement pour les classes -1 et $+1$. La marge autour de l'hyperplan H à une largeur de $d_- + d_+$. Le problème est considéré linéairement séparable, ce qui fait que chaque exemple x_i vérifie l'équation 2.15. La figure 2.20 illustre une séparation linéaire des exemples de la base d'apprentissage par un hyperplan H . Les vecteurs supports sont situés sur H_{+1} et H_{-1} avec $w \cdot x_i + b = \pm 1$. Les distances orthogonales de H_{+1} et H_{-1} avec l'origine sont respectivement $\frac{|-1-b|}{\|w\|}$ et $\frac{|1-b|}{\|w\|}$. Donc la marge entre H_{+1} , H_{-1} vaut $\frac{2}{\|w\|}$.

$$w \cdot x + b = 0 \quad (2.14)$$

$$\begin{cases} l_i = -1, & w \cdot x_i + b \leq -1 \\ l_i = 1, & w \cdot x_i + b \geq 1 \end{cases} \Rightarrow l_i(w \cdot x_i + b) - 1 \geq 0 \quad (2.15)$$

Afin d'optimiser les performances de la reconnaissance, les hyperplans H_{+1} et H_{-1} doivent maximiser la largeur de la marge $\frac{2}{\|w\|}$, ce qui est équivalent à minimiser $\frac{1}{2}\|w\|^2$. Cette minimisation est résolue par la programmation quadratique où le problème est défini par une représentation de Lagrange L (cf. (2.16)). Les α_i représentent les multiplicateurs de Lagrange qui remplacent l'inégalité de l'équation 2.15.

$$L = \frac{1}{2}\|w\|^2 - \sum_{i=1}^N \alpha_i (l_i(w \cdot x_i + b) - 1), \alpha_i \geq 0 \quad (2.16)$$

Les SVMs peuvent être généralisés aux données non linéairement séparables. Dans ce cas les données peuvent être transférées par une fonction non linéaire $\Phi(x)$ dans un espace de Hilbert \mathcal{H} de dimension supérieure afin d'être linéairement séparables (cf. Figure 2.21). Parmi les noyaux les plus utilisés on peut citer : le noyau linéaire (cf. (2.17)), le noyau polynômial (cf. (2.18)), et le noyau gaussien (cf. (2.19)).

$$K(x_i, x_j) = x_i \cdot x_j \quad (2.17)$$

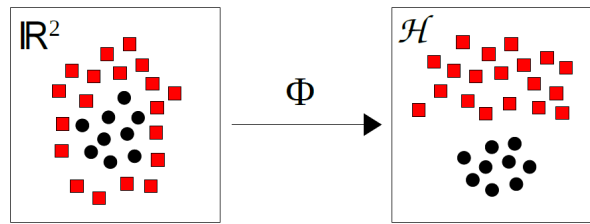


FIGURE 2.21 – La fonction non linéaire transférant les données de l'espace \mathbb{R}^2 vers l'espace de Hilbert \mathcal{H} [Zaklouta, 2011].

$$K(x_i, x_j) = (\gamma x_i \cdot x_j + r)^d, \gamma \in \mathbb{R}^+, r \in \mathbb{R}, d \in \mathbb{N} \quad (2.18)$$

$$K(x_i, x_j) = e^{-\gamma \|x_i \cdot x_j\|} \quad (2.19)$$

Pour déterminer l'étiquette d'un exemple x , le sign de la fonction f suivante est étudiée :

$$\begin{aligned} f(x) &= \text{sign}(w \cdot x + b) \\ &= \text{sign}\left(\sum_{i=1}^m \alpha_i l_i (\Phi(x_{si}) \cdot \Phi(x)) + b\right) \\ &= \text{sign}\left(\sum_{i=1}^m \alpha_i l_i K(x_{si}, x) + b\right) \end{aligned} \quad (2.20)$$

avec x_{si} les vecteurs supports.

[Creusen *et al.*, 2010] se basent sur les HOGs couleurs comme descripteurs de panneaux lors de l'apprentissage itératifs des SVMs. Dans un premier temps, un SVM est construit pour chaque catégorie de panneaux : d'interdiction (circulaires rouges), d'obligation (circulaires bleues), et de danger (triangulaires). Durant cette première phase d'apprentissage un nombre limité d'exemples est utilisé, 170, 74, et 53 exemples positifs pour chacune des catégories citées précédemment et 200 exemples négatifs. Ensuite, les classifieurs obtenus sont testés sur une base de données ne contenant que des exemples négatifs. A l'issue de cette évaluation, les FP sont récupérés et réutilisés lors d'une autre phase d'apprentissage. Ce processus itératif utilise moins de ressources puisque à chaque itération le nombre d'exemples n'est pas très grand. Au final le détecteur obtenu est évalué sur une base de 3000 images. [Creusen *et al.*, 2010] ne présentent ni le TDPI ni le nombre de FP mais seulement l'aire sous la courbe ROC. La détection de panneaux d'obligation et de danger procurent respectivement un aire de 85% et 75% par contre la détection de panneaux d'interdiction ne procure qu'un aire de 56% qui signifie qu'il y a autant de bonnes détections que de FP.

2.2.4 Conclusion

Le présent état de l'art de la *Détection* met en évidence le dynamisme de ce domaine. En effet, les nombreuses récentes publications témoignent de l'intérêt que suscite cette étape auprès des chercheurs (cf. Table 2.2). Cependant, dans les situations réelles la *Détection* n'est jamais parfaite.

Les mauvaises conditions de visibilité ainsi que les problèmes d’occultation rendent la tâche non-triviale et génèrent souvent des FP. L’idéal étant d’avoir un système TSR ayant un fort taux de détection avec un faible nombre de FP. C’est pourquoi, les chercheurs n’hésitent pas à combiner différentes approches (i.e. approches couleur et contours) afin d’améliorer les performances du système au détriment parfois du temps de traitement. Par contre, la méthode de *Détection* proposée dans cette thèse (cf. Chapitre 3) se base uniquement sur la forme des panneaux tout en procurant des résultats similaires aux récentes méthodes. Afin d’améliorer davantage les performances de notre détecteur nous préconisons l’ajout d’une étape de *Suivi temporel* basée sur la fusion de données. C’est pourquoi, la suite de ce chapitre présente le principe de la fusion de données, les étapes du *Suivi temporel*, ainsi qu’un état de l’art du *Suivi temporel* dans les systèmes TSR.

2.3 Fusion de données

La fusion d’informations est apparue dans le domaine militaire pour le traitement de données issues de sources hétérogènes. Depuis, cette approche a été étendue à d’autres domaines tels que la robotique, l’intelligence artificielle, l’imagerie, l’automatique, etc. Trois principaux types d’applications sont distingués [Martin, 2005] :

- la reconnaissance d’objets, détection et identification, à partir de plusieurs sources,
- l’estimation de paramètres par la combinaison de valeurs issues de différentes sources,
- l’association de données permettant d’identifier parmi les différentes mesures perçues celles émises par la même source.

Dans le cadre de travail de cette thèse, nous nous intéressons aux deux derniers types d’applications, en particulier l’association de données.

2.3.1 Processus de fusion

La figure 2.22 illustre les 4 principales étapes d’un processus de fusion : la modélisation, l’estimation, la combinaison, et la décision. La première étape consiste à modéliser l’information. Cette étape est déterminante pour la prise de décision, et nécessite parfois la connaissance externe (par exemple celle d’un expert) et d’éventuelles informations supplémentaires sur les sources (dépendance, fiabilité, etc.). L’étape d’estimation des données (à ne pas confondre avec l’estimation d’un filtre de Kalman par exemple) est dépendante de la modélisation, et consiste à donner une valeur numérique aux données modélisées. La troisième étape, la combinaison, fusionne les informations par un opérateur adapté au formalisme de modélisation choisi. Cette étape peut aussi nécessiter la connaissance de l’expert et des informations supplémentaires sur les sources. Les résultats de la combinaison sont de même nature que les informations de départ, ce qui garantit la même interprétation. La dernière étape du processus est la prise de décision selon un critère

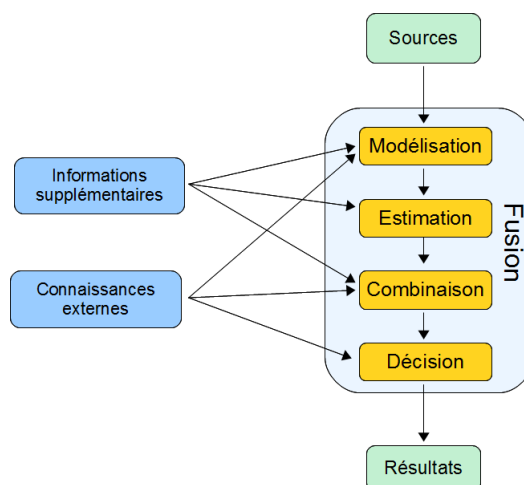


FIGURE 2.22 – Diagramme du processus de fusion [Martin, 2005].

prédéfini. Le choix de ce dernier dépend de la modélisation et de l'opérateur de combinaison utilisé en plus des connaissances externes.

2.3.2 Imperfections de l'information

Dans la pratique, les informations à fusionner ne sont jamais parfaites. Dans le contexte de la fusion d'informations, on cherchera au mieux à modéliser ces imperfections. L'imprécision, l'incertitude, et l'incomplétude sont les principaux défauts de l'information [Martin, 2005] :

- **Imprécision** : elle est caractérisée par le contenu de l'information. C'est un défaut quantitatif mesuré par l'erreur entre l'information issue des sources et la réalité de l'environnement qui doit être connue ou estimée.
- **Incertain** : elle représente le niveau de conformité d'une information avec la réalité. Une information incertaine fournit une connaissance partielle de la réalité, alors qu'une information certaine permet la connaissance de toute la réalité. L'incertitude est un défaut qualitatif qui est souvent par erreur, confondu avec l'imprécision, de nature quantitative.
- **Incomplète** : est le manque d'information apporté par une source. Cette absence d'information caractérise la plupart des systèmes de perception où la vision de l'environnement n'est que partielle.

D'autres imperfections sont rapportées à la fiabilité des sources comme le conflit et l'ambiguïté.

- **Conflit** : deux ou plusieurs informations contradictoires génèrent un conflit. Le conflit est dû à la fiabilité variable des sources ou à la complexité de l'environnement perçu.
 - **Ambiguïté** : l'ambiguïté d'une information est le fait que celle-ci ait plusieurs interprétations.
- Tenir compte de toutes ces imperfections lors du processus du *Suivi temporel* n'est pas du tout évident et la plupart des stratégies de suivi ne permettent pas leur modélisation. Il existe trois grands formalismes pour représenter l'information : théorie des probabilités, théorie des possibilités, et théorie des croyances. La théorie des probabilités ou l'approche bayésienne, telle que la méthode

du plus proche voisin global (GNN-*Global Nearest Neighbor*) et les méthodes probabilistes (PDA-*Probabilistic Data Association* et JPDA-*Joint PDA*) présentées dans la section 2.4.3, se base sur un cadre mathématique rigoureux où l'information est modélisée par des distributions de probabilités. Cette approche ne permet de représenter que l'incertitude de l'information. En effet, l'approche bayésienne n'offre pas les outils pour modéliser l'imprécision et ne gère pas le conflit et l'ambiguïté des informations. La théorie des possibilités se base sur une représentation floue de l'information ce qui lui procure une certaine flexibilité permettant ainsi de mieux modéliser les imperfections de l'information. Quant à la théorie des croyances, celle-ci représente l'information à partir des fonctions de croyance. Ces dernières prennent en comptes l'incertitude et l'imprécision lors de la modélisation. De plus, la théorie des croyances permet la gestion du conflit et de l'ambiguïté des sources. C'est pourquoi, nous l'avons choisi pour notre méthode de *Suivi temporel* présentée dans le chapitre 4.

2.4 Stratégies pour le suivi temporel

Cette section présente, dans un premier temps, les deux étapes du *Suivi temporel* : le filtrage et l'association de données. Ensuite, les travaux récents relatifs aux dispositifs TSR sont synthétisés et décrits selon deux catégories : approches suivi-*après*-reconnaissance et approches suivi-*après*-détection. La troisième partie de la section présente un état de l'art des principales méthodes d'association de données.

2.4.1 Étapes du suivi temporel

Le *Suivi temporel* d'un objet consiste à estimer ses paramètres dynamiques et/ou cinématiques, i.e. sa position et sa vitesse, à partir des mesures effectuées. Souvent cette tâche est réalisée grâce aux techniques de filtrage statistique comme le filtrage de Kalman. Toutefois, le suivi devient difficile dans un contexte dynamique, bruité et "multi-objets" comme pour les ADAS où le nombre d'objets à suivre (de panneaux pour un TSR) évolue à travers le temps du fait de leur apparition et disparition. L'objectif n'est plus de déterminer à partir des observations, fournies par les capteurs, les mesures issues d'un objet unique, mais d'associer à chaque instant les mesures liées à chaque objet présent dans la scène, ce qui permet d'estimer leurs paramètres dynamiques. Le diagramme d'un processus de suivi multi-objets est présenté par la figure 2.23. En premier, le *Filtrage* assure la *Prédiction* des états des objets connus (identifiés à l'instant $k - 1$), communément appelés "*pistes*" de l'instant $k - 1$ à l'instant k . Ensuite, l'*Association de données* met en correspondance les *pistes* avec les mesures perçues à l'instant k , communément appelées "*cibles*". Ce processus d'association est souvent précédé par un *Fenêtrage temporel* afin de réduire le nombre d'associations possibles. Par la suite, la *Maintenance* des pistes est effectuée en fonction des "*Meilleures*" associations

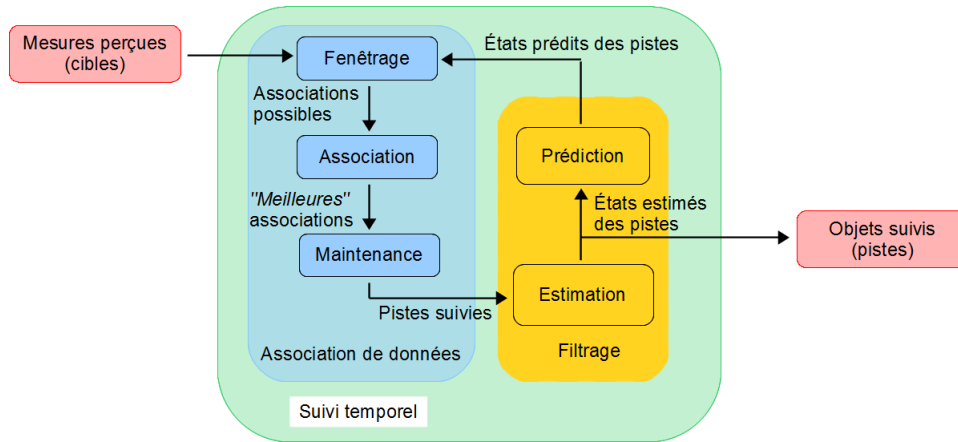


FIGURE 2.23 – Diagramme de construction d'un système de suivi multi-objets.

retenues. Cette étape permet l'initialisation et la suppression de pistes. Au final, en fonction des résultats de l'association l'*Estimation* des états des pistes est effectuée.

2.4.1.1 Filtrage

Le Filtre de Kalman (FK) a été appliqué avec succès dans de nombreuses applications de suivi d'objets dans des séquences vidéo [Weng *et al.*, 2006]. Le filtre estime récursivement l'état de l'objet (piste) en se basant sur une combinaison linéaire des estimations précédentes et des nouvelles mesures observées (cibles). Ce processus récursif se compose de deux étapes :

- l'étape de *Prédiction* qui exploite un modèle dynamique afin de prédire l'état de l'objet. La covariance d'erreur de prédiction associée est aussi calculée durant cette étape.
- l'étape d'*Estimation* où le filtre corrige l'état de l'objet et la covariance prédits en fonction des mesures associées.

Cette section introduit les principes du FK et les notions qui seront utilisés par la suite dans notre application TSR.

Soit le système à temps discret, décrit par le modèle d'état linéaire suivant :

$$\begin{cases} x_j(k) = Fx_j(k-1) + w(k) \\ z_i(k) = Hx_j(k) + v(k) \end{cases} \quad (2.21)$$

avec :

- $x_j(k) \in \mathbb{R}^l$ le vecteur d'état de la piste j à l'instant k ,
- $F \in \mathbb{R}^{l \times l}$ la matrice dynamique du système,
- $z_i(k) \in \mathbb{R}^p$ la mesure i à l'instant k
- $H \in \mathbb{R}^{p \times l}$ la matrice d'observation du système,
- $w(k)$ le bruit de processus à l'instant k , bruit Gaussien blanc de moyenne nulle et de matrice de covariance $Q \in \mathbb{R}^{l \times l}$

– $v(k)$ le bruit de mesure à l’instant k , bruit Gaussien blanc de moyenne nulle et de matrice de covariance $R \in \mathbb{R}^{p \times p}$.

Les bruits sont supposés indépendants entre eux, $E[w(k)v(k)^T] = 0$, et indépendants de l’état.

L’objectif est de calculer un estimateur (optimal) à l’instant k sachant les informations du modèle $(F, H, x_j(m), Q, R)$ à l’instant m de $x_j(k)$ noté $\hat{x}_j(k|m)$. Trois types de problèmes sont envisageables : prédiction ($k > m$), filtrage ($k = m$), et lissage ($k < m$).

À partir de l’estimée $\hat{x}_j(k-1|k-1)$, il est possible de prédire deux informations :

– x_j à l’instant k :

$$\hat{x}_j(k|k-1) = F\hat{x}_j(k-1|k-1) \quad (2.22)$$

– et sa mesure correspondante à l’instant k :

$$\hat{z}_j(k|k-1) = H\hat{x}_j(k|k-1) \quad (2.23)$$

Le filtre se base sur une correction linéaire afin d’améliorer l’estimation $\hat{x}_j(k|k-1)$ en fonction de la mesure perçue $z_i(k)$:

$$\hat{x}_j(k|k) = \hat{x}_j(k|k-1) + K_j(k)(z_i(k) - \hat{z}_j(k|k-1)) \quad (2.24)$$

La matrice $K_j(k)$ représente le gain du filtre qui est fonction de la matrice de covariance d’erreur $P_j(k|k-1)$:

$$P_j(k|k-1) = \mathbb{E}[(x_j(k) - \hat{x}_j(k|k-1))(x_j(k) - \hat{x}_j(k|k-1))^t] \quad (2.25)$$

L’algorithme 3 résume le FK avec sa phase d’initialisation et son processus récursif composé des deux étapes : *Prédiction* et *Estimation*. En premier, le filtre est initialisé à l’aide des premières mesures. Cette initialisation concerne l’état initial du système $\hat{x}_j(0|0)$, la matrice de covariance d’erreur de prédiction $P_j(0|0)$, et les matrices de covariance des bruits de processus et mesure respectivement Q et R . Ensuite, le processus récursif prédit l’état $\hat{x}_j(k|k-1)$ du système à partir des estimations précédentes $\hat{x}_j(k-1|k-1)$. À la réception de nouvelles mesures $z_i(k)$, l’état prédit est corrigé. La correction est fonction du gain $K_j(k)$ qui minimise la covariance d’erreur. Il est utile de préciser que dans un contexte multi-objets, une étape de mise en correspondance entre les mesures perçues et les objets est nécessaire pour la correction des filtres. Dans le chapitre 4, nous détaillerons la mise en oeuvre de ce filtre pour le *Suivi temporel* de panneaux routiers.

2.4.1.2 Association de données

L’association de données est composée de trois étapes : *Fenêtrage temporel*, *Association*, et *Maintenance* (cf. Figure 2.23).

Algorithme 3 Le filtre de Kalman**Donnée :** Mesure z_i **Résultat :** État du système \hat{x}_j 1. **Initialisation** $\hat{x}_j(0|0), P_j(0|0), Q$, et R 2. **Processus récursif****Pour** $k = 1 \dots \infty$ **faire***Prédiction*

– prédiction de l'état :

$$\hat{x}_j(k|k-1) = F\hat{x}_j(k-1|k-1) \quad (2.26)$$

– covariance de l'erreur de prédiction :

$$P_j(k|k-1) = FP_j(k-1|k-1)F^t + Q \quad (2.27)$$

Estimation

– gain du filtre :

$$K_j(k) = P_j(k|k-1)H^t[HP_j(k|k-1)H^t + R]^{-1} \quad (2.28)$$

– correction de l'état en fonction de la mesure associée $z_i(k)$:

$$\hat{x}_j(k|k) = \hat{x}_j(k|k-1) + K_j(k)[z_i(k) - H\hat{x}_j(k|k-1)] \quad (2.29)$$

– covariance de l'erreur d'estimation :

$$P_j(k|k) = [I - K_j(k)H]P_j(k|k-1) \quad (2.30)$$

Fin Pour*Fenêtrage temporel*

La première étape, le *Fenêtrage* ou *gating*, élimine les associations improbables entre cibles et pistes afin de réduire la complexité lors de l'étape d'*Association*. Ce processus s'opère par le calcul d'une distance statistique entre les cibles $z_i(k)$ ($i = 1, \dots, n$) et la piste considérée $\hat{x}_j(k|k-1)$. Les cibles générant une distance supérieure à un seuil prédéfini sont considérées comme improbables et sont donc supprimées. Les autres cibles, se trouvant dans la fenêtre d'association, sont considérées comme *possibles* et sont retenues pour la prochaine étape d'*Association*. Une des distances statistiques utilisée pour le calcul de la fenêtre d'association est la distance de Mahalanobis [Blackman et Popoli, 1999]. Cette distance notée $d_{i,j}^2$ est définie par :

$$d_{i,j}^2 = \tilde{z}_{ij}(k|k-1)^T S_j^{-1}(k) \tilde{z}_{ij}(k|k-1) \quad (2.31)$$

avec :

– i indice de la cible perçue $z_i(k)$ et j indice de la piste prédite dont la mesure est $\hat{z}_j(k|k-1)$

- S_j matrice d'innovation du FK, correspondant à la piste j , donnée par :

$$S_j(k) = HP_j(k|k-1)H^t + R \quad (2.32)$$

- \tilde{z}_{ij} innovation entre l'observation i ($z_i(k)$) et la prédiction j ($\hat{z}_j(k|k-1)$) définie par :

$$\tilde{z}_{ij}(k|k-1) = (z_i(k) - \hat{z}_j(k|k-1)) \quad (2.33)$$

Cette métrique prend en considération les erreurs sur les pistes prédites $P_j(k|k-1)$ lors de l'étape de prédiction du filtre j et les erreurs des observations (cibles) R . La fenêtre d'association peut être définie par $d_{i,j}^2 < G$ avec G le seuil de la fenêtre.

Association

L'association de données est une étape très importante pour le suivi d'objets car elle identifie les différents objets de la scène et leur associe les mesures correspondantes. De nombreuses approches, principalement bayésiennes existent dans la littérature (cf. Section 2.4.3.2). L'imperfection des mesures/données venant des capteurs et/ou des algorithmes de perception ne facilitent pas le traitement. La fusion d'informations peut s'avérer utile pour ce type de contexte où les données peuvent être multi-sources, incertaines, incomplètes, et imprécises. Le processus de fusion aboutit à une prise de décision sur les associations *piste-cible*, ce qui permet par la suite de mettre à jour l'estimation des états des objets à l'aide de la technique de filtrage retenue.

Maintenance

La dernière étape consiste à analyser les associations *piste-cible* retenues ("*Meilleures associations*") afin de suivre l'évolution de la scène qui se résume dans l'apparition et la disparition des pistes. En fonction des résultats de l'association des pistes peuvent être créés, ce qui se traduit par l'initialisation de nouveaux filtres à partir des mesures non-associées. La disparition d'une piste conduit à une suppression du filtre correspondant.

2.4.2 Suivi temporel : état de l'art

Dans un TSR, le *Suivi temporel* peut être intégré à deux niveaux différents : après la reconnaissance et/ou après la détection. Dans cette section, nous présentons les récentes et rares publications où une phase de *Suivi temporel* est intégrée afin d'améliorer les performances du système.

2.4.2.1 Approches suivi-après-reconnaissance

Plusieurs systèmes TSR insèrent la phase de *Suivi temporel* après la reconnaissance [Moutarde *et al.*, 2007], [Keller *et al.*, 2008], [Lafuente-Arroyo *et al.*, 2008], et [Meuter *et al.*, 2011]. [Garcia-Garrido *et al.*, 2006] considèrent qu'un panneau reconnu doit être suivi afin de ne plus le reconnaître



FIGURE 2.24 – Exemple du suivi temporel de panneaux routiers [Garcia-Garrido *et al.*, 2006].

à nouveau dans les images suivantes, ce qui permet de réduire le temps de traitement. Ce suivi est majoritairement basé sur le FK où un panneau est défini par sa position (x, y) dans l'image ainsi que par son échelle s . La figure 2.24 présente un exemple de résultats du *Suivi temporel* dans diverses conditions de luminosité illustrant la robustesse du suivi. Malheureusement, les auteurs ne quantifient pas l'apport de leur approche par rapport à un TSR classique.

Dans [Moutarde *et al.*, 2007], le *Suivi temporel* permet de valider la reconnaissance des panneaux. Lors de l'étape de reconnaissance, une valeur de confiance est attribuée à chaque panneau identifié. Cette valeur sera incrémentée si le panneau est identifié dans les images suivantes. Dans le cas contraire, cette valeur est décrémentée. La reconnaissance d'un panneau est confirmée si sa valeur de confiance est supérieure à un seuil prédéfini. Ce principe a été évalué sur une base de données privée de plus de 150 *min* de vidéo contenant 281 panneaux de limitation de vitesse. Les résultats présentent un taux d'erreur de classification inférieur à 1% mais sans évaluer l'impact direct du suivi sur le système. Dans ces travaux le suivi est assuré par un FK.

Dans un environnement complexe tel qu'une scène routière, un panneau peut être mal classifié ou non reconnu temporairement, sur une ou plusieurs images du fait d'occultations temporaires par exemple. Face à ce problème, [Keller *et al.*, 2008] proposent de fusionner les décisions de classification d'un même panneau par l'intégration de l'information temporelle. A l'aide d'un processus de vote, les décisions de classification du panneau suivi, obtenues depuis son apparition, sont fusionnées pour obtenir une décision unique. Cette approche permet d'éliminer ainsi les erreurs de classification accidentelle dues par exemple à une mauvaise visibilité. Ce principe est étendu dans [Meuter *et al.*, 2011] où la prise de décision est basée sur une approche Bayésienne. [Meuter *et al.*, 2011] proposent aussi de fusionner les panneaux identiques qui peuvent apparaître sur chaque côté de la route, ce qui permet de réduire encore les erreurs de classifications (cf. Figure 2.25). Les auteurs ont montré que cette approche de fusion permet une réduction de 83% des erreurs de classification.



FIGURE 2.25 – Exemple de fusion de panneaux identiques par l'approche Bayésienne [Meuter *et al.*, 2011].

[Lafuente-Arroyo *et al.*, 2008] proposent un système TSR pour les applications d'inventaire. Le système procure une liste de panneaux reconnus sur une section de réseau routier, ce qui permet la maintenance de la signalisation. Pour maximiser les performances du système, les auteurs appliquent plusieurs algorithmes de détection basés sur différentes techniques de segmentation couleur. Dans ce contexte, un même panneau peut être reconnu plusieurs fois dans la même image puisqu'il sera détecté par plus d'un détecteur. Face à ce problème de multiples détections, les auteurs ajoutent une étape de suivi après la reconnaissance. L'objectif est d'associer les panneaux reconnus précédemment avec les panneaux détectés dans l'image courante, ce qui permet de déterminer les nouveaux panneaux et de supprimer ceux qui sont déjà reconnus. L'association est basée sur la couleur des panneaux, leur forme géométrique, et leur position. [Lafuente-Arroyo *et al.*, 2008] n'apporte aucun détail sur l'algorithme d'association utilisé et son intérêt n'est pas quantifié objectivement.

Parallèlement à l'amélioration de l'étape de reconnaissance, le *Suivi temporel* peut également influencer la détection principalement pour la réduction du nombre de FP. Pour cela, il faut que l'étape de *Suivi temporel* soit intégrée directement après la détection.

2.4.2.2 Approches suivi-*après*-détection

Dans le contexte du suivi-*après*-détection, l'information temporelle permet d'améliorer le système TSR en réduisant le temps de traitement (de la détection et/ou de la reconnaissance) et en diminuant le nombre de FP, ce qui procure une stabilité accrue au système. [Fang *et al.*, 2003] proposent d'effectuer la reconnaissance d'un panneau uniquement si sa résolution est élevée. Les panneaux détectés sont suivis à travers les images jusqu'à ce qu'ils aient une certaine résolution, ce qui facilite leur reconnaissance et minimise les erreurs de classification. En se basant sur la vitesse du véhicule et la résolution réelle des panneaux, un FK prédit la position et la résolution de chaque panneau dans les images. Ces prédictions permettent de réduire la zone de recherche des panneaux et donc le temps de traitement de la détection.

Toujours dans le but de réduire le temps de calcul, [Ruta *et al.*, 2010] procèdent à une détection de panneaux sur l'intégralité de l'image de façon périodique uniquement. Ce processus détecte l'apparition des nouveaux panneaux dans la scène qui sont par la suite suivis à l'aide des FKs. Les

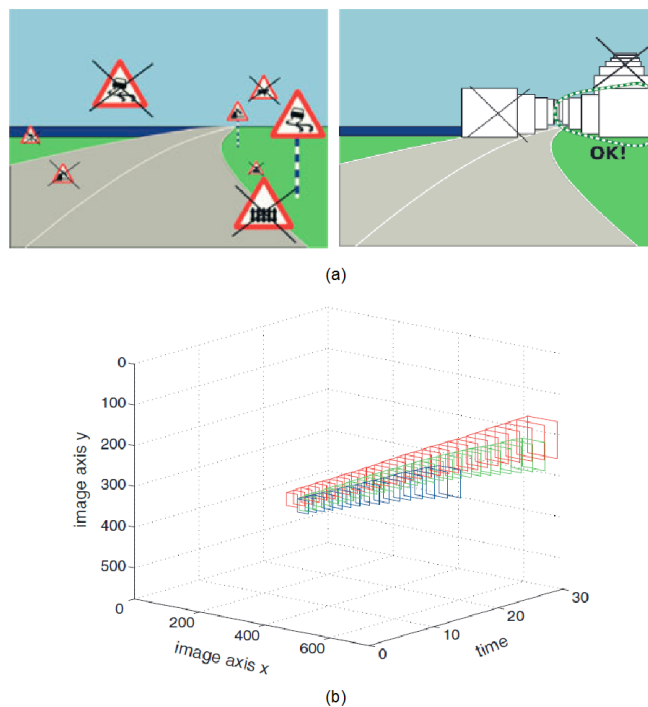


FIGURE 2.26 – Exploitation de l'information spatio-temporelle pour l'élimination de faux positifs [Šegvić *et al.*, 2014]. a) Le contexte spatio-temporel et les faux positifs. b) Exemple de trois trajectoires de panneaux dans un espace 3D (image 2D + temps).

panneaux sont représentés par deux vecteurs. Le premier vecteur décrit la position et la vitesse de déplacement des panneaux dans l'image $[x, y, v_x, v_y]^t$. Le deuxième vecteur symbolise la résolution des panneaux à travers leur largeur et hauteur $[w, h]^t$. Le processus de suivi détermine un *a priori* sur les positions et les résolutions des panneaux où le détecteur sera localement appliqué afin de confirmer ou non la présence du panneau. Dans [Ruta *et al.*, 2010], seule une évaluation globale du système est présentée, ce qui ne permet pas la comparaison de l'approche de suivi adoptée avec d'autres méthodes.

Dans [Nunn *et al.*, 2008] et [Meuter *et al.*, 2008], l'information temporelle est utilisée pour réduire le nombre de FP. [Nunn *et al.*, 2008] utilisent un FK pour le suivi afin de détecter et supprimer les panneaux candidats dont le mouvement n'est pas linéaire ou dont l'apparition n'est pas stable à travers la séquence d'images. [Meuter *et al.*, 2008] se basent sur un filtre particulaire combiné à un modèle de caméra pour projeter les résultats dans le plan réel. Lors du processus de suivi, chaque panneau détecté est défini par sa position dans l'espace 3D et par sa résolution réelle. L'élimination des FP est basée sur la durée de vie des panneaux candidats qui est représentée par le nombre d'images où le panneau candidat a été détecté. Un seuil minimum d'apparition est nécessaire afin de valider la détection d'un panneau, qui dans le cas contraire est considéré comme un FP. Avec un seuil de 5 apparitions, [Meuter *et al.*, 2008] réduisent de 50% le nombre de FP. Les vrais panneaux n'ayant pas encore atteint ce seuil d'apparition sont considérées comme des FP, ce qui dégrade légèrement le taux de détection (TDPI).

En plus de l'information temporelle, [Šegvić *et al.*, 2014] exploitent l'information spatiale qui caractérise le mouvement des panneaux dans la séquence d'images pour réduire le nombre de FP (cf. Figure 2.26 (a)). Pour cela, les trajectoires des panneaux suivies sont classifiées afin d'éliminer celles liées aux FP. Contrairement à ces derniers, les panneaux ne se déplacent pas de manière aléatoire et leur résolution augmente au fil du temps. Ce contexte spatio-temporel est illustré dans la figure 2.26 (b). Afin d'identifier les trajectoires des panneaux du reste, les auteurs se basent sur un classifieur. Évaluée sur une base constituée de 268 trajectoires positives et 601 trajectoires négatives, l'approche proposée permet une réduction de 82% de FP tout en sauvegardant 98% des panneaux détectés. Ces performances nécessitent a contrario un alignement parfait des panneaux détectés. Dans la pratique, cela n'est pas toujours possible à cause des mouvements et vibrations de la caméra embarquée.

Gu *et al.* [Gu *et al.*, 2011] proposent un système basé sur une caméra multi-focale. Le dispositif est composé d'une caméra CCD standard de 640×480 pixels et d'un téléphoto zoom de 320×240 pixels. L'idée est de détecter les panneaux à partir des images de la caméra CCD, ensuite d'augmenter leur résolution en se focalisant dessus afin de faciliter leur reconnaissance. Le temps nécessaire à ce procédé n'est pas négligeable. Par conséquent, les auteurs proposent de se focaliser une seule fois sur chaque panneau, ce qui nécessite leur suivi. Si un panneau est détecté sur trois images successives alors le dispositif se focalisera une seule fois dessus et l'image obtenue sera transmise à l'étape de reconnaissance. L'évaluation présentée dans [Gu *et al.*, 2011] ne quantifie pas l'apport du suivi.

2.4.2.3 Conclusion

L'état de l'art présenté ci-dessus montre clairement l'intérêt du *Suivi temporel* pour l'amélioration des systèmes TSR. Néanmoins, son intégration reste limitée comme cela a été souligné par [Møgelmoose *et al.*, 2012]. En effet, sur les 41 travaux discutés dans [Møgelmoose *et al.*, 2012] seulement 14 d'entre eux utilisent le *Suivi temporel*. De plus, pour la plupart, le *Suivi temporel* est intégré uniquement pour réduire le temps de traitement alors qu'une meilleure exploitation de l'information temporelle peut améliorer davantage les étapes de détection et/ou de reconnaissance. Nous constatons également que contrairement aux étapes de détection et de reconnaissance, l'impact du *Suivi temporel* sur le système TSR n'est que rarement quantifié, négligeant ainsi son importance. Dans ce contexte, nous considérons que le *Suivi temporel* mérite davantage d'intérêt, particulièrement pour l'amélioration de la détection. C'est pourquoi, nous proposons, dans le chapitre 4, une nouvelle approche de *Suivi temporel* pour remédier aux imperfections de la détection comme les FP. L'approche proposée met davantage l'accent sur l'association que sur le filtrage. Par conséquent, on détaillera davantage les techniques usuelles d'association dans la suite de ce chapitre.

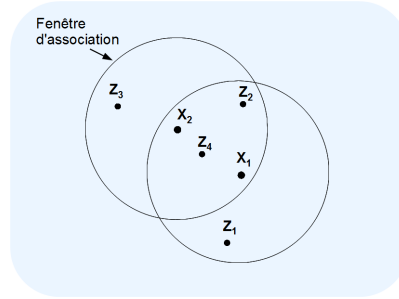


FIGURE 2.27 – Exemple d’association multiple. la cible z_4 est associée à deux pistes x_1 et x_2 .

2.4.3 Méthodes d’association de données

Plusieurs ouvrages discutent les méthodes d’associations classiques : méthode du plus proche voisin, méthodes probabilistes, et le pistage à hypothèses multiples [Blackman, 1986], [Blackman et Popoli, 1999]. Ces méthodes ont été développées initialement pour les applications militaires et depuis, elles sont étendues à d’autres domaines comme la robotique et les véhicules intelligents où l’association d’objets perçus permet une meilleure perception de l’environnement. Cette section décrit les approches les plus répandues.

2.4.3.1 Méthode du plus proche voisin

La méthode du plus proche voisin (NN-*Nearest Neighbor*) se base uniquement sur la distance de Mahalanobis (cf. (2.31)) pour déterminer les associations piste-cible. L’algorithme associe à la piste x_j la plus proche cible observée z_i à l’intérieur de la fenêtre d’association. Bien que cette méthode est simple à mettre en oeuvre, elle peut générer des associations multiples où une mesure observée peut être associée à plusieurs pistes. La figure 2.27 illustre un cas d’associations multiples où la mesure z_4 est proche des deux pistes x_1 et x_2 . Par conséquent, la mesure perçue z_4 sera associée aux deux pistes x_1 et x_2 en même temps. Face à cet inconvénient, le problème d’association doit être résolu d’un point de vue global. Au lieu de déterminer une association de manière individuelle pour chaque piste, la méthode GNN [Blackman, 1986] génère l’ensemble des associations en minimisant un coût global. Ce dernier représente la somme des coûts des différentes associations. Dans [Blackman et Popoli, 1999], le coût d’une association entre la piste x_j et la mesure z_i est quantifié par la distance statistique $d_{G_{i,j}}^2$ (cf. (2.34)). Cette distance est définie par la somme entre la distance de Mahalanobis $d_{i,j}^2$ et le logarithme du déterminant de la matrice d’innovation S_j (cf. (2.32)). L’objectif est de déterminer les associations ($piste_j, cible_i$) minimisant la somme des distances $d_{G_{i,j}}^2$.

$$d_{G_{i,j}}^2 = d_{i,j}^2 + \ln |S_j| \quad (2.34)$$



FIGURE 2.28 – Détection des formes triangulaires d’un panneau de cédez le passage.

2.4.3.2 Méthodes probabilistes

Dans un environnement réel, les observations ne sont pas parfaites mais souvent bruitées et parfois perturbées par la présence de fausses mesures. Ces dernières peuvent être proches (au sens de la distance de Mahalanobis) des mesures prédites et donc associées aux objets connus par les méthodes (G)NN. Ce choix d’association entraîne des erreurs d’estimation des pistes concernées.

Ce type de problème peut être traité par les méthodes probabilistes où l’ensemble des observations de la fenêtre d’association contribuent à la mise à jour de la piste considérée, au lieu de la plus proche seulement, réduisant ainsi l’influence des fausses mesures. Cette approche prend en charge aussi le cas des multiples observations où plusieurs mesures peuvent être générées pour un même objet, par exemple la double détection d’un panneau cédez la passage (cf. Figure 2.28). Le fait d’intégrer l’ensemble des observations lors de la mise à jour améliore le suivi de la piste. Introduite par Bar-Shalom et Tse [Bar-Shalom et Tse, 1975], l’approche probabiliste sous-optimale d’association piste-à-multi-observations porte le nom d’association probabiliste de données (PDA). La méthode PDA nécessite une reformulation du filtre afin d’estimer l’état de la piste à partir de l’ensemble des associations candidates, présentes dans la fenêtre d’association. La solution \hat{x}_j est de calculer une estimation globale à partir d’une somme pondérée d’estimations locales relatives aux mesures z_i . Une probabilité d’association p_{ji} est définie pour l’hypothèse associant la piste x_j aux observations z_i où $i = 1 \dots n$, et n le nombre d’observations présentes dans la fenêtre d’association correspondante à la piste x_j :

$$p_{ji} = \begin{cases} \frac{b}{b + \sum_{l=1}^n \alpha_{il}} & i = 0 \\ \frac{\alpha_{ij}}{b + \sum_{l=1}^n \alpha_{il}} & 1 \leq i \leq n \end{cases} \quad (2.35)$$

où p_{j0} représente la probabilité de non-association de la piste x_j et :

$$\begin{cases} b &= (1 - P_D)\beta(2\pi)^{\frac{M}{2}}\sqrt{|S_j|} \\ \alpha_{ij} &= P_D \exp \frac{-d_{i,j}^2}{2} \\ \beta &= \beta_{NT} + \beta_{FT} \end{cases} \quad (2.36)$$

avec :

- P_D , la probabilité de détection *a priori*,
- β_{NT} , la densité de nouvelles pistes, supposée distribuée selon la loi de Poisson,
- β_{FT} , la densité de fausses mesures (alarmes), supposée distribuée selon la loi de Poisson,
- M , la dimension de la mesure,
- S_j , la matrice d'innovation de la piste x_j ,
- $d_{i,j}^2$, la distance de Mahalanobis entre la piste x_j et l'observation z_i .

Lors de l'estimation de la piste x_j , le processus d'innovation est pondéré par les probabilités d'association ($p_{ji}, i = 1 \dots n$) entre la piste considérée et les n observations présentes dans la fenêtre d'association :

$$\tilde{z}_j(k) = \sum_{i=1}^n p_{ji} \tilde{z}_{ij}(k) \quad (2.37)$$

La correction de la matrice de covariance notée $P_j(k|k)$ est modifiée aussi afin de prendre en considération l'ensemble des hypothèses d'association :

$$P_j(k|k) = p_{j0}P_j(k|k-1) + (1 - p_{j0})P_j^c(k|k) + \tilde{P}_j(k) \quad (2.38)$$

avec :

- \tilde{P}_j matrice prenant en compte l'effet des n observations validées :

$$\tilde{P}_j(k) = K_j(k) \left(\sum_{i=1}^n p_{ji} \tilde{z}_j(k) \tilde{z}_{ij}^t(k) - \tilde{z}_j(k) \tilde{z}_j^t(k) \right) K_j^t(k), \quad (2.39)$$

- P_j^c matrice de correction de la covariance d'erreur similaire à celle de KF classique :

$$P_j^c(k|k) = (I - K_j(k)H_k)P_j(k|k-1) \quad (2.40)$$

La méthode PDA est une approche mono-objet où une unique piste est associée à l'ensemble des observations disponibles sans prendre en considération les autres pistes existantes. Dans un environnement complexe, plusieurs pistes peuvent se croiser générant ainsi un chevauchement des fenêtres d'association. Par conséquent, des observations non-associées à la piste considérée peuvent être utilisées pour la correction du filtre, ce qui génère des erreurs d'estimation. Pour résoudre ce problème une extension de la méthode PDA a été proposée dans [Fortmann *et al.*, 1980] et connue sous le nom d'association probabiliste conjointe de données (JPDA).

La méthode JPDA prend en considération toutes les hypothèses d'association envisageables entre les pistes et les observations disponibles, et leur attribue une probabilité d'association p_{ji} . La méthode prend également en compte le cas d'observations correspondant à des fausses alarmes. La différence par rapport à la méthode PDA réside dans le calcul des probabilités p_{ji} , qui doivent tenir compte de toutes les pistes ayant des observations valides en commun.

Dans le domaine des ADAS, particulièrement des TSR, il est très intéressant, dans le cas de *Suivi temporel*, de savoir si des panneaux sont apparus ou ont disparu de notre environnement. Cependant, les méthodes d'association présentées précédemment se basent sur une hypothèse restrictive d'un nombre connu et invariant de pistes à l'instant k , ce qui empêche toute création et suppression de pistes. Dans ce contexte, la gestion des apparitions et disparitions d'objets n'est pas prise en charge, ce qui procure des erreurs d'estimation des pistes et conduit ainsi à une mauvaise interprétation de notre environnement routier. Face à cette problématique, une des solutions consiste à prendre en considération l'ensemble des hypothèses possibles où chaque observation peut être à l'origine : d'une fausse alarme, d'une nouvelle piste, ou d'une piste existante.

2.4.3.3 Le pistage à hypothèses multiples

L'approche du pistage à hypothèses multiples (MHT-*Multiple Hypothesis Tracking*) a été initiée par Reid [Reid, 1979] afin de mieux gérer l'initialisation des pistes (apparition de nouvelles pistes) et l'ambiguïté dans les associations. Il existe deux implementations du MHT : l'une orientée hypothèse dite HOMHT (*Hypothesis Oriented MHT*) et l'autre orientée piste dite TOMHT (*Track Oriented MHT*) [Chen et al., 2014]. L'HOMHT ne cherche plus à considérer les pistes mais à y associer les observations. Des hypothèses d'associations sont générées pour chaque observation et une probabilité *a posteriori* leur est attribuée. Le principe de HOMHT est de maintenir les hypothèses d'association (cible-piste) les plus probables. Trois hypothèses sont considérées pour chaque observation :

- l'observation est issue d'une piste existante,
- l'observation correspond à une nouvelle piste,
- l'observation est une fausse alarme.

Contrairement au HOMHT, la méthode TOMHT construit un arbre pour chaque piste et les branches représentent les différentes observations (mesures) qui peuvent lui être associées à chaque itération. Dans ce contexte, la construction des arbres se base sur deux hypothèses seulement :

- l'observation est issue d'une piste existante,
- l'observation correspond à une nouvelle piste,

L'approche MHT suit la croissance des hypothèses par une structure d'arbre dont chaque branche est une hypothèse générée à partir d'une observation disponible à chaque itération. Dans ce contexte, il est évident que la taille de l'arbre croît de manière exponentielle au cours du temps ce qui rend très complexe sa gestion. Pour réduire cette complexité deux solutions sont envisageables :

- garder les N hypothèses les plus probables,
- garder les hypothèses dont la probabilité est supérieure à un seuil prédéfini.

Le MHT permet l'initialisation de nouvelles pistes ce qui constitue un point fort lorsque le nombre de cibles n'est pas connu *a priori* comme dans les systèmes TSR. Néanmoins, dans un environnement complexe et bruyant tel qu'une scène routière, les informations provenant des capteurs et des algorithmes de perception sont très souvent imprécises, incomplètes et incertaines. Ce contexte n'est pas propice aux méthodes probabilistes qui se basent sur la certitude des informations et/ou la connaissance des modèles d'erreurs pour calculer la probabilité des associations possibles. Une des solutions est l'association crédale basée sur la *théorie de l'évidence*, dite aussi *théorie des croyances* ou *théorie de Dempster-Shafer* [Shafer, 1976]. En effet, cette théorie est particulièrement adaptée à ce type de problématique en offrant des mécanismes de représentation de l'incertitude, de l'imprécision des informations, de la méconnaissance ou encore du conflit entre les sources. De plus, l'association crédale gère les apparitions ainsi que les disparitions des pistes.

2.4.3.4 Association crédale

Introduite par Dempster [Dempster, 1967] et formalisée mathématiquement par Shafer [Shafer, 1976], la théorie des croyances repose sur la modélisation de la croyance en un événement. Cette modélisation se réalise à partir de fonctions de masse permettant la représentation de la connaissance acquise, ainsi que l'ignorance sur l'environnement (cf. Section 4.3). Une telle technique de modélisation rend la phase de décision plus pertinente et plus juste [Mourllion, 2006].

La théorie des croyances a été utilisée dans plusieurs travaux d'association multi-objets [Rombaut, 1998], [Royère *et al.*, 2000], [Mourllion, 2006], [Mercier *et al.*, 2011], [Daniel et Lauffenburger, 2012] et [Denoeux *et al.*, 2014]. L'association crédale évalue les différentes hypothèses d'association à partir de leur croyance. Cette approche permet non seulement d'évaluer la pertinence ou confiance en une association mais également de considérer les ambiguïtés et le conflit existants. Cette méthode présente plusieurs avantages [Cherfaoui, 2009] :

- c'est une méthode globale qui permet de prendre en compte toutes les pistes et toutes les mesures.
- elle gère les apparitions et les disparitions des pistes.
- elle gère l'incertitude sur les données sans faire d'hypothèse sur les modèles d'imprécision utilisés.
- elle peut être appliquée à d'autres type d'association (*piste-à-piste* ou *mesure-à-mesure*).

Dans le contexte des applications TSR, il est clair que l'association crédale semble être la plus adaptée. En effet, contrairement aux autres méthodes d'association, cette approche permet de modéliser aisément l'imprécision et l'incertitude des résultats de la détection ce qui procure une meilleure interprétation de la scène routière. C'est pourquoi, nous avons opté pour le développement d'un système de pistage croyantiste. Plus précisément, nous avons choisi le formalisme du "*Modèle de Croyances Transférables*" de Smets [Smets et Kennes, 1994] présenté en détail dans le chapitre 4 pour le développement de notre association crédale.

2.5 Conclusion

Ce chapitre a dressé un état de l'art complet des étapes de *Détection* et de *Suivi temporel* basé sur les récentes avancées du domaine. Dans la première partie du chapitre, nous avons synthétisé les récentes méthodes de *Détection* en les classant dans trois catégories selon la principale caractéristique visuelle (couleur, forme, apparence) utilisée. La première catégorie concerne les approches couleurs qui exploitent l'information couleur des panneaux afin de les localiser dans les images par des techniques de segmentation. La deuxième catégorie regroupe les approches contours qui recherchent les formes géométriques des panneaux en exploitant l'information contour. La dernière catégorie concerne les approches descripteurs basées sur un classifieur afin d'apprendre les différentes apparences que peut avoir un type de panneau. Dans une situation réelle, la *Détection* est influencée par plusieurs facteurs. En effet, les problèmes d'occultation, les variations climatiques et d'intensité de lumière ainsi que l'absence de l'information de profondeur de la scène rendent cette opération non-triviale. Dans ce contexte, la *Détection* n'est jamais parfaite et un nombre important de FP dégradent les performances du système TSR. L'idéal étant d'avoir un système ayant un fort taux de détection (TDPI et/ou TDPP) avec un faible nombre de FP. Nous présenterons par la suite (chapitre 3) une nouvelle approche de *Détection* basée sur les contours et dont les performances sont similaires à ceux des récents algorithmes.

La deuxième partie de ce chapitre a mis en évidence que le suivi multi-cibles pouvait permettre une meilleure robustesse des applications TSR face aux erreurs de *Détection* et de *Reconnaissance*. Pourtant, l'état de l'art des travaux récents a montré que le *Suivi temporel* n'était que peu utilisé dans ce contexte, encore moins pour la suppression de FP. Ces derniers sont considérés comme des hallucinations dont le mouvement est aléatoire dans les séquences d'images fournies par la caméra. À partir de ce constat, nous avons opté pour l'intégration d'un système de *Suivi temporel* à notre détecteur (cf. Chapitre 3.2). La méthode proposée met l'accent sur le processus d'association de données afin de faire face aux problèmes liés à la détection : détections multiples d'un panneau, occultations temporaires, FP, etc. Ce processus doit gérer l'apparition et la disparition des panneaux ainsi que la mise en correspondance entre les panneaux connus (pistes) et les panneaux perçus (cibles). Compte tenu de l'imperfection (imprécision et incertitude) de la *Détection* dans la pratique, nous avons choisi l'utilisation de la théorie des croyances pour l'association de données afin de pouvoir modéliser notre ignorance sur l'environnement. Le chapitre 4 présente le modèle transférable de croyances que nous avons utilisé pour l'association de données, le principe de notre stratégie de *Suivi temporel*, ainsi qu'une quantification de l'impact apporté par notre approche de suivi sur la *Détection*.

Chapitre 3

Détection de panneaux par codage de contours

3.1 Introduction

Les panneaux se caractérisent par deux principaux aspects visuels : la couleur et la forme, facilitant ainsi leur distinction par les usagers (conducteurs et piétons). Néanmoins, pour un système TSR la *Détection* de panneaux est une étape non-triviale. En effet, plusieurs facteurs comme les conditions de visibilité, les problèmes d'occultation, et l'altération de la couleur des panneaux rendent la distinction difficile entre ces derniers et l'arrière-plan, générant ainsi un certain nombre de FP. Contrairement aux récentes méthodes qui combinent l'ensemble des caractéristiques visuelles (couleur et forme) afin d'améliorer la *Détection*, nous adoptons dans cette thèse une approche basée uniquement sur la forme.

La première partie du chapitre est consacré à la nouvelle méthode de *Détection*. Tout d'abord, nous présentons son principe qui consiste à reconnaître la symétrie des panneaux en se basant sur le codage de contours. Ce processus de codage est fonction des orientations du gradient et permet de reconnaître les formes triangulaires et circulaires des panneaux. La deuxième partie du chapitre montre l'intérêt de notre approche en la comparant avec des méthodes récemment publiées sur deux base de données publiques et une base de données privée.

3.2 Approche proposée : codage de contours

Dans le cadre de ce travail de thèse, nous avons développé une nouvelle approche pour la *Détection* de panneaux. Au cœur de cette approche, l'idée est que les panneaux triangulaires possèdent deux côtés symétriques, et qu'en inversant l'un des deux côtés par rapport au sommet du haut du triangle,

ces deux cotés s'alignent pour former un segment suffisamment long pour être détectable. Cette idée a été vérifiée, et nous l'avons modélisé au sein d'une méthode que nous avons appelé RSLD (*Ransac Symmetric Lines Detection*). Notre modélisation s'appuie sur un codage de contours, et a été publiée dans la revue *Machine Vision and Application* [Boumediene et al., 2013]. Par la suite nous avons repris le principe du codage de contours pour étendre la *Détection* aux panneaux circulaires. Nous pouvons inscrire notre méthode comme une approche contours, pour la lier à l'état de l'art présenté dans ce mémoire (cf. Chapitre 2).

Cette section présente ces contributions suivant l'organisation suivante. Le principe général est décrit, puis le codage de contours sur lequel nous nous basons pour vérifier le critère de symétrie. Ensuite, nous présentons la classification des coins qui identifie les sommets supérieurs et inférieurs des triangles localisant ainsi la position d'inversion d'un des deux cotés. Par la suite, nous décrivons notre méthode de détection de panneaux triangulaires qui est basée sur l'algorithme RSLD ainsi que la méthode de détection de panneaux circulaires.

3.2.1 Principe

Un panneau est riche en informations visuelles, outre la couleur et les contours, il est possible d'extraire des points d'intérêt (coins) qui représentent une double discontinuité de l'intensité des pixels. Ces coins peuvent être observés au niveau des contours externes d'un panneau, notamment d'un panneau triangulaire où les coins correspondent à ses sommets. Le centre des panneaux reflète lui aussi des coins qui sont liés à la présence des pictogrammes (cf. Figure 3.1). À partir de ce constat nous proposons d'utiliser les coins pour définir les positions candidates des panneaux. Autour de ces dernières une distribution particulière des contours est recherchée afin de vérifier la présence d'un panneau. La figure 3.2 illustre le diagramme de notre approche. En premier, la *Détection de coins* permet de définir les positions candidates qui seront par la suite classées, lors de l'étape de *Classification de coins*, en deux catégories : les sommets supérieurs et inférieurs des triangles et les centres des cercles. Les sommets sont exploités par l'algorithme *RSLD* afin de vérifier la présence d'un panneau triangulaire en détectant ses deux segments qui sont symétriques par rapport au sommet considéré. Ce processus utilise *l'image codée* qui représente le résultat du *Codage de contours*. Cette dernière permet de prendre en considération uniquement les contours appropriés à la forme recherchée. La détection de panneaux circulaires exploite les centres candidats autour desquels une distribution particulière des contours est recherchée à l'aide d'une approche de *template matching*.

3.2.2 Détection de coins

Un point d'intérêt ou coin est un point dans l'image où son voisinage est riche en terme d'information gradient. Le détecteur de points d'intérêt le plus connu est le détecteur de Harris [Harris et



FIGURE 3.1 – Détection des points d'intérêt liés aux panneaux.

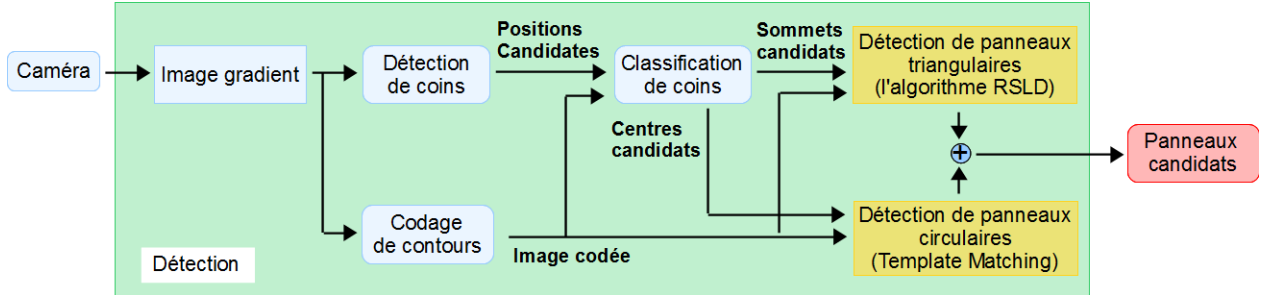


FIGURE 3.2 – Diagramme de l'approche proposée.

[Stephens, 1988]. Ce détecteur est basé sur une matrice de covariance M (cf. (3.1)) qui décrit la distribution gradient (I_x, I_y) dans le voisinage local d'un pixel de l'image (I) . Ce voisinage dépend de la dimension de la fenêtre Gaussienne W utilisée. Les valeurs propres, λ_1 et λ_2 , de M , représentent les principales composantes du gradient dans deux directions orthogonales. Deux valeurs propres significatives indiquent la présence de deux directions principales dans le voisinage, ce qui représente un coin à la position correspondante. Dans [Harris et Stephens, 1988], une alternative au calcul des valeurs propres est proposée avec un critère R décrit dans (3.2) où k est fixé à 0.04. Les coins sont détectés en recherchant les maxima locaux dans la matrice R . Un exemple de détection de coins est présenté par la figure 3.1.

$$\left\{ \begin{array}{l} I_x = \frac{\partial I}{\partial x}, I_y = \frac{\partial I}{\partial y} \\ \overline{I_x^2} = I_x^2 \otimes W, \overline{I_y^2} = I_y^2 \otimes W \\ \overline{I_x I_y} = I_x I_y \otimes W \\ M = \begin{bmatrix} \overline{I_x^2} & \overline{I_x I_y} \\ \overline{I_x I_y} & \overline{I_y^2} \end{bmatrix} \end{array} \right. \quad (3.1)$$

$$\left\{ \begin{array}{l} (\lambda_1 + \lambda_2)^2 = Tr^2(M) \\ \lambda_1 * \lambda_2 = Det(M) \\ R = Det(M) - k * Tr^2(M) \end{array} \right. \quad (3.2)$$

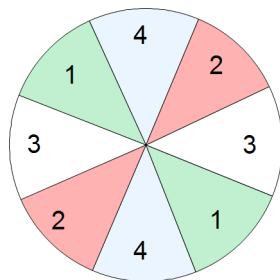


FIGURE 3.3 – Illustration de l’espace de direction des contours divisé selon les 4 principales classes $\{1, 2, 3, 4\}$.

3.2.3 Codage de contours

Notre approche détecte les panneaux en reconnaissant leur forme, ce qui nécessite la détection de contours. Un contour représente une discontinuité de l’intensité des pixels mais ne donne pas d’information sur sa direction, qui peut être pertinente lors de la reconnaissance. En effet, cette information permet par exemple de distinguer les contours liés à la base d’un triangle des contours représentant ses deux côtés diagonaux, ce qui peut raffiner la détection tout en minimisant les erreurs. Par conséquent, nous proposons de générer une nouvelle image, nommée *image codée*, où chaque pixel est codé selon l’orientation de son gradient. Ce processus de codage se base sur la définition de cinq classes $\{0, 1, 2, 3, 4\}$. Les pixels non-contours sont codés par la classe 0, et les pixels contours sont codés selon les 4 directions illustrées par la figure 3.3. Les contours diagonaux sont représentés par les classes 1 et 2. Finalement, les contours horizontaux et verticaux sont représentés par les classes 3 et 4.

Cette étape présente l’intérêt de s’effectuer au moyen des calculs intermédiaires effectués pour la détection de coins, optimisant ainsi le temps de calcul. Le processus de codage peut se résumer avec l’équation suivante :

$$Classe = A B (2 - C) + 3A (1 - B) + 4B (1 - A) \quad (3.3)$$

avec

- $A = 1$ Si $I_x^2 > T$ Sinon 0,
- $B = 1$ Si $I_y^2 > T$ Sinon 0,
- $C = 1$ Si $I_x I_y < 0$ Sinon 0,
- T : seuil prédéfini qui permet de retenir les gradients significatifs.

3.2.4 Classification de coins

Les points d’intérêt sont souvent utilisés pour la mise en correspondance d’images ou le recalage d’images. Ces méthodes décrivent chaque coin par un vecteur de descripteurs afin de le comparer avec des vecteurs représentatifs contenus dans une base de données [Lin et al., 2010]. Nous adoptons

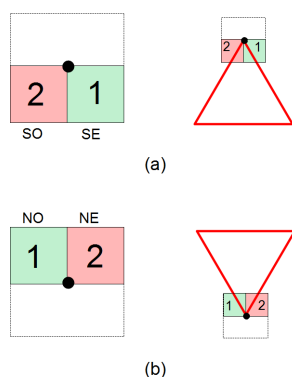


FIGURE 3.4 – Distribution de contours utilisée pour le codage des coins. (a) La distribution de contours pour la détection d’un sommet supérieur (triangle). (b) La distribution de contours pour la détection d’une pointe (triangle inversé).

ce type d’approche pour distinguer les coins liés aux triangles des autres, ce qui permet d’affiner les positions candidates et d’optimiser le temps de traitement. En effet, la *Détection* de panneaux devient sélective puisque la forme recherchée à une position donnée dépendra de la classe de cette dernière (coin lié à un triangle ou à un cercle). Pour ce faire, nous définissons le vecteur descripteur d’un coin à partir de *l’image codée*. Un voisinage est défini autour de chaque position candidate afin de détecter les sommets (supérieurs et inférieurs) des triangles du reste des coins selon une distribution des contours. Cependant, ce type d’approche n’est pas adapté à la *Détection* de panneaux puisque le vecteur descripteur ne contient pas uniquement des informations liées au panneau mais aussi celles de l’arrière plan de la scène. Ce constat justifie la proposition d’un voisinage adapté pour une meilleure description.

Deux distributions de contours sont proposées dans un voisinage adapté pour détecter les positions des panneaux triangulaires et triangulaires inversés. La figure 3.4 présente les sous-régions pour la classification des coins. Dans ce cas, le voisinage considéré est défini par deux sous-régions positionnées sous le sommet supérieur où l’information liée au panneau triangulaire doit être présente. Pour chaque coin, si la classe 2 est dominante dans la sous-région *Sud-Ouest* (SO) et la classe 1 dans la sous-région *Sud-Est* (SE) (cf. Figure 3.4 (a)), alors le coin est considéré comme étant le sommet supérieur d’un triangle. Un processus similaire est appliqué sur le voisinage illustré dans la figure 3.4 (b) pour la détection des pointes (sommet inférieur) liées aux triangles inversés. Un exemple de classification de coins sur une image de synthèse est présenté dans la figure 3.5. A l’issue de la classification, les coins non reconnus sont considérés comme les centres candidats des panneaux circulaires.

La classification de coins donne une information sur le type de panneau à rechercher ce qui permet d’optimiser le temps de traitement. En effet, pour chaque position il est possible de sélectionner l’algorithme approprié (détecteur de triangles ou de cercles) en fonction du type de coin (sommet supérieure d’un triangle, sommet inférieur d’un triangle inversé, ou centre d’un cercle). Une comparaison a été effectuée entre le nombre total des positions candidates, détectées par le détecteur

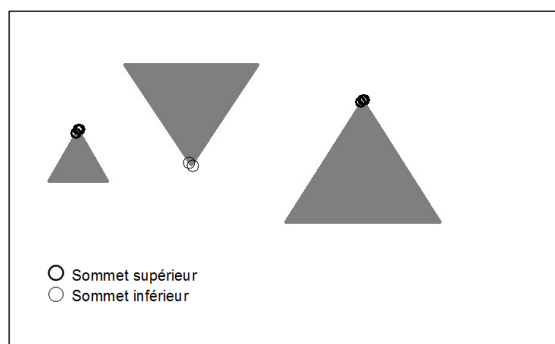
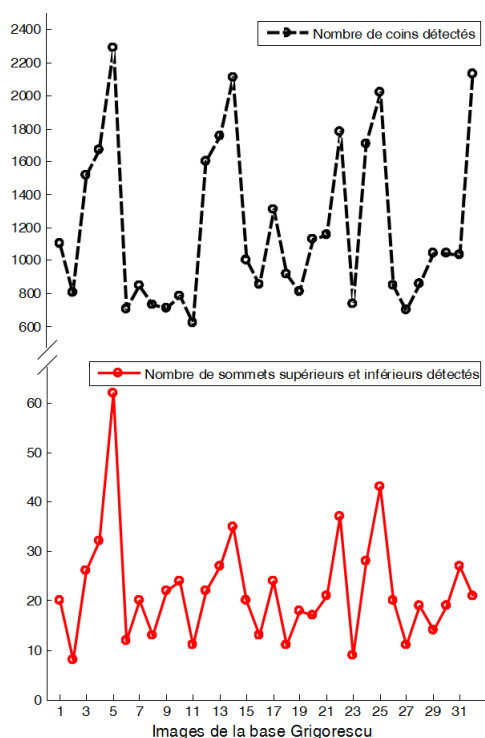


FIGURE 3.5 – Illustration d'un exemple de codage de coins sur une image de synthèse.

FIGURE 3.6 – Résultats de la classification de coins sur 32 images de la base de *Grigorescu* [Grigorescu et Petkov, 2003]. La courbe supérieure représente le nombre total des coins détectés par le détecteur de Harris. La courbe inférieure représente le nombre de positions retenues suite à la classification de coins pour la *Détection* de panneaux triangulaires

de Harris, et le nombre des positions retenues pour la *Détection* des panneaux triangulaires. Cette comparaison a été réalisée sur 32 images de la base *Grigorescu* [Grigorescu et Petkov, 2003] contenant 40 panneaux triangulaires. La figure 3.6 présente les résultats où il est utile de préciser que l'ensemble des sommets supérieurs et points respectivement des panneaux triangles et triangles inversés ont été correctement sélectionnés suite à la classification des coins. Ces résultats montrent clairement que le nombre de positions sélectionnées pour la *Détection* de panneaux triangulaires (et triangulaire inversés) a été réduit significativement par rapport au nombre total des coins détectés par le détecteur de Harris, ce qui représente un gain en temps de traitement par la suite.

3.2.5 Détection de panneaux triangulaires

3.2.5.1 Principe

Pour la *Détection* de panneaux triangulaires nous exploitons le critère de symétrie. En effet, un panneau de danger ou de cédez le passage est caractérisé par une symétrie axiale, liée à ces deux côtés diagonaux, que nous détectons par *l'algorithme RSLD*. Ce dernier se base sur le codage de contours pour vérifier la présence de symétrie autour des sommets (supérieurs et inférieurs) identifiés lors de la classifications de coins. Néanmoins, sans l'information de profondeur de la scène il est impossible de connaître l'échelle d'un panneau et donc difficile de détecter correctement sa symétrie. Face à cet inconvénient, nous proposons une approche de *Détection multi-échelle* basée sur un processus itératif qui estime l'échelle du panneau tout en détectant sa symétrie. En plus du critère de symétrie, nous nous basons sur la *Détection de ligne de base* de la forme triangulaire afin de réduire le nombre de FP.

3.2.5.2 L'algorithme RSLD

L'idée simple est d'aligner les deux côtés diagonaux de la forme triangulaire, ce qui permet la détection d'un seul segment au lieu de deux par l'algorithme RANSAC (*RANdom SAmples Consensus*) [Fischler et Bolles, 1981]. La figure 3.7 (a) illustre ce principe où une simple inversion des coordonnées y appliquée sur les pixels de la classe 2 transforme les deux segments diagonaux en un seul. Pour les sommets inférieurs, la même procédure est appliquée sur les pixels de la classe 1 (cf. Figure 3.7 (b)). Notre approche transforme la détection de symétrie en une simple détection de ligne tout en préservant le critère de symétrie. Par conséquent, si une position candidate est réellement positionnée sur un sommet supérieur ou inférieur d'une forme triangulaire, la méthode proposée procure un segment centré autour de cette position. Si le nombre des inliers, points contours constituant le segment, est en-dessous d'un seuil s_{seg} , le segment obtenu sera court et la position candidate n'est pas valide. Le seuil s_{seg} est en fonction de l'échelle de détection (cf. Section 3.2.5.4).

Nous avons montré que cette méthode permet de réduire significativement le temps de traitement tout en ayant un bon taux de détection en utilisant RANSAC pour détecter les segments. L'algorithme RANSAC est une méthode robuste pour l'estimation de modèle à partir d'un ensemble de N points contenant des données relatives au modèle, (*inliers*), et des données non liées au modèle, (*outliers*). Cet algorithme est itératif, composé de deux étapes : génération du modèle et son évaluation. La première étape consiste à générer un modèle à partir de n points choisis aléatoirement dans l'ensemble des N points. La deuxième étape évalue le modèle obtenu pour l'ensemble des N points.

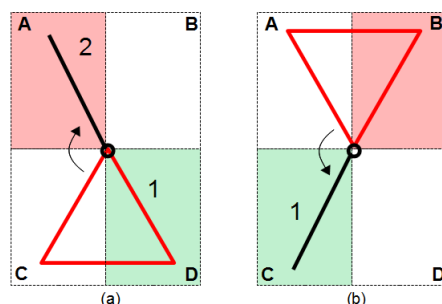


FIGURE 3.7 – Illustration de l’inversion des coordonnées y autour des coins sélectionnés. (a) représente le résultat de l’inversion verticale des coordonnées y autour d’un sommet supérieur. (b) représente le résultat de l’inversion verticale des coordonnées y autour d’une pointe.

3.2.5.3 Estimation des gains de RSLD

La méthode proposée a deux avantages : une meilleure estimation, indépendamment du pourcentage des outliers, et un temps de traitement réduit par rapport à d’autres méthodes.

Gain sur la précision

L’estimation du modèle, obtenu après l’inversion des coordonnées y , a été comparée avec celle des deux modèles représentant les côtés diagonaux d’un triangle. La comparaison est effectuée en ajoutant des “outliers” à une image de synthèse, contenant les “inliers” représentant les points contours liés aux segments diagonaux d’une forme triangulaire. Une erreur-type (RMSE-*Root-Error-Mean-Square*) de détection est calculée en fonction du pourcentage d’outliers ajouté à l’image de synthèse. Cette erreur représente la racine carré de la distance quadratique moyenne entre le modèle réel (modèle connu) des côtés diagonaux et les lignes détectés. La comparaison est effectuée entre deux approches :

- sans inversion des coordonnées y : détection des deux segments par l’application de la méthode RANSAC pour chaque segment. Deux sous-régions adaptées sont définies autour des sommets candidats et la méthode RANSAC est appliquée sur chaque sous-région séparément (cf. Figure 3.8 (a)).
- avec inversion des coordonnées y : détection d’un seul segment après la procédure d’inversion des coordonnées y (cf. Figure 3.7 (a)). La méthode RANSAC est appliquée une seule fois sur les sous-régions appropriées (cf. Figure 3.8 (b)).

Un exemple de ces deux approches est présenté dans la figure 3.8. Les inliers, liées aux deux segments diagonaux, sont générés par (3.4) avec un bruit Gaussien ($\sigma = 3$) ajouté sur les coordonnées x et y . Le bruit représente une perturbation, ce qui permet des simulations réalistes. Les mesures sont effectuées en fonction de l’augmentation du taux d’outliers. Pour cette simulation, la région d’intérêt est de 50×50 pixels avec un nombre d’inliers proche de la réalité (30). Afin d’obtenir des résultats

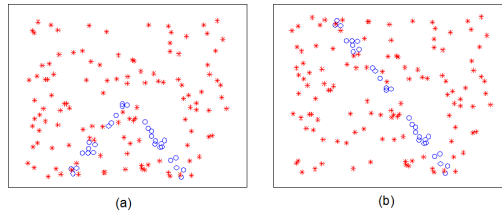


FIGURE 3.8 – Images de synthèse. les points (o) représentent les inliers, les données liées aux côtés diagonaux du triangle et bruitées par un bruit Gaussien ; les points ($*$) représentent les outliers qui sont distribués uniformément. (a) représente les côtés diagonaux d'un triangle et (b) représente ces côtés après une inversion des coordonnées y .

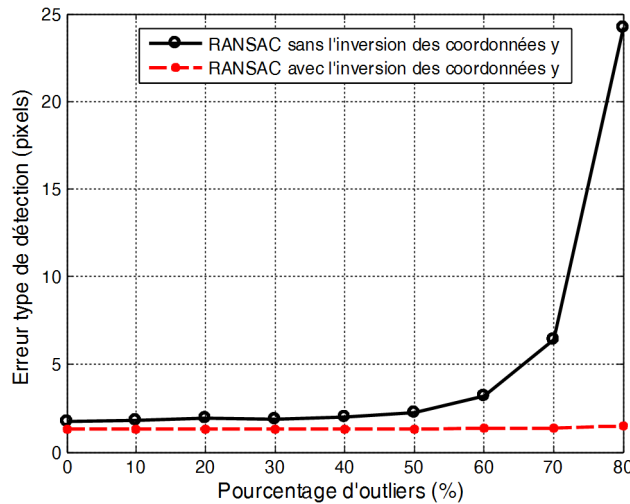


FIGURE 3.9 – Comparaison entre la détection par RANSAC avec et sans l'inversion des coordonnées y .

représentatifs, l'expérience est répétée 1000 fois pour chaque taux d'outliers.

$$\begin{cases} \tan(\frac{\pi}{3})x + y = 0 \\ -\tan(\frac{\pi}{3})x + y = 0 \end{cases} \quad (3.4)$$

La figure 3.9 illustre les résultats des expériences. La possibilité d'avoir une meilleure précision en détectant un seul segment au lieu de deux segments courts est confirmée. La courbe représentant une faible erreur type malgré l'augmentation du taux d'outliers correspond à l'approche basée sur l'inversion des coordonnées y .

Gain en temps de calcul

Pour confirmer le gain en temps de calcul de la proposition, le nombre d'itérations des deux approches, sans et avec inversion, sont comparés. Le nombre d'itérations K de la méthode RANSAC dépend de trois paramètres : pourcentage des inliers w (cf. (3.5)), complexité du modèle n , et la probabilité p que RANSAC détecte le modèle recherché [Fischler et Bolles, 1981]. Une ligne peut être modélisée à l'aide de 2 points seulement, $n = 2$. La probabilité p est définie telle que $(1 - p) = 10^{-3}$

(valeur par défaut de RANSAC). Avec cette configuration, $K = 24$ si le taux d'inliers $w = 50\%$, et $K = 687$ si les inliers représentent que 10% de l'ensemble des données.

$$K = \frac{\log(1-p)}{\log(1-w^n)} \quad (3.5)$$

L'algorithme RANSAC s'exécute en deux étapes : génération et évaluation. La première consiste à tirer au sort deux points à partir de l'ensemble des données (inliers et outliers) pour générer un modèle (une droite), la seconde détermine le nombre de données proches du modèle généré afin d'évaluer sa pertinence. Le temps d'exécution de RANSAC T est estimé à l'aide de l'équation :

$$T = K (t_g + N t_e) \quad (3.6)$$

avec t_g le temps de la génération d'un modèle et t_e le temps d'évaluation de ce modèle.

Sans inversion des coordonnées y , il faut détecter les deux côtés diagonaux du triangle ce qui nécessite deux fois l'application de RANSAC comme cela est illustré dans la figure 3.7. Il est appliqué dans la sous-région C pour la détection du premier côté, et ensuite dans la sous-région D pour le deuxième côté. Les sous-régions A et B ne sont pas utilisées. Le temps de calcul global T_{CD} de la détection est défini par la somme du temps de traitement dans les deux sous-régions C et D contenant alors chacune $\frac{N}{2}$ points (cf. (3.7)).

$$\begin{aligned} T_{CD} &= T_C + T_D, \\ &= 2 K (t_g + \frac{N}{2} t_e), \\ &= K t_g + T. \end{aligned} \quad (3.7)$$

Avec inversion des coordonnées y , RANSAC est appliqué une seule fois directement sur les sous-régions A et D où le nombre global de points est N . Dans ce cas, le temps de calcul de la détection T_{AD} est obtenu par l'addition du temps d'inversion au temps de traitement T (cf. (3.8)).

$$T_{AD} = t_i + T \quad (3.8)$$

Le fait que le processus d'inversion ne soit appliqué qu'une seule fois et non pas à chaque itération RANSAC, implique que $t_i < K t_g$. Par conséquent, on déduit que $T_{AD} < T_{CD}$ ce qui signifie que la méthode est plus rapide que la détections successives de deux segments. La figure 3.10 compare le temps de calcul des deux approches en fonction du taux d'outliers. D'après les courbes de la figure 3.10 la méthode proposée nécessite moins de temps de calcul que la détection de deux segments ce qui confirme la conclusion précédente. Pour une image réelle obtenue par la caméra embarquée, le temps de calcul est réduit de 40% .

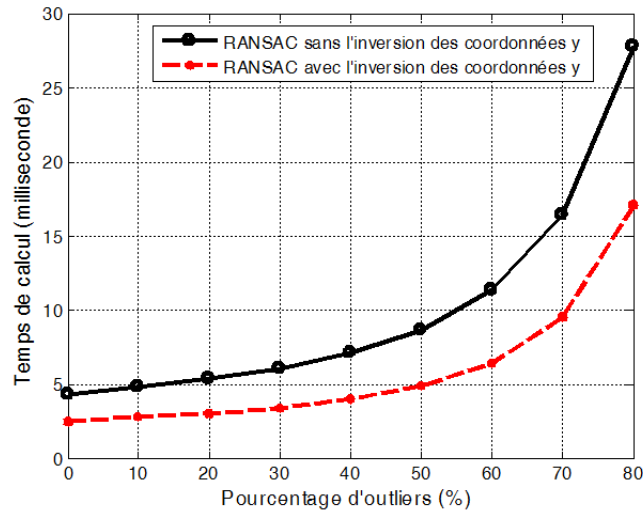


FIGURE 3.10 – Illustration du temps de calcul pour la détection (RANSAC) avec et sans l'inversion des coordonnées y . Le nombre d'inliers est de 30.

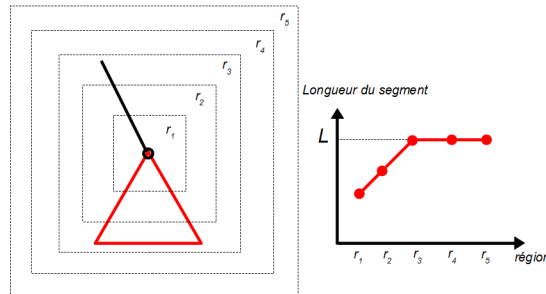


FIGURE 3.11 – Détection Multi-échelle.

3.2.5.4 Détection multi-échelle

Dans une image de scène routière, l'échelle d'un panneau dépend de sa distance à la caméra. La région autour de chaque sommet supérieur et pointe doit être définie correctement afin que la méthode RSLD puisse détecter les panneaux. L'idéal est de définir cette région en fonction de la taille du panneau mais sans l'information de profondeur il est impossible d'estimer cette taille. Cet inconvénient est résolu par un processus de détection multi-échelle. En premier, une petite région r_1 est centrée sur une position candidate où les points nécessaires à la méthode RANSAC seront extraits, comme illustré sur la figure 3.11. Si aucune ligne n'est détectée alors la position ne correspond pas à un coin d'un triangle. Sinon la longueur de la ligne détectée en r_1 est retenue. Ensuite, ce processus est répété dans les autres régions de r_2 à r_5 , présentées dans la figure 3.11, jusqu'à la stabilité de la longueur de la ligne détectée L .

La détection multi-échelle permet de détecter les côtés diagonaux d'un éventuel panneau triangulaire tout en estimant sa résolution. Cette dernière est représentée par la longueur du côté diagonal ($\frac{L}{2}$). La figure 3.13 (c) présente un exemple de résultats de la détection multi-échelle où tout objet composé de deux segments symétriques est détecté.

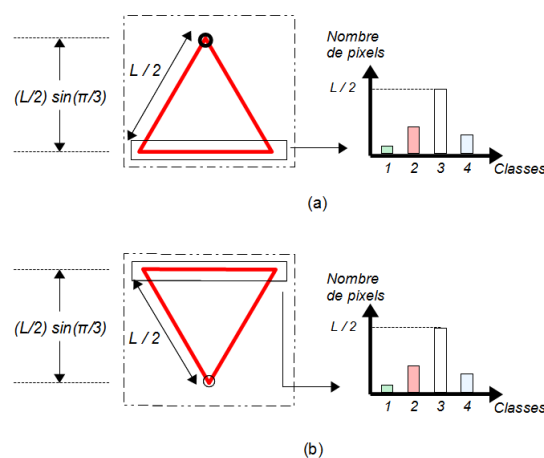


FIGURE 3.12 – Détection de la ligne de base. (a) La sous région horizontale et son histogramme pour la *Détection* d'un triangle. (b) La sous région horizontale et son histogramme pour la *Détection* d'un triangle inversé.

3.2.5.5 Détection de la ligne de base

La figure 3.13 c illustre la détection des deux diagonales d'un triangles. On constate que d'autres objets de la scène routière peuvent vérifier cette symétrie liée aux côtés et sont donc détectés, i.e. FP. Afin de les éliminer, la forme d'un triangle doit être confirmée par la présence de la ligne de base. Pour cela, on utilise la longueur $\frac{L}{2}$ d'un segment afin de déterminer la position d'une éventuelle ligne de base. En premier, une sous-région horizontale est positionnée à une distance de $\frac{L}{2} \sin(\frac{\pi}{3})$ du sommet supérieur (*pointe*) d'un triangle (*triangle inversé*) (cf. Figure 3.12). Ensuite, comme les cotés d'un triangle équilatéral sont égaux, la ligne de base est détectée si le nombre de pixels appartenant à la classe 3 est égal à $\frac{L}{2} \pm \epsilon$ avec ϵ une précision que nous avons fixé à la valeur de $0.05L$. Un exemple des résultats de la détection de la ligne de base est présenté par la figure 3.13 (d). La méthode a été aussi évaluée sur trois banques d'images et les résultats sont discutés dans la section 3.3.

3.2.6 Détection de panneaux circulaires

3.2.6.1 Principe

À partir de la *Détection* de panneaux triangulaires présentée, nous avons utilisé l'idée du codage de contours pour la *Détection* de panneaux circulaires. Chaque panneau est identifié par son pictogramme, à partir duquel des coins peuvent être extraits et sont plus au moins positionnés au centre du panneau. Par conséquent, les coins non reconnus en tant que sommets supérieurs ou inférieurs, lors de la classification de coins, sont exploités comme les centres candidats des panneaux circulaires. L'idée est de vérifier la présence d'une forme circulaire autour de chaque centre candidat par une approche de *template matching*. Cette dernière se base sur un critère de similarité entre

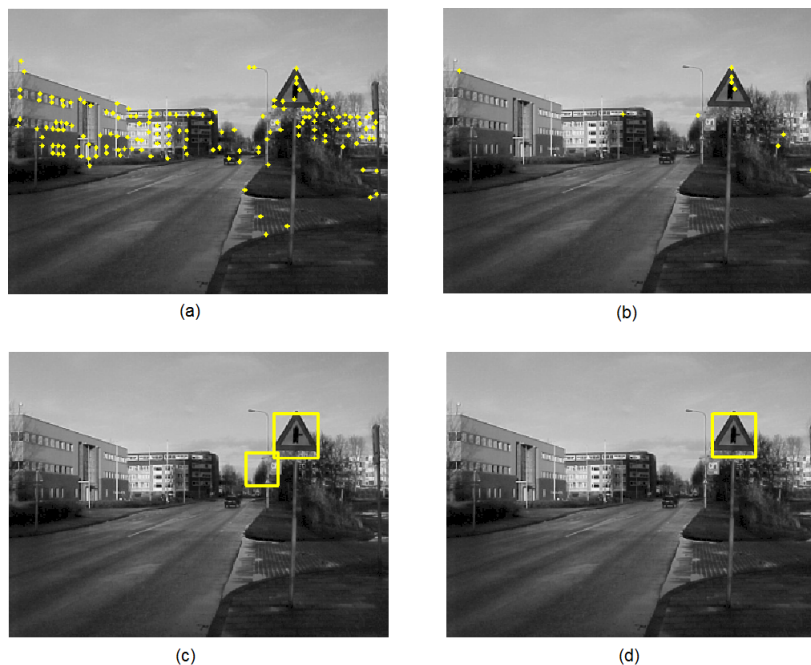


FIGURE 3.13 – Exemples de *Détection* de panneaux triangulaires. (a) Détection des positions candidates par le détecteur de Harris. (b) La classification des coins. (c) Résultat de la détection de segments diagonaux par RANSAC. (d) Résultats de la détection de la base ligne.

un modèle et une région définie autour du centre candidat. Le modèle utilisé pour représenter la forme géométrique du cercle est basé sur l'image codée calculée précédemment.

3.2.6.2 Détection par *template matching*

Le *template matching* consiste à mesurer le degré de ressemblance entre un modèle et des régions candidates. La connaissance sur les classes ou les objets recherchés est modélisée par des modèles explicites. Ici le choix du modèle est une étape importante et délicate puisqu'il doit être robuste aux variations d'illumination et à la variabilité intra-classe. L'objectif est d'approcher au mieux l'apparence des objets recherchés. Pour la détection de cercles nous proposons un modèle basé sur le codage de contours. L'idée est de décrire la forme circulaire par le type de contours détectés sur son périmètre. Le modèle est défini par un ensemble de sous-régions positionnées de manière circulaire à l'intérieur d'une région candidate englobant le panneau. Chaque sous-région doit contenir un type approprié de contour (cf. Figure 3.14).

Le processus de *Détection* consiste à mesurer la similarité de chaque région candidate, dont le centre est défini par un coin non utilisé précédemment, au modèle. Un simple comptage des contours codés permet de mesurer cette similarité. Si le nombre de contours recherchés dans une sous-région est supérieur à un seuil prédéfini alors cette dernière est marquée comme positive. La correspondance avec le modèle est validée si 85% des sous-régions sont positives. Afin de pouvoir détecter les panneaux à plusieurs échelles, le modèle est défini en plusieurs tailles. Puisque la résolution des

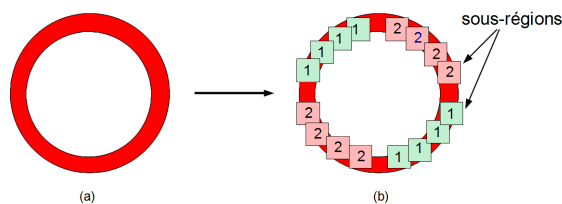
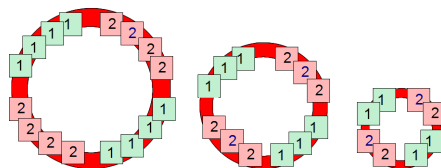
FIGURE 3.14 – Le modèle (*template*) proposé pour la *Détection* de panneaux circulaires.

FIGURE 3.15 – Exemples de modèle à plusieurs échelles.

sous-régions est fixe, leur nombre varie selon la taille du modèle (cf. Figure 3.15). Cependant, les coins liés aux pictogrammes ne sont pas forcément localisés au centre du panneau. Face à cet inconvénient, chaque région candidate est glissée sur un voisinage défini autour du coin considéré.

3.2.7 Groupement des résultats de la détection

La sortie d'un détecteur est un ensemble de régions, contenant les panneaux candidats, définies par leur position et échelle dans l'image courante. Normalement, chaque région doit encadrer un panneau qui sera identifié lors de l'étape de reconnaissance. Cependant, dans la pratique un détecteur peut générer plusieurs régions dans le voisinage d'un même panneau (cf. Figure 3.16 (a et b)), ce qui peut être interprété comme une redondance d'information. Reconnaître le même panneau plusieurs fois à partir de la même image n'a aucun sens ni d'intérêt pour le système, bien au contraire il constitue une perte de temps de calcul. Par conséquent, il est nécessaire de grouper les régions proches, liées au même panneau, avant de les transmettre à l'étape de classification. L'algorithme *Mean Shift* a été utilisé dans plusieurs applications de vision pour le groupement des détections (régions) [Dalal, 2005], [Negri, 2008], et [Ruta et al., 2011]. Le principe est de représenter, dans un premier temps, les régions dans un espace 3D en fonction de leur position et échelle. Ensuite, un processus itératif groupe les régions proches dans cet espace 3D pour déterminer une position et une échelle moyennes.

Nous avons appliqué l'algorithme *Mean Shift* tel qu'il a été implémenté dans [Negri, 2008] pour la détection de véhicules. La représentation des régions dans l'espace 3D nécessite une normalisation de leur échelle. En effet, comme l'idée de base est de grouper les régions dont les centres coïncident, il ne faut pas que la différence d'échelle soit trop importante sinon le groupement ne se fera pas correctement. Par conséquent, nous proposons de normaliser l'échelle s de chaque région par la formule suivante :

$$z = \frac{\log(s/s_{min})}{\log(s)} \quad (3.9)$$

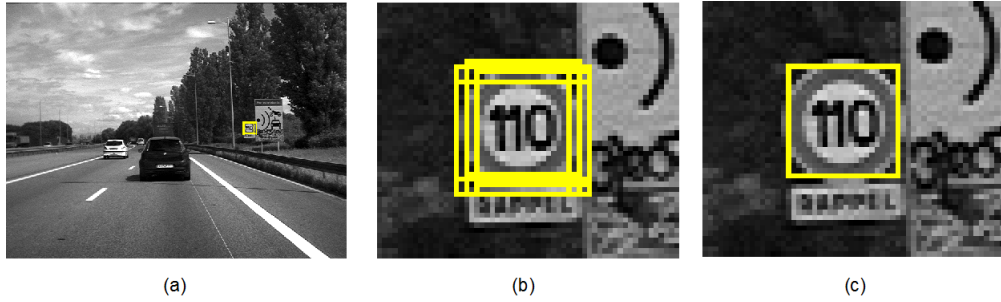


FIGURE 3.16 – Exemple de groupement des détections. (a) *Détection* d'un panneau circulaire à partir d'image de scène routière, (b) zoom sur le résultat de la *Détection*, et (c) celui du processus de groupement.

où s_{min} est l'échelle de la plus petite région détectée par le système. Cette normalisation permet de rapprocher dans l'espace 3D les régions dont le centre coïncide bien qu'il y ai une différence d'échelle. Après cette première étape de normalisation, les positions et échelles moyennes sont déterminées à l'aide du processus itératif de *Mean Shift* (cf. Algorithme 4). Le principe est de sélectionner de manière itérative les régions ($r_i = [x_i, y_i, z_i]$) proches d'une régions moyenne r_m^t selon une distance euclidienne ($|r_i - r_m^t|$)(étapes 1 – 3). A chaque itération t une position et une échelle moyennes sont calculées (étapes 4 – 6). Lors de ce calcul nous utilisons un vecteur de pondération w défini par le facteur d'échelle ($w_i = \frac{s_i}{s_{min}}$) des régions, ce qui permet de diminuer l'influence des régions trop grandes. L'algorithme converge lorsque l'ensemble des régions sont visitées ($T(i) = 1$). La figure 3.16 (c) illustre un exemple de ce processus de groupement de régions.

Algorithme 4 Mean Shift

Donnée : Ensemble des N régions $r_i = [x_i, y_i, z_i]$

Résultat : Ensemble des groupes de régions

Initialisation des poids $w_i = \frac{s_i}{s_{min}}, i = 1, \dots, N$

Initialisation du vecteur des visites, $T(i) = 0, i = 1, \dots, N$

Tant que il existe un $T(i) == 0$ **faire**

1. Choisir un r_i tel que $T(i) == 0$
2. Définir la région moyenne $r_m^t = r_i$
3. Trouver l'ensemble des r_j proches de $r_m^t, |r_m^t - r_j| < D$
4. Calculer la nouvelle région r_m^{t+1} , pondérée par w
5. Si, $|r_m^{t+1} - r_m^t| > \epsilon, r_m^t = r_m^{t+1}$ et on répète (3) et (4)
6. $T(i) = 1, \forall i$ de l'ensemble de la moyenne
7. La moyenne r_m^{t+1} est conservée dans la liste des régions groupées.

Fin Tant que

3.3 Évaluation expérimentale

Nous proposons dans cette section une évaluation des performances de nos méthodes de *Détection* de panneaux. Pour cela nous les comparons avec les récents algorithmes [Loy et Barnes, 2004], [Belaroussi et Tarel, 2009] et [Houben, 2011]. Aussi nous évaluons leur robustesse en présence de différentes dégradations du système d’acquisition d’images.

3.3.1 Bases de données

Nos algorithmes de *Détection* sont évalués au moyen de trois bases de données dont deux sont publiques et utilisées par certains auteurs comme benchmark :

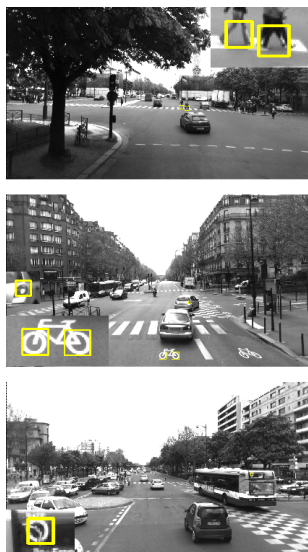
- la base de *Grigorescu* est composée de 48 images couleurs d’une résolution de 360×270 pixels et contenant 57 panneaux dont 40 triangulaires et 17 circulaires [Grigorescu et Petkov, 2003].
- la base de *Stéréopolis* est constituée d’une séquence de 847 images couleurs d’une résolution de 1920×1080 pixels [iTowns, 2009]. La base est réalisée en milieu urbain seulement, ville de Paris. Les panneaux triangulaires apparaissent sur 27 images et les panneaux circulaires apparaissent sur 179 images. L’application de nos méthodes sur ces deux premières bases de données nécessite une étape de conversion en niveau de gris.
- la dernière base utilisée est une base de données privée du laboratoire MIPS. Elle est composée de deux séquences vidéos, plus de 18000 images, d’une résolution de 640×480 pixels. Cette base a été réalisée avec un des véhicules instrumentés du laboratoire MIPS dans l’un des départements de France, le Haut-Rhin. La caméra embarquée a pour référence ECK100 (*Evaluation Camera Kit*) de *Sansata Technologies*, elle est monochrome et développée pour les applications automobiles.

3.3.2 Analyse des performances

Dans un premier temps, nous avons comparé les performances de nos méthodes avec les résultats publiés dans [Houben, 2011]. L’auteur utilise la base de données *Stéréopolis* pour mesurer le TDPI des méthodes DPR, TSB, STVUT pour la *Détection* de panneaux triangulaires et des méthodes TrSR, TCB, et STVUE pour la *Détection* de panneaux circulaires. Il est utile de préciser que [Houben, 2011] n’aborde pas le problème des FP et ne présente aucune évaluation dans ce sens. L’auteur exploite l’information couleur en intégrant des méthodes de prés-traitement : LCG, EO, et GMT (cf. Section 2.2.2). La table 3.1 compare le TDPI des méthodes DPR, TSB et STVUT présentées dans [Houben, 2011] avec l’algorithme RSLD. Ce dernier permet un TDPI de 81% avec 0.15 de FPPI. Les résultats montrent que notre proposition basée uniquement sur l’information gradient (contour) a des performances équivalentes à la meilleure méthode STVUT présentée dans [Houben, 2011]. Par conséquent, l’intégration de l’information couleur à l’algorithme RSLD peut accroître ses performances. Les FP sont générés principalement par la présence d’objets reflétant la forme

TABLE 3.1 – Comparaison des taux de détection des panneaux triangulaires sur la base *Stéréopolis*

Méthode	Taux de Détection Par Image (TDPI) (%)
DPR (LCG)	52
TSB (LCG)	74
STVUT (LCG)	78
DPR (EO)	48
TSB (EO)	81
STVUT (EO)	81
RSLD	81 (FPPI = 0.15)

FIGURE 3.17 – Exemples de Faux Positifs (FP) lors de la *Détection* de panneaux triangulaires et circulaires sur la base de données *Stéréopolis*.

triangulaire. La figure 3.17 illustre un exemple de FP. Nous montrons dans le chapitre suivant que ces derniers, qui apparaissent et disparaissent de manière aléatoire, peuvent être éliminés par l'intégration de l'information spatio-temporelle.

La table 3.2 compare notre méthode de *Détection* de panneaux circulaires avec les méthodes TrSR, TCB, et STVUE [Houben, 2011]. La méthode STVUE appliquée aux contours sélectionnés par la méthode LCG (près-traitement) procure le meilleur TDPI (91%), suivie de la TrSR et de la TCB. Le fait que la base *Stéréopolis* ne contienne que des panneaux circulaires dont le fond est de couleur rouge ou bleu permet de mieux sélectionner les contours pertinents, ce qui explique ces taux de détection. En revanche, notre proposition qui n'exploite pas cette information couleur ne procure qu'un taux de 38% avec 0.3 de FPPI. Ce taux est dû aussi à la définition des positions candidates qui n'est pas performante sur cette base (*Stéréopolis*). En effet, 20% des panneaux circulaires de cette dernière possèdent un pictogramme à partir duquel il est impossible de détecter des coins localisés au centre du panneau. La figure 3.18 présente deux exemples de ce type de panneaux où aucun coin n'est détecté au centre. Un exemple de FP est représenté dans la figure 3.17.

TABLE 3.2 – Comparaison des taux de détection des panneaux circulaires sur la base *Stéréopolis*

Méthode	Taux de Détection Par Image (TDPI) (%)
TrSR (LCG)	90
TCB (LCG)	86
STVUE (LCG)	91
TCB (GMT)	44
STVUE (GMT)	62
Méthode proposée	38 (FPPI = 0.3)

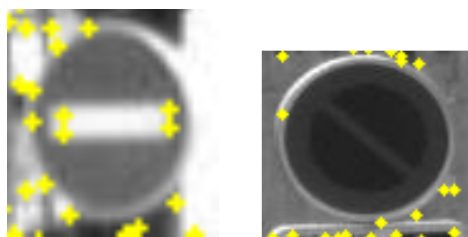


FIGURE 3.18 – Exemples de panneaux circulaires sans coin détecté au centre.

TABLE 3.3 – Séquences de test capturées sur le réseau routier français.

	Vidéo 1	Vidéo 2
Nombre de images	12 500	6 250
Distance (km)	9.3	5.2
Autoroute (%)	26.3	7.7
Nationale (%)	0.0	27.3
Départementale (%)	73.7	64.9
Ville (%)	20.2	16.2
Nombre de panneaux circulaires	31	18
Nombre de panneaux triangulaires	11	13
Vitesse Min du véhicule (km/h)	0	50
Vitesse Max du véhicule (km/h)	110	110
< 90 km/h (%)	20.9	45.7
Changement de voie (sorties)	8(3)	4(3)
Autre	1 feu tricolore	

Pour assurer une évaluation plus proche du terrain, nous avons testé nos deux méthodes sur des séquences vidéos plus longues (15 *min*), plus variées, et complètes. Dans ce cas, chaque image est traitée par le système, ce qui nous permet de mieux mesurer le FPPI qui risque d'influencer les performances globales du système. Lors de cette évaluation nous avons utilisé la base de données privée du laboratoire MIPS. Cette dernière est constituée de deux séquences que nous avons annoté à la main de sorte que, dans chaque image, les positions des panneaux ayant une résolution entre 20×20 et 60×60 pixels sont définies. Les deux vidéos de la base reflètent différentes situations de trafic (cf. Table 3.3). En effet, la première séquence (12500 images) concerne principalement

les routes départementales et autoroutes avec plusieurs changements de voie de roulement (8 dont 3 sorties d'autoroute), ainsi une large variation de la vitesse du véhicule est observée (de 0 à 110 km.h^{-1} avec un faible ratio de vitesses en-dessous de 90 km.h^{-1}). Dans cette première séquence, 42 panneaux sont présents et visibles sur 1450 images. La deuxième séquence est plus variée, plusieurs types de routes sont représentés avec un large ratio de vitesses en-dessous de 90 km.h^{-1} . Cette variation est confirmée par le ratio du nombre de panneaux présents par rapport à la longueur de la vidéo (31 panneaux sur 6250 images). En plus des mesures standards TDPI et FPPI, nous nous sommes basés sur le TDPP pour évaluer notre approche. En effet, nous estimons qu'un système n'est pas obligé de détecter un panneau dans toutes les images de la séquence vidéo pour être performant. De plus un panneau détecté et reconnu ne doit être signalé qu'une seule fois au conducteur. Dans le cas contraire ce dernier risque d'être distrait par sa recherche du deuxième panneau alors qu'il en a qu'un seul [Møgelmoose *et al.*, 2012].

La table 3.4 résume les résultats de l'évaluation sur les vidéos 1 et 2. Les méthodes proposées assurent en moyenne la *Détection* de 93% des panneaux (68 sur les 73 présents dans les 2 séquences) avec un TDPI de 78% et un FPPI de 0.23. Il est important de noter que notre solution permet la *Détection* de deux types de panneaux : circulaires et triangulaires dont les panneaux de cédez le passage (triangulaires inversés). L'évaluation démontre que la *Détection* de panneaux circulaires est légèrement meilleure que celle des panneaux triangulaires contrairement à l'évaluation précédente sur la base *Stéréopolis*. En effet, tous les panneaux circulaires des vidéos 1 et 2 contiennent un pictogramme noire sur fond blanc ce qui facilite la détection de coins. De plus, la majorité des panneaux sont des panneaux de limitation de vitesse où les pictogrammes reflètent des coins proches du centre des panneaux. Sur la séquence 1, la méthode proposée permet de détecter 27 panneaux circulaires des 31 (TDPP= 87%) présents avec un très bon TDPI de 87.95% et seulement 0.28 de FPPI. Nous observons la même tendance sur la séquence 2 avec un TDPI de 92.81% et 0.27 de FPPI. Concernant la *Détection* de panneaux triangulaires les résultats sont satisfaisants aussi avec 23 panneaux détectés sur les 24 présents sur les deux vidéos soit un TDPP de 96%. Bien que l'algorithme RSLD procure un TDPI de 64.15% seulement sur la séquence 1 il ne manque qu'un seul panneau. Même constat pour la séquence 2 avec 69.54% de TDPI. En effet, la distance entre le panneau et le véhicule influe sur la classification des coins liés aux panneaux triangulaires, ce qui peut nuire considérablement à la performance de l'algorithme RSLD. Plus le panneau est proche du véhicule plus les sommets supérieurs (et les pointes inférieures) sont mieux reconnus et le TDPI augmente. Une évaluation a été menée où la plage de résolution des panneaux pris en considération était de 30×30 (au lieu de 20×20) à 60×60 pixels. Les résultats démontrent que dans ces conditions le TDPI passe de 64.15% à 68.02% pour la séquence 1 et de 69.54% à 73.78% pour la séquence 2 mais sans influencer le TDPP. Les figures 3.19, 3.20, et 3.21 illustrent des exemples de détections à différents moments de la journée.

Les panneaux non détectés sont principalement placés au niveau des sorties d'autoroutes où leur distance par rapport à la caméra rend la tâche de *Détection* assez complexe. Le fait de les manquer

TABLE 3.4 – Performances pour les séquences 1 et 2

Panneaux		Séquence 1	Séquence 2
Circulaires (ROIs > 20 × 20)	<i>TDPP</i> (%)	87	95
	<i>TDPI</i> (%)	87.95	92.81
	<i>FPPI</i>	0.28	0.27
Triangulaires (ROIs > 20 × 20)	<i>TDPP</i> (%)	91	100
	<i>TDPI</i> (%)	64.15	69.54
	<i>FPPI</i>	0.26	0.11
Triangulaires (ROIs > 30 × 30)	<i>TDPP</i> (%)	91	100
	<i>TDPI</i> (%)	68.02	73.78
	<i>FPPI</i>	0.20	0.11

FIGURE 3.19 – Exemples de *Détection* de panneaux de danger.

ne constitue pas un réel danger puisqu'ils concernent que les véhicules quittant l'autoroute. La figure 3.22 illustre un exemple de fausses détections générées durant les tests. On constate clairement que la scène routière peut regrouper des objets ayant une forme similaire à celle des panneaux, ce qui augmente le nombre de FP.

3.3.3 Robustesse des méthodes proposées aux dégradations d'images

Pour une meilleure évaluation, nous avons mesuré la sensibilité de nos méthodes à trois types de dégradations : présence de bruit additif, le flou des images, et la rotation des panneaux. Lors de



FIGURE 3.20 – Exemples de *Détection* de panneaux cédez le passage.



FIGURE 3.21 – Exemples de *Détection* de panneaux d'interdiction.

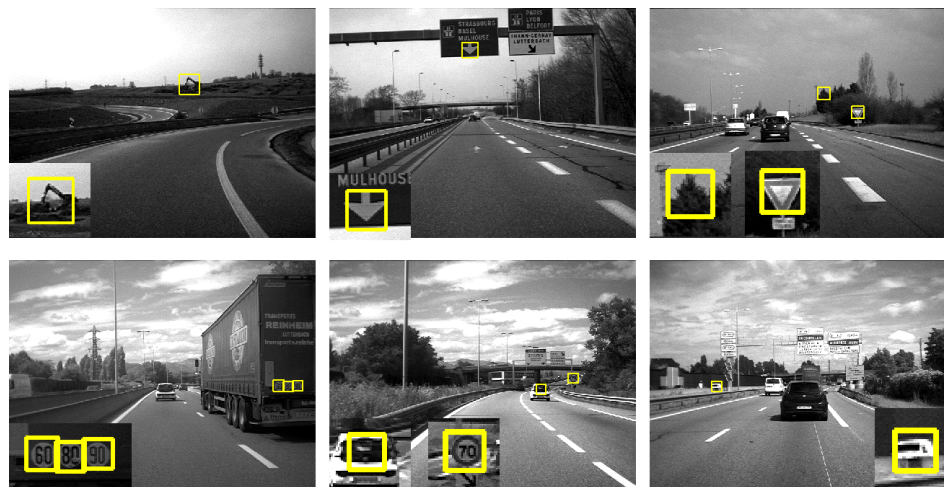


FIGURE 3.22 – Exemples de Faux Positifs (FP) lors de la *Détection* de panneaux triangulaires et circulaires sur la base de données du MIPS.

cette évaluation de sensibilité nous avons utilisé la base de données de *Grigorescu* et nous avons reproduit les mêmes conditions d'expériences rapportées dans [Belaroussi et Tarel, 2010].

Bruit

Dans un premier temps, nous avons mesuré la tolérance au bruit additif de l'algorithme RSLD et nous l'avons comparés avec celle de l'algorithme TSB de [Belaroussi et Tarel, 2010]. L'étude concernait différents niveaux de bruit Gaussien. A chaque pixel de chaque image de la base de données de *Grigorescu* est ajouté un bruit blanc Gaussien d'écart type σ . La figure 3.23 (a) représente le TDPI en fonction de σ variant de 0 à 40. Notre algorithme RSLD procure les meilleurs TDPI avec un décroissement plus lent que celui de TSB lorsque l'écart type augmente, par exemple pour $\sigma = 5$; RSLD permet de détecter 100% des panneaux triangulaires contre seulement 90% pour l'algorithme TSB. Avec $\sigma = 15$, notre méthode permet un TDPI de 90%. Par conséquent, on peut conclure que l'algorithme RSLD est plus robuste au bruit que l'algorithme TSB. En effet, RSLD est basé sur l'algorithme RANSAC qui est performant dans un environnement bruité et l'utilisation de l'image codée permet l'extraction des pixels liés seulement aux côtés diagonaux de la forme triangulaire. Sans ajout de bruit additif, l'algorithme TSB et notre RSLD procurent respectivement 95% et 100% de TDPI avec 2 et 1 FP.

La même évaluation a été réalisée pour mesurer la robustesse de notre méthode de *Détection* de panneaux circulaires. La figure 3.23 (b) compare la performance de notre proposition avec celle de TCB en fonction de différent niveau de bruit additif. Notre méthode est globalement plus robuste que la TCB bien que cette dernière procure un meilleur TDPI lorsque $\sigma < 10$. Par exemple, avec $\sigma = 5$ la TCB détecte 90% des panneaux contre 78% pour notre méthode. Au delà d'un certain niveau de bruit ($\sigma = 10$) notre méthode procure de meilleurs TDPI que la TCB. En effet, la TCB utilise un processus de vote basé sur le critère des symétries centrale et axiale (précisions angulaires δ et β , voir la description de la méthode dans 2.2.2), ces dernières ne sont plus vérifiées si un bruit

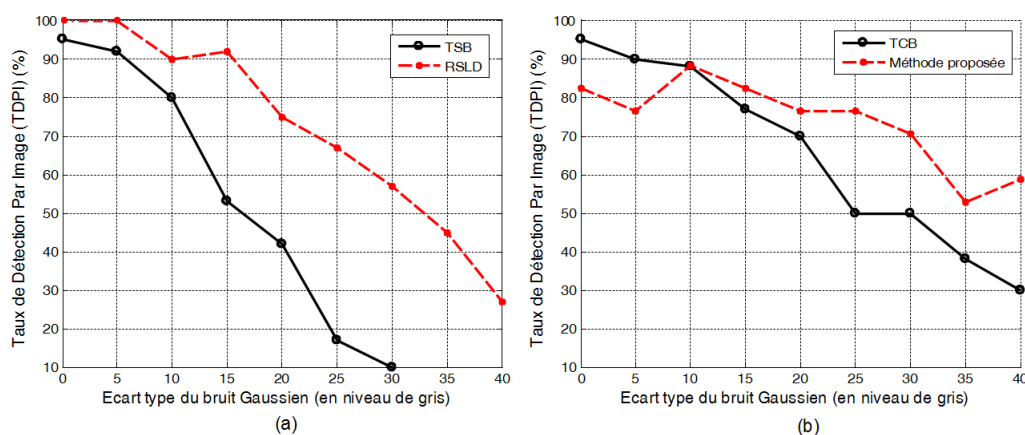


FIGURE 3.23 – Taux de détection de panneaux ((a) triangulaires et (b) circulaires) sur des images dégradées par l'ajout d'un bruit Gaussien.

aléatoire est ajouté aux périmètres des cercles. En revanche, la méthode proposée est basée sur la classification des pixels qui exploite elle aussi les orientations des gradients mais où les plages des directions gradient sont assez large et donc plus robuste au bruit (cf. Figure 3.3).

Flou

Le deuxième type de dégradation étudiée est le flou. Les images de la base sont convoluées avec un filtre Gaussien ayant un écart type variant de 0 à 5. Ces images floues ou défocalisées dégradent les performances des méthodes proposées (cf. Figure 3.24). En effet, notre approche est basée sur l'information gradient et le détecteur de Harris qui permet la détection des positions candidates. En présence de flou, les coins deviennent difficile à détecter et donc les positions candidates ne sont pas extraites correctement. Contrairement à l'algorithme RSLD l'algorithme TSB est moins sensible au flou puisqu'il est basé sur un détecteur de symétrie. La figure 3.24 (a) compare les TDPI de RSLD et TSB en fonction de l'écart type du filtre Gaussien. Si l'écart type est inférieure à un pixel, les TDPI sont proches de 95% pour les deux algorithmes, par contre pour des valeurs supérieures l'algorithme RSLD est moins performant que l'algorithme TSB. Le TDPI de l'algorithme TSB décroît considérablement pour un écart type supérieur à 2. Nous constatons le même comportement pour notre deuxième proposition (cf. Figure 3.24 (b)). Sur les images floues il est difficile voir impossible de définir correctement les positions candidates, ce qui nuit considérablement à la *Détection*. Néanmoins, nous estimons qu'un système TSR doit utilisé une lentille focalisée à l'infini avec une distance focale fixe, ce qui rend très rare les images floues.

Rotation

Finalement, la robustesse à la rotation a été évaluée. L'objectif est de reproduire l'effet de rotation qui peut être observé dans le cas où le panneau est mal orienté suite à une dégradation. Notre approche se base sur un codage des contours selon l'orientation du gradient, l'orientation du panneau peut influencer les performances. L'algorithme RSLD est plus sensible à ce type de

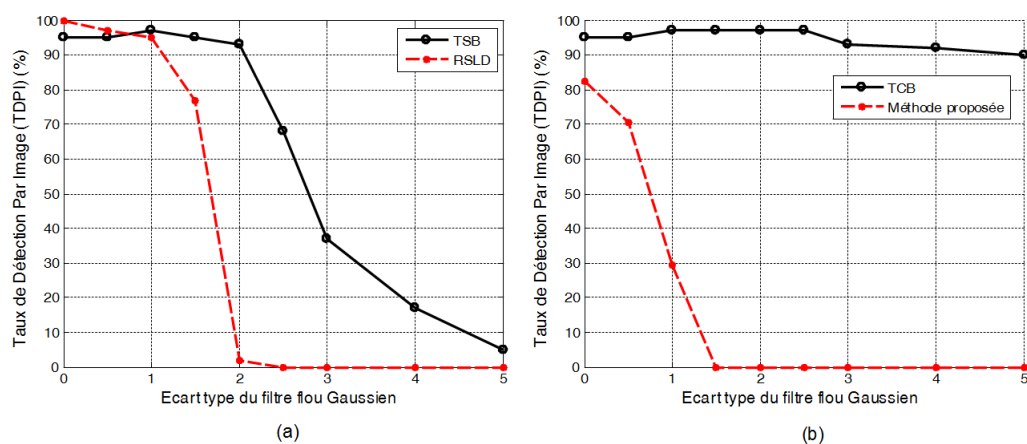


FIGURE 3.24 – Taux de détection de panneaux ((a) triangulaires et (b) circulaires) sur des images floutées à l'aide d'un filtre Gaussien.

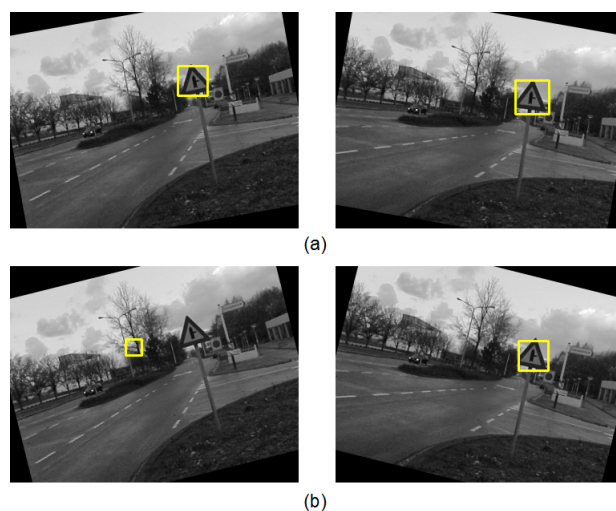


FIGURE 3.25 – Exemples de *Détection* de panneaux triangulaires sur les images de *Grigorescu* après application d'une rotation. (a) L'image subie une rotation de +10 degrés (*image de gauche*) et -10 degrés (*image de droite*). (b) L'image subie une rotation de +14 degrés (*image de gauche*) et -14 degrés (*image de gauche*).

dégradation puisqu'il est basé sur la détection des côtés diagonaux. Ces derniers sont représentés par des contours dont le codage sera modifié suite à l'effet d'orientation. Si les points contours appropriés ne sont pas extraits correctement pour l'algorithme RANSAC alors la détection échoue. La figure 3.25 illustre les résultats de la *Détection* sur des images de la base de *Grigorescu* après l'application d'une rotation. Notre conclusion est que les performances de RSLD reste inchangées si l'angle de la rotation ne dépasse pas les 10 degrés. Le constat est différent pour la *Détection* de panneaux circulaires. En effet, la forme de ces derniers reste inchangée si la rotation n'est que horizontale ou dans le plan image sinon le panneau perçu change de forme et il ressemblera à une ellipse. La figure 3.26 illustre des exemples de détections après application d'une rotation de 10 et de 20 degrés.

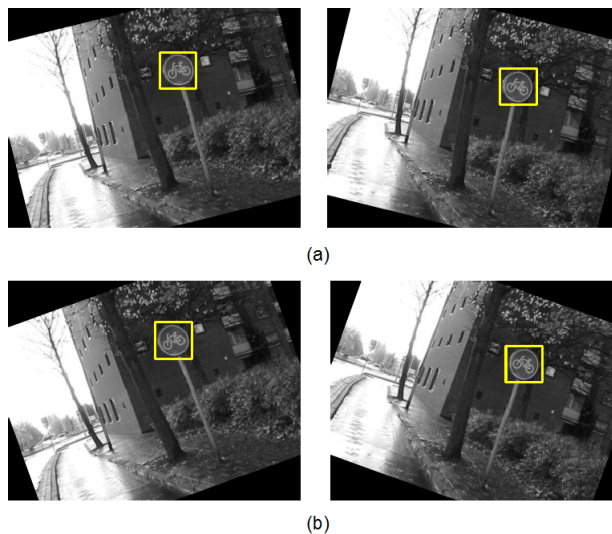


FIGURE 3.26 – Exemples de *Détection* de panneaux circulaires sur les images de *Grigorescu* après application d’une rotation. (a) L’image subie une rotation de +14 degrés (*image de gauche*) et -14 degrés (*image de droite*). (b) L’image subie une rotation de +20 degrés (*image de gauche*) et -20 degrés (*image de gauche*).

3.4 Conclusion

L’étude des algorithmes de *Détection* présentée précédemment (cf. Chapitre 2) a permis la proposition d’une nouvelle approche originale pour la *Détection* de panneaux triangulaires et circulaires. Cette approche est basée sur la détection de coins et le codage de contours. Les coins sont détectés par le détecteur de Harris et définissent les positions candidates des panneaux dans l’image. Quant au codage de contours, il est basé sur une classification des pixels en fonction de leur orientation du gradient. Ensuite autour de chaque position candidate, un panneau est détecté si une distribution particulière des contours est présente. La méthode proposée a été évaluée avec des méthodes récemment publiées et les résultats sont similaires aux résultats des meilleures méthodes. Nous estimons qu’il est possible d’améliorer ces performances par l’intégration de l’information temporelle à la *Détection*. Pour cela nous proposons l’ajout d’une phase de *Suivi temporel* après l’étape de *Détection*.

Chapitre 4

Suivi temporel de panneaux

4.1 Introduction

Développé initialement dans le domaine de la robotique, principalement pour des applications militaires, le *Suivi temporel d'objets* (“*Target Tracking*”) est devenu un sujet de recherche très dynamique. En particulier, la communauté de la vision par ordinateur exploite cette technique pour par exemple des applications dédiées aux véhicules intelligents comme le suivi des voies de roulement [Mèche, 2013] ou de véhicules [Sivaraman et Trivedi, 2013]. Plusieurs applications de vision se basent sur le *Suivi temporel* afin d’améliorer leurs performances, e.g. les applications de vidéosurveillance, les interfaces homme-machine, ou les applications d’analyse vidéo. En effet, le *Suivi temporel* permet de mieux analyser le mouvement des objets en faisant face par exemple au problème d’occultation temporaire de ces derniers dans une séquence d’images.

Ce processus de suivi nécessite deux étapes : le filtrage et l’association de données. Le filtrage a pour but d’estimer les paramètres dynamiques et/ou cinématiques des objets, e.g. la position, la vitesse, etc. Quant à l’association de données, elle analyse l’aspect dynamique de la scène, e.g. gestion des apparitions et des disparitions d’objets, grâce à la mise en correspondance des objets connus et stockés dans une base de données locale avec les objets perçus par un ou plusieurs capteurs. Cette technique a été étendue aux véhicules intelligents afin d’améliorer leurs capacités de perception. Parmi les travaux récents, il y a ceux de [Meuter *et al.*, 2011], [Mercier *et al.*, 2011], et [El Zoghby *et al.*, 2013]. Le principe est d’utiliser la redondance et la complémentarité des données fournies par les capteurs et/ou les applications de perception en fusionnant ces données.

Pour les systèmes TSR, la caméra embarquée fournit des images de la scène routière où une redondance de la signalisation peut être observée. En effet, un panneau, après son apparition, reste visible sur plusieurs images avant de disparaître du champ de vision de la caméra et du conducteur. Dans ce contexte, il est intéressant d’exploiter cette redondance temporelle d’informations, par

l'intégration d'une phase de *Suivi temporel* après la *Détection* et/ou après la *Reconnaissance*, afin d'améliorer les taux de détection et/ou de classification tout en réduisant le nombre de FP et/ou le taux d'erreur de classification.

Ce chapitre présente, dans la première partie, le principe de notre approche suivi-*après*-détection basée sur la fusion de données. L'idée est de fusionner les résultats de la *Détection* afin d'améliorer les performances de notre détecteur (cf. Chapitre 3) et cela en réduisant le nombre de FP tout en garantissant un taux de détection élevé. Ce principe a été modélisé dans un algorithme dit "*algorithme de Suivi Multi-ROIs*" (SMR) basé sur une association crédale. Puisque les résultats de la *Détection* sont imparfaits, nous avons opté pour l'utilisation de la théorie des croyances qui offre les mécanismes nécessaires pour modéliser au mieux les différentes imperfections liées à une telle application.

La deuxième partie de ce chapitre est consacrée à la présentation de la théorie des croyances. Dans un premier temps, nous décrivons l'utilisation des croyances, en particulier le Modèle des Croyances Transférables (MCT), pour l'association de données en présentant les différentes étapes nécessaires à la prise de décision. Par la suite, un état de l'art des algorithmes de prise de décision développés pour la mise en correspondance d'objets est présenté.

Finalement, dans la troisième partie, nous évaluons l'impact de l'algorithme de Suivi Multi-ROIs proposé sur les performances de la *Détection*.

4.2 Association crédale multi-objets pour le suivi de ROIs

Présenté dans le chapitre 2, l'état de l'art du *Suivi temporel* a mis en évidence le fait que cette étape est peu intégrée dans les systèmes TSR. En effet, peu de travaux mettent l'accent sur son intégration malgré la robustesse qu'elle peut apporter aux systèmes face aux erreurs de *Détection* et de *Reconnaissance* [Møgelmoose *et al.*, 2012]. Souvent exploité après la *Reconnaissance*, le *Suivi temporel* a pour principal intérêt la réduction du temps de traitement bien que sa mise en oeuvre permet également d'améliorer les performances de la *Détection*. C'est pourquoi, nous proposons une nouvelle approche de suivi-*après*-détection pour traiter les imperfections de la *Détection* : détections multiples pour le même panneau, échec de détection dû aux occultations temporaires, et les FP.

En général, ces FP (ou hallucinations) sont caractérisées par un mouvement aléatoire et donc imprévisible dans les images (contrairement aux trajectoires des panneaux dans le référentiel image) ce qui permet leur identification. Non filtrés, leur présence engendre des calculs inutiles durant la *Reconnaissance* et risque d'augmenter le taux d'erreur de classification d'où l'intérêt de les supprimer en amont. C'est pourquoi, nous avons intégré le concept de *Suivi temporel* à notre détecteur de panneaux (cf. Section 3.2) en développant une nouvelle méthode de suivi basée sur la

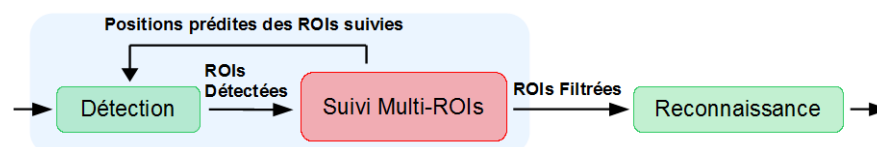


FIGURE 4.1 – Diagramme du flux de données entre le SMR et les étapes de *Détection* et *Reconnaissance*.

théorie des croyances. Le but est d'assurer, dans un contexte mono-capteur basé sur une caméra, le pistage de ROIs obtenues par le détecteur afin de supprimer les incohérences de détection et d'être robuste aux occultations en garantissant un taux de détection maximal. La généralité de cette méthode permettra son application au pistage d'objets statiques (panneaux, amers, etc.) ou dynamiques (véhicules, piétons, etc.) dans un environnement bruité.

Cette section présente le principe de l'algorithme de Suivi Multi-ROIs composé de trois étapes : *le Filtrage*, *l'Association de données*, et *l'Analyse des pistes*. La première étape, le *Filtrage*, est basée sur le FK dont les modèles dynamique et d'observation sont décrits par la suite. Pour la deuxième étape, une association évidentielle pour la mise en correspondance entre les objets connus (ROIs suivies) et les objets détectés (ROIs détectées) est proposée. La dernière étape de l'algorithme reconnaît les FP en analysant les pistes des ROIs.

4.2.1 Principe du Suivi de Multi-ROIs

Contrairement aux panneaux, les fausses détections sont souvent caractérisées par des apparitions aléatoires et une durée de vie courte dans la séquence d'images. Par conséquent, pour réduire leur nombre, la *Détection* d'un panneau est confirmée seulement si ce dernier est suivi sur plusieurs images consécutives. Cette méthode s'inspire de celle de Meuter [Meuter *et al.*, 2008] où on ne confirme la présence d'un panneau qu'après plusieurs détections. Un autre avantage du suivi est son habilité à résoudre le problème d'occultations temporaires par la prédiction des positions et résolutions, dans le plan image, des panneaux à travers le temps.

Dans ce contexte, notre système est composé d'une *Détection* couplée à l'algorithme SMR (cf. Figure 4.1). La *Détection* procure un ensemble de panneaux candidats (i.e. ROIs détectées) pouvant contenir des FP. Les ROIs détectées constituent les objets perçus que le SMR doit suivre temporellement et spatialement dans le plan image. Ce suivi permet de gérer les apparitions et disparitions d'objets (panneaux) à travers le temps. Quant à l'algorithme SMR, il renvoie au détecteur, avec rétroaction, les positions prédites des ROIs suivies (i.e. objets connus) dans les prochaines images permettant ainsi de localiser plus facilement les panneaux. Les ROIs filtrées, objets confirmés en tant que panneau, sont transmises à l'étape de *Reconnaissance*.

Le SMR est composé de trois fonctions : *Filtrage*, *Association de données*, et *Analyse des pistes* (cf. Figure 4.2). Le *Filtrage* assure la prédiction des futures positions des ROIs suivies dans les images.

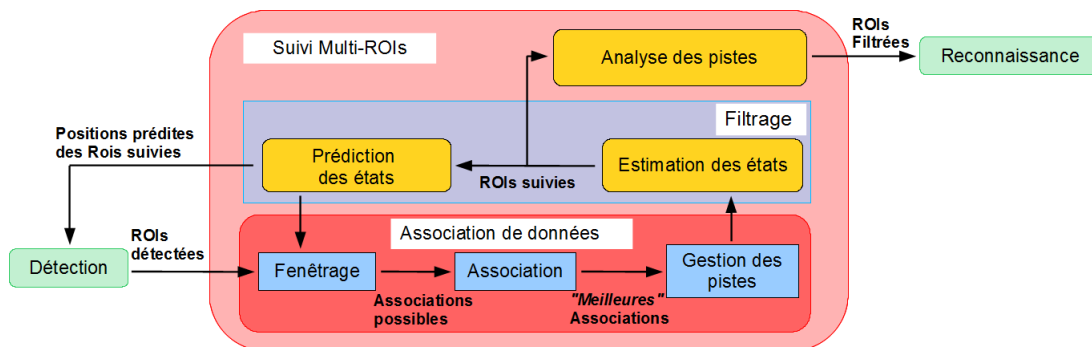


FIGURE 4.2 – Diagramme du SMR.

Cette première étape estime aussi les positions courantes des ROIs en fonction des résultats de l'étape d'*Association de données*. Cette dernière met en correspondance chaque mesure générée par le détecteur (i.e. ROI détectée) à une piste (i.e. ROI suivie) tout en gérant les apparitions et les disparitions des pistes. La dernière étape, *Analyse des pistes*, exploite le contexte temporel pour valider/rejeter les ROIs transmises à la *Reconnaissance*. À partir de chaque image, le détecteur extrait l'ensemble des panneaux candidats (i.e. ROIs détectées) représentant les n cibles pour le SMR. Les ROIs suivies, panneaux connus, constituent les m pistes qui seront associées aux n cibles.

4.2.2 Filtrage

Très utilisé dans les solutions de localisation et navigation, le filtrage statistique est implémenté aussi dans les applications de suivi d'objets. Pour l'algorithme SMR, le *Filtrage* permet d'estimer les paramètres dynamiques et cinématiques des ROIs détectées et de les suivre dans la séquence d'images. Ce processus est réalisé à l'aide du FK (cf. Section 2.4.1.1). La prédiction, première étape du filtre, exploite un modèle dynamique des ROIs afin de prédire leur position dans les prochaines images. Ensuite l'étape d'estimation utilise les mesures observées pour corriger l'état courant des ROIs suivies (i.e. les pistes). Cette section s'attache à décrire les modèles dynamique et d'observation retenus dans ce travail ainsi que les raisons de ces choix.

4.2.2.1 Modèle dynamique

Le modèle dynamique le plus répandu dans les applications de suivi d'objets statiques et dynamiques est le modèle à vitesse quasi-constante. La plupart des systèmes TSR intégrant une phase de suivi et cités dans [Møgelmoose *et al.*, 2012] utilisent ce modèle [Garcia-Garrido *et al.*, 2011], [Prisacariu *et al.*, 2010], [González *et al.*, 2011], et [Ruta *et al.*, 2010]. Le choix de ce modèle est justifié par le contexte des applications TSR. En effet, le mouvement des objets dans une séquence vidéo dépend fortement du mouvement de la caméra et de leur propre déplacement. Dans les applications TSR, le seul mouvement pris en considération est celui de la caméra puisque les objets (panneaux) sont statiques. Compte tenu de la fréquence de *Détection* élevée du système (25 Hz), le

mouvement du véhicule peut être considéré à vitesse quasi-constante, avec des variations de vitesse négligeables entre deux images. C'est pourquoi, nous avons opté nous aussi pour un modèle dynamique à vitesse quasi-constante pour décrire le mouvement des ROIs dans le plan image. Ce modèle a été démontré globalement satisfaisant pour les applications utilisant une caméra embarquée sur véhicule [Cifuentes *et al.*, 2012]. Cependant, le modèle montre ses limites face aux effets de perspectives. Dans ce cas, même si le véhicule se déplace à une vitesse constante, le mouvement des panneaux dans les images n'est plus constant. Néanmoins, grâce à la fréquence d'échantillonnage élevée, notre système n'est pas très affectés par cette distorsion.

À partir de ce constat, nous définissons une piste, à l'instant k , par le vecteur d'état suivant : $x_j(k) = [x, y, s, v_x, v_y, v_s]^t \in \mathcal{X}$ avec $j = 1, 2, \dots, m$ où $[x, y, s]$ représente la position et la résolution de la ROI suivie dans l'image k , et $[v_x, v_y, v_s]$ définit ses vitesses relatives et sa variation de résolution entre deux images successives (k et $k - 1$) [Weng *et al.*, 2006].

Le modèle dynamique utilisé est présenté par l'équation (4.1). Le déplacement des ROIs $[v_x(k), v_y(k)]$, au niveau de l'image k , est défini par un processus Gaussien centré sur $[v_x(k - 1), v_y(k - 1)]$ avec un écart type (σ_x, σ_y) . Pour des raisons de simplification, les effets de perspective sont ignorés et le changement d'échelle des panneaux est considéré quasi constant. $v_s(k)$ est défini Gaussien avec une moyenne $v_s(k - 1)$ et un écart type σ_s .

$$\begin{cases} x(k) = x(k - 1) + v_x(k) \\ y(k) = y(k - 1) + v_y(k) \\ s(k) = s(k - 1) + v_s(k) \end{cases} \quad (4.1)$$

avec :

$$\begin{cases} v_x(k) \sim \text{Norm}(v_x(k - 1), \sigma_x^2) \\ v_y(k) \sim \text{Norm}(v_y(k - 1), \sigma_y^2) \\ v_s(k) \sim \text{Norm}(v_s(k - 1), \sigma_s^2) \end{cases}$$

À partir du modèle d'évolution (cf. (4.1)) et l'équation d'état du filtre (cf. (2.21)) on obtient la matrice dynamique suivante :

$$F = \begin{pmatrix} I_{3 \times 3} & T_e \cdot I_{3 \times 3} \\ 0_{3 \times 3} & I_{3 \times 3} \end{pmatrix} \quad (4.2)$$

avec T_e le temps d'échantillonnage, il vaut 1 si toutes les images de la séquence sont traitées.

Quant à la matrice de covariance du bruit du système, elle est définie par :

$$Q = \begin{pmatrix} 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & \sigma \end{pmatrix} \quad (4.3)$$

avec $\sigma = \text{diag}(\sigma_x^2, \sigma_y^2, \sigma_s^2)$.

4.2.2.2 Modèle d'observation

Les observations, i.e. ROIs détectées à l'instant k , sont définies par les vecteurs $z_i(k) = [x, y, s]^t \in \mathcal{Z}$ avec $i = 1, 2, \dots, n$ où (x, y) représente la position de la ROI détectée dans le plan image, et s sa résolution. Pour des raisons de simplification, les cibles $z_i(k)$ et les pistes $x_j(k)$ sont considérées linéairement dépendantes. Par conséquent, la matrice d'observation du système linéaire (cf. (2.21)) est définie par :

$$H = \begin{pmatrix} I_{3 \times 3} & 0_{3 \times 3} \end{pmatrix} \quad (4.4)$$

Quant à la matrice de covariance du bruit de mesure, elle est définie par :

$$R = \begin{pmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_y & 0 \\ 0 & 0 & \sigma_s \end{pmatrix} \quad (4.5)$$

4.2.3 Association par fusion de données

Une fois l'étape de *Détection* effectuée il convient de mettre en correspondance les mesures observées, ROIs détectées, avec les pistes existantes, ROIs suivies. L'association de données est une étape préalable à la mise à jours des filtres d'estimation. L'objectif est d'associer chaque mesure à une piste, ce qui permet de corriger son état dynamique. Plusieurs situations sont possibles :

- mettre à jour les pistes en fonction des associations déterminées,
- créer de nouvelles pistes à partir des mesures non-associées,
- supprimer les pistes n'ayant pas été associées sur un nombre prédéfini d'images et donc considérées comme des FP.

Parmi les approches d'association présentées dans le chapitre 2, nous avons opté pour l'association crédale basée sur la théorie des croyances. Plusieurs travaux montrent l'intérêt de cette théorie pour la fusion de capteurs et en particulier pour le problème d'association de données [Dallil *et al.*, 2013], [Daniel et Lauffenburger, 2013], et [El Zoghby *et al.*, 2013]. Dans son ouvrage de référence, [Blackman, 1986] soutient que, en tant qu'une généralisation de l'inférence bayésienne et de la règle bayésienne de combinaison, le raisonnement évidentiel a été construit pour faire face aux limites du cadre probabiliste : inapproprié à l'ignorance, incertitude et incomplétude de la modélisation, nécessité de définir des hypothèses sur les erreurs de mesures, etc. Par ailleurs, certains travaux ont montré les performances de l'association crédale par rapport à des approches plus classiques telles que les GNN et JPDA [Dallil *et al.*, 2013].

4.2.4 Analyse des pistes

L'étape d'analyse des pistes de l'algorithme SMR consiste à évaluer leur évolution afin d'identifier celles liées aux panneaux réellement existant dans l'image de celles engendrées par les fausses ROIs. Chaque piste x_j est évaluée à partir de son indice de confiance $C(j)$ (cf. (4.6)). Cet indice est défini par le ratio entre le compteur de mise à jour $c_{maj}(j)$ et la longueur de la piste $longueur(j)$. $c_{maj}(j)$ représente le nombre de fois où la piste x_j a été associée à une mesure, autrement dit c'est le nombre d'images où la détection de la ROI est confirmée. $longueur(j)$ est la durée de vie d'une piste, représentée par le nombre d'images où la ROI suivie x_j est présente. Les ROIs suivies ayant des indices $C(j)$ inférieur à un seuil prédéfini $T_{confiance}$ sont considérées comme de fausses détections. Les autres ROIs ($C(j) > T_{confiance}$) représentent les ROIs filtrées et sont transmises à l'étape de *Reconnaissance*.

$$C(j) = \frac{c_{maj}(j)}{longueur(j)} \quad (4.6)$$

4.3 Théorie des croyances

À partir d'une perception imparfaite de l'environnement, le suivi d'objets (panneaux, véhicules, etc.) n'est pas une tâche simple. Dans la pratique, les capteurs et/ou les algorithmes de perception, embarqués sur le véhicule, fournissent des informations imprécises et incertaines ce qui nuit à l'interprétation des événements de la scène. Sans tenir compte de ces imperfections, l'association de données risque d'aboutir à une mauvaise interprétation de la dynamique de la scène. Par conséquent, il est nécessaire de modéliser au mieux ces imperfections afin d'aboutir à des décisions pertinentes [Mourllion, 2006].

4.3.1 Introduction

Initialement issue des travaux de [Dempster, 1967], la théorie des croyances modélise des connaissances imparfaites par une distribution de masses de croyance représentant la confiance accordée aux solutions d'un problème donnée. Plusieurs modèles d'interprétation de cette théorie ont été proposés : le modèle de Dempster-Shafer [Shafer, 1976], le modèle des "Hints" [Kohlas et Monney, 1995] basés sur les probabilités imprécises, et le MCT de [Smets et Kennes, 1994] basé sur une interprétation non-probabiliste où les connaissances sont caractérisées par un degré de croyance. Chacun de ces modèles s'appuie en outre sur des opérateurs de combinaison (cf. Section 4.3.4) et/ou de décision (cf. Section 4.3.5) propres.

Dans cette section, nous présentons le processus de fusion à l'aide de la théorie des croyances. En premier lieu, le contexte étudié est modélisé en identifiant les solutions élémentaires possibles (hypothèses singletons). Ensuite, des masses de croyance représentant la confiance accordée aux

hypothèses sont estimées tout en tenant compte de l'imprécision et l'incertitude des informations. Une combinaison de ces masses est effectuée afin d'exploiter la redondance et la complémentarité des informations. Cette combinaison génère aussi des masses de croyance sur les unions d'hypothèses en plus des singletons. Finalement, les croyances des unions d'hypothèses sont transférées sur les hypothèses singletons afin de pouvoir prendre une décision à partir des solutions élémentaires définies initialement.

4.3.2 Modélisation du contexte

En théorie des croyances, le contexte étudié est modélisé par un *cadre de discernement* noté Θ et représentant les k hypothèses $(H_i, i = 1, 2, \dots, k)$ susceptibles d'être la solution du problème posé (cf. (4.7)). Dans le modèle de Dempster-Shaper, ce cadre doit être exhaustif [Shafer, 1976], la solution du problème étant forcément une des hypothèses H_i . Θ doit aussi être exclusif, i.e. $H_i \cap H_j = \emptyset, \forall i \neq j$.

$$\Theta = \{\{H_1\}, \{H_2\}, \dots, \{H_k\}\} = \bigcup_{i=1}^N \{H_i\} \quad (4.7)$$

Le *référentiel de définition* noté 2^Θ défini à partir de Θ comprend l'ensemble des 2^k disjonctions (ou unions) possibles des hypothèses H_i :

$$2^\Theta = \{\emptyset, \{H_1\}, \dots, \{H_k\}, \dots, \{H_1, H_2, H_3\}, \dots, \Theta\} \quad (4.8)$$

Ainsi, l'ensemble 2^Θ contient, en plus des hypothèses singletons de Θ , toutes les disjonctions possibles à partir de ces hypothèses. Il traduit toutes les solutions qui peuvent être rencontrées face au problème considéré. L'hypothèse vide \emptyset représente la solution impossible, et l'hypothèse Θ représente l'ignorance, l'union de toutes les hypothèses H_i . Une proposition $A = \{H_1, H_2\}$ représente la disjonction " H_1 ou H_2 " et signifie que l'hypothèse $\{H_1\}$ ou l'hypothèse $\{H_2\}$ peuvent être une solution au problème, traduisant ainsi l'imprécision de l'information. Chaque proposition A de 2^Θ est caractérisée par sa *fonction de masse* ou *masse de croyance élémentaire* (*basic belief mass* (*bbm*)) notée $m^\Theta(A)$. Cette dernière évalue la véracité de la proposition A et définit l'incertitude de l'information. *bbm* est définie dans l'intervalle $[0, 1]$ par :

$$m^\Theta : 2^\Theta \rightarrow [0, 1] \quad (4.9)$$

avec

$$m^\Theta(\emptyset) = 0, \text{ et } \sum_{A \in 2^\Theta} m^\Theta(A) = 1 \quad (4.10)$$

Puisque les k hypothèses de Θ sont supposées exclusives et exhaustives, on dit que le cadre Θ est un *monde fermé*. La solution au problème est nécessairement l'une des hypothèses de Θ et la masse sur \emptyset est nulle (cf. (4.10)).

L'originalité de cette modélisation, face aux autres théories de l'incertain telles l'approche bayésienne ou possibiliste réside dans le fait que la masse unitaire est distribuée non seulement sur les hypothèses simples (singletons) mais également sur les disjonctions de celles-ci.

La répartition des masses obtenues est appelée *jeu de masses*. Les propositions A dont $m^\Theta(A) \neq 0$ sont appelées *éléments focaux* et constituent le *noyau* N_Θ :

$$N_\Theta = \{A \in 2^\Theta / m^\Theta(A) > 0\} \quad (4.11)$$

Quand le noyau N_Θ est composé uniquement d'hypothèses singletons, la notion de masse bbm est assimilable à celle de probabilité et porte le nom de jeu de masses bayésien [Shafer, 1976].

Une proposition complexe A ($|A| > 1$) avec une masse de croyance non nulle ($m^\Theta(A) \neq 0$), traduit une sorte d'*ignorance partielle* parce que la masse n'a pas pu être affectée à un sous-ensemble plus spécifique à A [Mourllion, 2006]. Cette souplesse de la théorie des croyances permet aussi de représenter l'*ignorance totale* $m^\Theta(\Theta) = 1$ et $m^\Theta(A) = 0, \forall A \neq \Theta$, et donc les méconnaissances peuvent être modélisées.

4.3.3 Estimation des croyances

Après la modélisation du contexte, il convient d'évaluer les jeux de masses modélisés pour les combiner et ensuite prendre la décision. Plusieurs techniques d'estimation ont été proposées et leur choix dépend de la modélisation et du problème considéré.

[Shafer, 1976] propose une fonction de masses très simple où toute la masse d'une source Src_j est répartie sur une seule proposition A de 2^Θ tel que $A \neq \Theta$ et $A \neq \emptyset$, ainsi que sur l'ignorance :

$$\begin{cases} m_j^\Theta(A) = s \\ m_j^\Theta(\Theta) = 1 - s \\ m_j^\Theta(B) = 0, \forall B \in 2^\Theta, B \neq A, B \neq \Theta \end{cases} \quad (4.12)$$

où $s \in [0, 1]$. Si $s = 0$ alors $m_j^\Theta(\Theta) = 1$, ce qui représente l'ignorance totale.

Une approche similaire, appelée fonction de croyance complémentaire, est présentée dans [Yager, 1987]. Dans ce cas, la masse de la source Src_j est répartie sur une seule proposition A et sur son

complément (contraire) \bar{A} :

$$\begin{cases} m_j^\Theta(A) = s \\ m_j^\Theta(\bar{A}) = 1 - s \\ m_j^\Theta(B) = 0, \forall B \in 2^\Theta, B \neq A, B \neq \bar{A} \end{cases} \quad (4.13)$$

avec $s \in [0, 1]$.

[Denoeux, 1995] propose une bbm issue de la reconnaissance des formes, où seuls les singletons H_i et l'ignorance Θ forment le noyau. Une classe (H_i) est définie par son centre x_i estimé à partir d'exemples d'apprentissage. Chaque paire (x_i, H_i) est une source d'information pour la classification d'un vecteur donné x . Les fonctions de masses sont définies à partir des distances de x aux centres x_i :

$$\begin{cases} m_j^\Theta(H_i) = \alpha_{i,j} \varphi_i(d(x, x_i)) \\ m_j^\Theta(\Theta) = 1 - \alpha_{i,j} \varphi_i(d(x, x_i)) \end{cases} \quad (4.14)$$

avec :

- $\alpha_{i,j} \in]0, 1[$ coefficient d'affaiblissement (cf. Section 4.4.1.2),
- d une distance statistique à définir en fonction des données,
- φ_i une fonction décroissante vérifiant :

$$\begin{cases} \varphi_i(0) = 1 \\ \lim_{d \rightarrow +\infty} \varphi_i(d) = 0 \end{cases} \quad (4.15)$$

Plusieurs fonctions φ_i peuvent vérifier ces propriétés. Dans le cas où d est une distance euclidienne, [Denoeux, 1995] propose l'emploi d'une fonction exponentielle :

$$\varphi_i(d) = \exp(-\gamma_i d^\beta) \quad (4.16)$$

avec $\gamma_i > 0$ et $\beta \in \{1, 2, \dots\}$. Ces paramètres sont liés au contexte de l'application et déterminés de manière heuristique.

Une autre approche proposée par [Rombaut, 1998] et reprise par [Gruyer et Berge-Cherfaoui, 1999] et [Royère et al., 2000] se base sur des sources dites “spécialisées”. Une source spécialisée Src_i se prononce uniquement sur un singleton H_i , son inverse \bar{H}_i et l'ignorance. Le cadre de discernement d'une source spécialisée est alors défini par $\Theta = \{\{H_i\}, \{\bar{H}_i\}\}$. Les fonctions de masse sont décrites par :

$$\begin{aligned} m_i^\Theta(H_i) &= \begin{cases} 0 & C_v \in [0, \tau] \\ \varphi_1(\alpha_0, C_v) & C_v \in [\tau, 1] \end{cases} \\ m_i^\Theta(\bar{H}_i) &= \begin{cases} \varphi_2(\alpha_0, C_v) & C_v \in [0, \tau] \\ 0 & C_v \in [\tau, 1] \end{cases} \\ m_i^\Theta(\Theta) &= \begin{cases} 1 - \varphi_2(\alpha_0, C_v) & C_v \in [0, \tau] \\ 1 - \varphi_1(\alpha_0, C_v) & C_v \in [\tau, 1] \end{cases} \end{aligned} \quad (4.17)$$

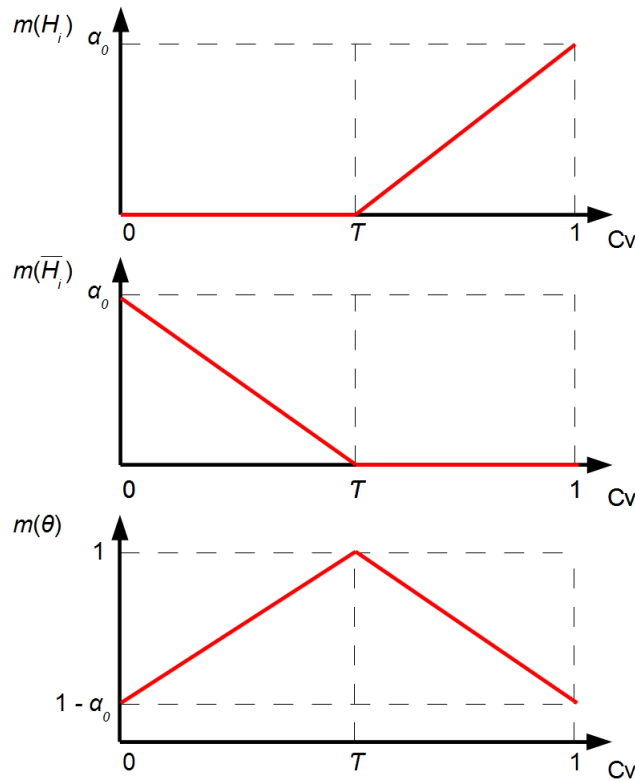


FIGURE 4.3 – Jeu de masses d’une source spécialisée selon [Rombaut, 1998].

avec :

- α_0 valeur limite autorisée pour la masse d’un singleton (H_i) ou de son inverse (\overline{H}_i).
- C_v variable d’évaluation, elle définit le niveau de confiance de l’information fournie par la source.
- φ_1 et φ_2 deux fonctions décrivant l’évolution des masses de croyance. Ces fonctions peuvent être linéaires comme illustré sur la figure 4.3 avec des conditions aux limites :

$$\begin{aligned}
 C_v = 0, \quad \varphi_2(\alpha_0, 0) &= \alpha_0 \\
 C_v = \tau, \quad \varphi_2(\alpha_0, \tau) &= 0 \\
 C_v = \tau, \quad \varphi_1(\alpha_0, \tau) &= 0 \\
 C_v = 1, \quad \varphi_1(\alpha_0, 1) &= \alpha_0
 \end{aligned}
 \tag{4.18}$$

- $\tau \in [0, 1]$ paramètre important influençant la distribution des masses. Si $\tau < 0.5$ on sera optimiste, puisque l’hypothèse H_i aura une masse de croyance même avec de faibles valeurs de C_v . Si $\tau = 0.5$ alors on aura une position neutre. Enfin si $\tau > 0.5$ on sera pessimiste puisque l’hypothèse H_i n’aura de masse non nulle qu’avec des valeurs élevées de C_v .

Le principal avantage de ce jeu de masses non antagonistes est qu’il interdit à une source d’information d’être confiante à la fois (pour un même C_v) en une hypothèse et en son contraire. Ainsi, la combinaison de deux sources spécialisées dont les jeux de masses sont identiques ne génère pas de conflit ¹ [Mourllion, 2006].

1. Classiquement, sans contrainte posée, la combinaison de deux jeux de masses identiques génère du conflit

4.3.4 Combinaison des sources

La combinaison est une étape très importante de la fusion de données. Les informations fournies par les sources (capteurs, experts, etc.) sont souvent imparfaites (imprécises, incertaines, conflictuelles, et ambiguës), ce qui nécessite leur fusion afin d'obtenir une information globale et/ou de plus haut niveau d'abstraction plus fiable. Cette fusion consiste à assembler les masses m_i^Θ des différentes sources Src_i en une nouvelle dite masse combinée. Il existe trois types d'opérateurs pour la combinaison des masses : les opérateurs conjonctifs, les opérateurs disjonctifs, et les opérateurs mixtes.

4.3.4.1 Combinaison conjonctive

L'opérateur le plus souvent utilisé est l'opérateur conjonctif de Dempster-Shafer [Shafer, 1976]. Il combine les masses élémentaires $m_1^\Theta, m_2^\Theta, \dots, m_m^\Theta$ issues des m sources Src_i en considérant les intersections des éléments de 2^Θ . Cet opérateur noté \oplus est défini par :

$$\begin{cases} m^\Theta(A) = m_1^\Theta(A) \oplus m_2^\Theta(A) \oplus \dots \oplus m_m^\Theta(A) = K_\Theta \sum_{\substack{B_1 \cap \dots \cap B_m = A \\ A \neq \emptyset}} \prod_{i=1}^m m_i^\Theta(B_i) \\ m^\Theta(\emptyset) = 0 \end{cases} \quad (4.19)$$

avec

$$\begin{cases} K_\Theta = \frac{1}{1-k_\Theta} \\ k_\Theta = \sum_{B_1 \cap \dots \cap B_m = \emptyset} \prod_{i=1}^m m_i^\Theta(B_i) \end{cases} \quad (4.20)$$

Afin de respecter la notion du monde fermé (exclusivité des solutions)(cf. (4.10)), [Shafer, 1976] redistribue la masse conflictuelle sur les autres masses ce qui permet d'avoir une masse nulle sur l'ensemble vide ($m^\Theta(\emptyset) = 0$) (cf. (4.19)). Cette redistribution consiste à normaliser le jeu de masses par le facteur de normalisation K_Θ (cf. (4.20)) où k_Θ représente le conflit entre les sources. Si $k_\Theta = 0$ alors les sources sont en total accord. Si $k_\Theta = 1$ alors les sources sont en conflit total.

Bien que la normalisation est nécessaire sous l'hypothèse du monde fermé, elle supprime une partie de l'information. En effet, le fait de redistribuer la masse conflictuelle masque le conflit existant entre les sources et il ne sera pas considéré lors de la prise de décision. Par conséquent, d'autres mondes ont été proposés afin de prendre en considération ce problème de conflit.

Monde ouvert Contrairement au monde fermé de [Shafer, 1976] où les hypothèses sont exclusives et exhaustives, [Smets, 1990] propose *le monde ouvert* caractérisé par la non-exhaustivité des solutions. Le cadre de discernement contient toutes les hypothèses connues au stade de la modélisation mais pas forcément toutes les hypothèses possibles pour un problème donné. Ce monde ouvert

s'adapte mieux aux applications pratiques, telles que la reconnaissance d'objets et l'interprétation des scènes routières, où il est souvent impossible de définir précisément l'ensemble des solutions. Dans ce cadre de modélisation, le conflit peut avoir deux interprétations. Soit, il est lié au dysfonctionnement d'une source (non-fiabilité), soit il provient de l'absence d'une ou plusieurs solutions dans Θ . Lors de la combinaison, les masses ne sont pas normalisées et la masse conflictuelle est affectée à l'ensemble vide \emptyset . L'opérateur conjonctif de Smets [Smets, 1990] noté \odot est défini par :

$$m^\Theta(A) = m_1^\Theta(A) \odot m_2^\Theta(A) \odot \dots \odot m_m^\Theta(A) = \sum_{B_1 \cap \dots \cap B_m = A} \prod_{i=1}^m m_i^\Theta(B_i) \quad (4.21)$$

Dans le contexte du monde ouvert où les sources sont considérées fiables, l'origine du conflit reflète alors la non-exhaustivité du cadre de discernement. Dans ce cas, le conflit ($m^\Theta(\emptyset)$) est interprété comme une classe de rejet regroupant toutes les hypothèses non envisagées lors de la définition du cadre de discernement. Cependant, dans la réalité, les sources ne sont pas toujours fiables et il peut y avoir des dysfonctionnements. Par conséquent, le monde ouvert est inapproprié pour les systèmes réels où l'information est imparfaite et où le cadre de discernement évolue au cours du temps.

Monde ouvert étendu Sur les bases du MCT de Smets, [Gruyer et Berge-Cherfaoui, 1999] et [Royère *et al.*, 2000] définissent un nouveau monde nommé *le monde ouvert étendu*. Ce dernier inclut une hypothèse singleton exclusive $*$ proposée initialement par [Rombaut, 1998]. L'hypothèse $*$ regroupe toutes les solutions non prises en compte ou inconnues lors de la création du cadre de discernement. Par conséquent, cette proposition rend le cadre de discernement à nouveau exhaustif, respectant ainsi les deux conditions du monde fermé : l'exclusivité et l'exhaustivité. L'avantage du monde ouvert étendu est le fait de pouvoir distinguer le conflit dû aux dysfonctionnements des sources, représenté par l'ensemble vide \emptyset , de celui lié à l'apparition d'une ou plusieurs nouvelles hypothèses, représenté par l'hypothèse $*$. Si une observation différente des hypothèses H_i est effectuée, alors elle sera affectée à l'hypothèse $*$. L'hypothèse \emptyset représentera quant à elle exclusivement le conflit lié à la non-fiabilité des sources. [Smets, 2007] souligne néanmoins un inconvénient de cette approche lors de la prise de décision puisque la cardinalité de l'hypothèse $*$, nécessaire pour la détermination de la probabilité pignistique (cf. Section 4.3.5.3) n'est pas connue.

4.3.4.2 Combinaison disjonctive

La combinaison disjonctive prend en considération les unions, au lieu des intersections, entre les éléments focaux des m sources [Smets, 1993]. Dans ce cas, l'opérateur disjonctif noté \oslash est défini pour tout $A \in 2^\Theta$ par :

$$m^\Theta(A) = m_1^\Theta(A) \oslash m_2^\Theta(A) \oslash \dots \oslash m_m^\Theta(A) = \sum_{B_1 \cup B_2 \cup \dots \cup B_m = A} \prod_{i=1}^m m_i^\Theta(B_i) \quad (4.22)$$

Le fait de considérer uniquement les unions permet d'être en accord avec le contexte du monde fermé puisque la masse sur l'ensemble vide \emptyset est toujours nulle. Dans ce cas, le conflit n'apparaît pas ce qui peut être utile lorsqu'on a des sources non-fiables ou en désaccord. L'opérateur disjonctif est moins précis que les opérateurs conjonctifs puisqu'il réduit la précision de l'information en dispersant la confiance sur les unions de propositions [Bernet, 2012].

4.3.4.3 Combinaison mixte

[Dubois et Prade, 1988] proposent une combinaison mixte afin de conserver les avantages de la combinaison conjonctive et disjonctive. L'opérateur mixte attribue pour toute proposition $A \in 2^\Theta$ la masse issue de la combinaison conjonctive ajoutée à la masse de l'union des hypothèses qui la compose et qui génèrent du conflit (cf. (4.23)). La partie conjonctive de la combinaison permet de déterminer précisément les masses sur les hypothèses et la partie disjonctive redistribue le conflit existant sur les unions des propositions. Par conséquent, la combinaison mixte constitue un bon compromis entre la précision et la fiabilité.

$$m^\Theta(A) = \sum_{B_1 \cap \dots \cap B_m = A} \prod_{i=1}^m m_i^\Theta(B_i) + \sum_{\substack{B_1 \cup \dots \cup B_m = A \\ B_1 \cap \dots \cap B_m = \emptyset}} \prod_{i=1}^m m_i^\Theta(B_i) \quad (4.23)$$

4.3.5 Prise de décision

Dans la théorie des croyances, la résolution d'un problème donné consiste à prendre une décision parmi toutes les hypothèses possibles à l'issue de la combinaison. Cette prise de décision doit être effectuée parmi les hypothèses du cadre de discernement, c'est-à-dire uniquement sur les propositions simples ($|A| = 1$). Cependant, la combinaison génère aussi des masses de croyance pour des unions de propositions ($|A| > 1$). Par conséquent, il est nécessaire de redistribuer les masses de ces unions sur les singletons pour pouvoir prendre une décision. La prise de décision nécessite par conséquent une étape de transformation de 2^Θ dans Θ . Cette section décrit les transformations les plus courantes ainsi que les critères de décision les plus employés.

4.3.5.1 Maximum de crédibilité

La première approche choisit le singleton H_i le plus crédible :

$$\max_i [Cr_\Theta(H_i)] \quad (4.24)$$

avec

$$Cr_\Theta(H_i) = \sum_{A \subseteq H_i} m^\Theta(A) \quad (4.25)$$

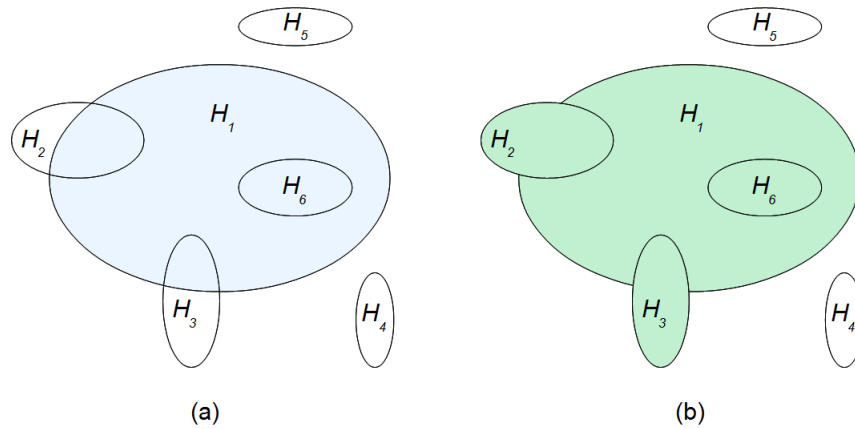


FIGURE 4.4 – Représentation des sous-ensembles. (a) de la crédibilité de H_1 ; (b) de la plausibilité de H_1 [Bernet, 2012].

La fonction de crédibilité $Cr_{\Theta}(H_i)$ représente l'intensité avec laquelle les informations données par les sources confirment l'hypothèse H_i . Elle additionne toutes les croyances placées sur les propositions impliquant l'hypothèse H_i , c'est-à-dire celles qui sont plus spécifiques que H_i . La figure 4.4 (a) représente la crédibilité de H_1 où seules les propositions incluses dans le sous-ensemble H_1 , c'est-à-dire celles qui soutiennent totalement H_1 , sont retenues (H_1 et H_6). La prise de décision par maximum de crédibilité est une approche pessimiste où l'hypothèse retenue est celle qui procure le minimum de chance à chacune d'entre elles.

4.3.5.2 Maximum de plausibilité

Cette deuxième approche consiste à choisir le singleton H_i donnant le maximum de plausibilité :

$$\max_i [Pl_{\Theta}(H_i)] \quad (4.26)$$

avec

$$Pl_{\Theta}(H_i) = 1 - Cr_{\Theta}(\overline{H_i}) = \sum_{A \cap H_i = \emptyset} m^{\Theta}(A) \quad (4.27)$$

La fonction de plausibilité $Pl_{\Theta}(H_i)$ mesure l'intensité avec laquelle les informations fournies par la source ne contredisent pas H_i . La plausibilité traduit la crédibilité d'une hypothèse à laquelle s'ajoute toute la masse potentielle que peut recevoir cette proposition. La figure 4.4 (b) illustre la plausibilité de H_1 où toutes les propositions ayant une intersection avec l'hypothèse H_1 sont retenues (H_1 , H_2 , H_3 , et H_6). Cette approche de décision est dite optimiste puisqu'elle donne le maximum de chance à chacun des singletons.

4.3.5.3 Maximum de probabilité pignistique

[Smets et Kennes, 1994] propose, pour son MCT (et donc dans le contexte d’un monde ouvert où la masse sur l’ensemble vide est tolérée), la probabilité pignistique comme un compromis aux critères du maximum de crédibilité et de plausibilité. La probabilité pignistique tend à se rapprocher de la théorie bayésienne en offrant un critère de sélection semblable à une probabilité. Cette approche redistribue de manière équitable les masses de chaque union de propositions sur les singletons qui la composent(cf. (4.28)). Cette méthode est qualifiée de “prudente”, la probabilité pignistique se situe entre la crédibilité et la plausibilité.

$$BetP_{\Theta}(H_i) = \sum_{A \in 2^{\Theta}, H_i \subseteq A} \frac{1}{|A|(1 - m_{\cap}^{\Theta}(\emptyset))} m^{\Theta}(A) \quad (4.28)$$

Ainsi l’hypothèse retenue est celle ayant le maximum de probabilité pignistique :

$$\max_i [BetP_{\Theta}(H_i)] \quad (4.29)$$

Notons que pour la prise de décision, la masse sur l’ensemble vide doit être redistribuée lors du calcul des probabilités pignistiques [Smets, 2005].

4.4 Association multi-objets par MCT

La théorie des croyances a été exploitée dans plusieurs applications de suivi multi-objets dont l’objectif est l’interprétation des événements de la scène à partir de données imparfaites [Gruyer et Berge-Cherfaoui, 1999], [Mercier *et al.*, 2011], [Dallil *et al.*, 2013] et [El Zoghby *et al.*, 2013]. Dans ce contexte, la problématique considérée est d’associer les objets perçus X (cibles), à l’instant k , avec les objets connus Y (pistes) à l’instant $k - 1$ afin d’analyser les événements de la scène (cf. Figure 4.5). Dans un environnement dynamique comme une scène routière, cette étape d’association doit impérativement gérer l’apparition de nouveaux objets et la disparition d’objets connus (e.g. un cycliste qui sort du champ de vision du conducteur). En plus, ce processus d’association doit tenir compte de l’ignorance qu’on peut avoir sur l’environnement. En effet, dans un environnement complexe et bruité, le système de perception (i.e. capteurs et/ou algorithmes de perception) est caractérisé par une fiabilité variable selon les conditions extérieures (météorologie, densité du trafic, etc.) ce qui ne permet pas d’observer systématiquement l’ensemble des événements, d’où l’intérêt d’utiliser la théorie des croyances. [Rombaut, 1998] est le premier à avoir utilisé cette théorie pour reconstruire l’environnement d’un véhicule intelligent en mesurant la confiance des hypothèses d’association entre obstacles perçus X et obstacles connus Y par la combinaison de jeux de masses. Par la suite, [Gruyer et Berge-Cherfaoui, 1999] ont introduit une mesure de similarité (basée sur la distance, la vitesse, etc.) pour l’estimation des masses. Dans ces travaux, le problème de mise

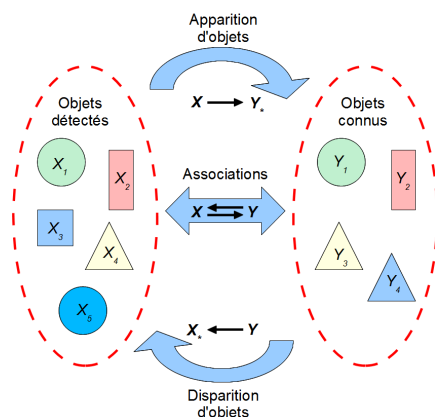


FIGURE 4.5 – Principe d’association de données par théorie des croyances [Daniel et Lauffenburger, 2012].

en relation est résolu selon deux points de vues : $X \rightarrow Y$ et $X \leftarrow Y$ (cf. Figure 4.5). En plus d’une meilleure gestion des apparitions et disparitions, cette mise en relation pistes/cibles et cibles/pistes permet de détecter le conflit qui peut exister entre les sources d’information. Ces travaux ont été repris par [Mourllion, 2006] et [Mercier *et al.*, 2011] pour le suivi de véhicules où le processus d’association est basé sur le MCT de [Smets et Kennes, 1994]. [El Zoghby *et al.*, 2013] et [Denoeux *et al.*, 2014] exploitent également la théorie des croyances pour améliorer la perception de l’environnement en fusionnant deux types de capteurs : une caméra et un scanner laser embarqués sur le véhicule. Dans ces travaux, une formulation différente est proposée visant à optimiser la plausibilité d’association par l’intermédiaire de la programmation linéaire. Les auteurs proposent également d’estimer les masses en tenant compte de la classe des objets perçus (piétons, véhicule, etc.) ainsi que d’une distance statistique, la distance de Mahalanobis [Blackman et Popoli, 1999]. Cette dernière est normalisée dans [Dallil *et al.*, 2013] avant l’estimation des masses ce qui permet une meilleure distribution des croyances sur l’ensemble des hypothèses d’association.

Dans cette thèse, nous avons étendu l’application du MCT au suivi de panneaux afin d’améliorer les performances du système TSR. Le MCT de [Smets et Kennes, 1994] est une interprétation subjective et non-probabiliste de la théorie des croyances. Son formalisme décrit un modèle de raisonnement incertain et de prise de décision basé sur deux niveaux : crédal et pignistique. Au niveau crédal, les masses de croyance sont utilisées pour représenter et combiner les différentes parties de l’information (i.e. connaissance fournie par les capteurs), tandis qu’au niveau pignistique, ces masses sont transformées en mesures de probabilité pour la prise de décision.

4.4.1 Niveau crédal

Dans ce premier niveau, les informations perçues par les sources sont modélisées et évaluées par des bbm pour être combinées par la suite et obtenir une information globale plus fiable.

4.4.1.1 Modélisation

La première étape consiste à définir le cadre de discernement pour le problème d'association de données. Soit n cibles (X) détectées (objets perçus) à l'instant k et m pistes (Y) prédites (objets connus). Une hypothèse d'association définit une relation entre une cible X_i , $i \in [1, n]$, et une piste Y_j , $j \in [1, m]$. Cette association peut être exprimée comme “la cible X_i est associée à la piste Y_j ” avec $i \in [1, n]$ et $j \in [1, m]$. Un ensemble d'associations représente une solution possible au problème de mise en relation à l'instant k comme par exemple celle fournie par le MHT ou la JPDA.

Dans la théorie des croyances, le processus d'association de données est bidirectionnel [Gruyer et Berge-Cherfaoui, 1999], c'est-à-dire traité en deux étapes : déterminer en premier lieu les associations *cible-piste* ($X \rightarrow Y$), ensuite les associations *piste-cible* ($X \leftarrow Y$) (cf. Figure 4.5). Cette méthode de résolution permet une meilleure gestion des apparitions et disparitions d'objets mais nécessite la définition de deux types de cadres de discernement dans le monde ouvert étendu : $\Theta_{i..}$, qui contient l'ensemble des pistes, Y_j avec $j = 1, \dots, m$, qui peuvent être associées à la cible X_i et $\Theta_{..j}$ contenant l'ensemble des cibles, X_i avec $i = 1, \dots, n$, qui peuvent être associées à la piste Y_j . Ainsi les cadres de discernement seront :

$$\begin{aligned}\Theta_{i..} &= \{\{Y_1\}, \{Y_2\}, \dots, \{Y_m\}, \{Y_*\}\} \\ \Theta_{..j} &= \{\{X_1\}, \{X_2\}, \dots, \{X_n\}, \{X_*\}\}\end{aligned}\quad (4.30)$$

avec

- $\{Y_j\}$: hypothèse d'association de la cible X_i à la piste Y_j .
- $\{X_i\}$: hypothèse d'association de la piste Y_j à la cible X_i .
- $\{Y_*\}$: hypothèse de non association de la cible X_i , ce qui représente l'apparition d'une nouvelle piste (cf. Figure 4.5).
- $\{X_*\}$: hypothèse de non association de la piste Y_j , ce qui représente la disparition de cette dernière (cf. Figure 4.5).

Dans un processus d'association multi-objets, la question à laquelle il faut répondre est la suivante : “la cible X_i est-elle associée à la piste Y_j ?”. Pour cette question, trois réponses sont possibles : oui l'association est effective, non elle ne l'est pas, ou il y a un doute alors on ne se prononce pas. Dans le MCT de [Smets et Kennes, 1994], ces réponses sont définies sur le noyau $N_\Theta = \{\{Y_j\}, \{\overline{Y_j}\}, \Theta_{i..}\}$ et évaluées par trois masses initiales de croyance $m_j^{\Theta_{i..}}(\cdot)$:

- $m_j^{\Theta_{i..}}(\{Y_j\})$: masse de croyance sur l'hypothèse “ X_i est associée à Y_j ”,
- $m_j^{\Theta_{i..}}(\{\overline{Y_j}\})$: masse de croyance sur l'hypothèse “ X_i n'est pas associée à Y_j ”,
- $m_j^{\Theta_{i..}}(\Theta_{i..})$: masse de croyance sur l'ignorance de l'association.

avec $\{\overline{Y_j}\} = \{\{Y_1\}, \dots, \{Y_{j-1}\}, \{Y_{j+1}\}, \dots, \{Y_m\}, \{Y_*\}\} = \Theta_{i..} \setminus Y_j$.

De la même manière, le jeu de masses initial pour les associations $X \leftarrow Y$ est défini par :

- $m_i^{\Theta_{..j}}(\{X_i\})$: masse de croyance sur l'hypothèse “ Y_j est associée à X_i ”,

- $m_i^{\Theta_{\cdot,j}}(\{\overline{X}_i\})$: masse de croyance sur l’hypothèse “ Y_j n’est pas associée à X_i ”,
- $m_i^{\Theta_{\cdot,j}}(\Theta_{\cdot,j})$: masse de croyance sur l’ignorance de l’association.

Ce type de distribution de masses a été formalisée par [Rombaut, 1998] (cf. Section 4.3.3). L’approche consiste à convertir une mesure de similarité, entre une cible X_i et une piste Y_j , en une croyance sur une hypothèse, sur son contraire, ainsi que sur l’ignorance.

4.4.1.2 Estimation

L’étape d’estimation quantifie les masses de croyance à partir des mesures de similarité entre les cibles et les pistes. Ces mesures représentent les distances des cibles observées (X) aux pistes connues (Y) [Denoeux, 1995], [Mercier *et al.*, 2011], et [El Zoghby *et al.*, 2013]. Dans la plupart des applications d’association d’objets, la mesure utilisée est la distance de Mahalanobis $d_{i,j}$ entre la cible X_i et la piste Y_j (cf. (2.31)) [Blackman et Popoli, 1999]. Cette distance statistique à l’avantage de prendre en considération les erreurs de mesures des cibles et d’estimation des pistes. L’initialisation du jeu de masses est basée sur l’approche présentée dans [Denoeux, 1995] adaptée au triplet $\{\{Y_j\}, \{\overline{Y}_j\}, \Theta_{i,\cdot}\}$ (cf. (4.14)). Les fonctions de masses générées sur $\Theta_{i,\cdot}$, (de la même manière pour $\Theta_{\cdot,j}$) sont définies par :

$$\begin{cases} m_j^{\Theta_{i,\cdot}}(\{Y_j\}) = \alpha \exp^{-\gamma d_{i,j}^\beta} \\ m_j^{\Theta_{i,\cdot}}(\{\overline{Y}_j\}) = \alpha(1 - \exp^{-\gamma d_{i,j}^\beta}) \\ m_j^{\Theta_{i,\cdot}}(\Theta_{i,\cdot}) = 1 - \alpha \end{cases} \quad (4.31)$$

avec

- $0 < \alpha < 1$ coefficient d’affaiblissement qui permet de tenir compte de la fiabilité a priori connue de la source.
- $\gamma \in \mathbb{R}^*$ et $\beta \in \mathbb{N}^*$ peuvent être fixés à de faibles valeurs [Denoeux, 1995].

Les tables 4.1 et 4.2 représentent respectivement les bbms initiales correspondant aux hypothèses d’association $X \longrightarrow Y$ et aux hypothèses d’association $X \longleftarrow Y$. Durant l’étape d’estimation, aucune masse n’est initialement attribuée aux hypothèses $\{Y_*\}$ et $\{X_*\}$. Celles-ci vont potentiellement apparaître durant la combinaison des bbms.

Exemple d’illustration Soit une cible observée (X_1) à l’instant k et deux pistes connues (Y_1 et Y_2) issues d’une précédente détection. Les distances de Mahalanobis déterminées entre la cible et les deux pistes sont données par :

$$\begin{cases} d_{1,1} = 4 \\ d_{1,2} = 12 \end{cases} \quad (4.32)$$

Ces distances montrent que la cible est plus proche de la piste Y_1 que de Y_2 . Ainsi, la confiance dans l’association $X_1 \longrightarrow Y_1$ devrait être élevée.

TABLE 4.1 – Masses initiales des associations $X \rightarrow Y$

	Y_1	Y_2	...	Y_m	Y_*
X_1	$m_1^{\Theta_{1..}}(\{Y_1\})$	$m_2^{\Theta_{1..}}(\{Y_2\})$...	$m_m^{\Theta_{1..}}(\{Y_m\})$	0
	$m_1^{\Theta_{1..}}(\{\overline{Y_1}\})$	$m_2^{\Theta_{1..}}(\{\overline{Y_2}\})$...	$m_m^{\Theta_{1..}}(\{\overline{Y_m}\})$	0
	$m_1^{\Theta_{1..}}(\Theta_{1..})$	$m_2^{\Theta_{1..}}(\Theta_{1..})$...	$m_m^{\Theta_{1..}}(\Theta_{1..})$	0
X_2	$m_1^{\Theta_{2..}}(\{Y_1\})$	$m_2^{\Theta_{2..}}(\{Y_2\})$...	$m_m^{\Theta_{2..}}(\{Y_m\})$	0
	$m_1^{\Theta_{2..}}(\{\overline{Y_1}\})$	$m_2^{\Theta_{2..}}(\{\overline{Y_2}\})$...	$m_m^{\Theta_{2..}}(\{\overline{Y_m}\})$	0
	$m_1^{\Theta_{2..}}(\Theta_{2..})$	$m_2^{\Theta_{2..}}(\Theta_{2..})$...	$m_m^{\Theta_{2..}}(\Theta_{2..})$	0
⋮	⋮	⋮	⋮	⋮	⋮
X_n	$m_1^{\Theta_{n..}}(\{Y_1\})$	$m_2^{\Theta_{n..}}(\{Y_2\})$...	$m_m^{\Theta_{n..}}(\{Y_m\})$	0
	$m_1^{\Theta_{n..}}(\{\overline{Y_1}\})$	$m_2^{\Theta_{n..}}(\{\overline{Y_2}\})$...	$m_m^{\Theta_{n..}}(\{\overline{Y_m}\})$	0
	$m_1^{\Theta_{n..}}(\Theta_{n..})$	$m_2^{\Theta_{n..}}(\Theta_{n..})$...	$m_m^{\Theta_{n..}}(\Theta_{n..})$	0

TABLE 4.2 – Masses initiales des associations $X \leftarrow Y$

	X_1	X_2	...	X_n	X_*
Y_1	$m_1^{\Theta_{..1}}(\{X_1\})$	$m_2^{\Theta_{..1}}(\{X_2\})$...	$m_n^{\Theta_{..1}}(\{X_n\})$	0
	$m_1^{\Theta_{..1}}(\{\overline{X_1}\})$	$m_2^{\Theta_{..1}}(\{\overline{X_2}\})$...	$m_n^{\Theta_{..1}}(\{\overline{X_n}\})$	0
	$m_1^{\Theta_{..1}}(\Theta_{..1})$	$m_2^{\Theta_{..1}}(\Theta_{..1})$...	$m_n^{\Theta_{..1}}(\Theta_{..1})$	0
Y_2	$m_1^{\Theta_{..2}}(\{X_1\})$	$m_2^{\Theta_{..2}}(\{X_2\})$...	$m_n^{\Theta_{..2}}(\{X_n\})$	0
	$m_1^{\Theta_{..2}}(\{\overline{X_1}\})$	$m_2^{\Theta_{..2}}(\{\overline{X_2}\})$...	$m_n^{\Theta_{..2}}(\{\overline{X_n}\})$	0
	$m_1^{\Theta_{..2}}(\Theta_{..2})$	$m_2^{\Theta_{..2}}(\Theta_{..2})$...	$m_n^{\Theta_{..2}}(\Theta_{..2})$	0
⋮	⋮	⋮	⋮	⋮	⋮
Y_m	$m_1^{\Theta_{..m}}(\{X_1\})$	$m_2^{\Theta_{..m}}(\{X_2\})$...	$m_n^{\Theta_{..m}}(\{X_n\})$	0
	$m_1^{\Theta_{..m}}(\{\overline{X_1}\})$	$m_2^{\Theta_{..m}}(\{\overline{X_2}\})$...	$m_n^{\Theta_{..m}}(\{\overline{X_n}\})$	0
	$m_1^{\Theta_{..m}}(\Theta_{..m})$	$m_2^{\Theta_{..m}}(\Theta_{..m})$...	$m_n^{\Theta_{..m}}(\Theta_{..m})$	0

À partir de (4.31) et avec $\alpha = 0.9$, $\gamma = 0.01$, et $\beta = 2$, les jeux de masses initiaux obtenus pour les associations : $X_1 \rightarrow Y_1$ et $X_1 \rightarrow Y_2$ sont :

$$X_1 \begin{cases} Y_1 \\ m_1^{\Theta_{1..}}(\{Y_1\}) = 0.7669 \\ m_1^{\Theta_{1..}}(\{\overline{Y_1}\}) = 0.1331 \\ m_1^{\Theta_{1..}}(\Theta_{1..}) = 0.1 \end{cases} \quad \begin{cases} Y_2 \\ m_2^{\Theta_{1..}}(\{Y_2\}) = 0.2132 \\ m_2^{\Theta_{1..}}(\{\overline{Y_2}\}) = 0.6868 \\ m_2^{\Theta_{1..}}(\Theta_{1..}) = 0.1 \end{cases} \quad (4.33)$$

Ces bbms sont en accord avec les distances observées (cf. (4.32)) puisqu'elles attribuent une confiance élevée à l'association de X_1 avec Y_1 (0.7669) ainsi qu'à la non association de X_1 avec Y_2 (0.6868).

4.4.1.3 Combinaison des bbms

Avant la prise de décision, il faut combiner les masses obtenues à l'issue de l'étape d'estimation afin d'obtenir une vue d'ensemble des solutions d'association. Puisque le problème de mise en relation est résolu de manière bidirectionnelle dans le MCT, la combinaison est effectuée sur les jeux de masses des tables 4.1 et 4.2 de façon indépendante. Dans la table 4.1, les jeux de masses déterminés à partir des sources spécialisées² concernant une cible X_i sont combinés, ce qui permet de calculer des masses de croyance, notées $m^{\Theta_{i..}}(A)$, sur toute proposition A du référentiel $2^{\Theta_{i..}}$. Le fait de définir les cadres de discernement ($\Theta_{i..}$ et $\Theta_{j..}$) dans le contexte du monde ouvert étendu influe sur le choix de l'opérateur de combinaison. En effet, l'opérateur choisi doit tenir compte de ce contexte et permettre aussi l'attribution de masse non nulle à l'ensemble vide \emptyset . Cette masse conflictuelle est considérée comme une source d'information pertinente pour détecter les associations ambiguës et contradictoires [Mourllion, 2006]. Diverses combinaisons ont été proposées dans la littérature. Nous résumons ci-dessous les plus employées.

[Gruyer et Berge-Cherfaoui, 1999] ne calculent que les masses de singletons et affectent les masses des propositions composites directement sur l'ignorance $\Theta_{i..}$. La combinaison de masses est ainsi définie par :

$$\begin{aligned}
m^{\Theta_{i..}}(\{Y_j\}) &= K_{i..} m_j^{\Theta_{i..}}(\{Y_j\}) \prod_{\substack{a=1 \\ a \neq j}}^m (1 - m_a^{\Theta_{i..}}(\{Y_a\})) \\
m^{\Theta_{i..}}(\Theta_{i..}) &= K_{i..} \left(\prod_{a=1}^m [m_a^{\Theta_{i..}}(\Theta_{i..}) + m_a^{\Theta_{i..}}(\{Y_a\})] - \prod_{a=1}^m m_a^{\Theta_{i..}}(\overline{\{Y_a\}}) \right) \\
m^{\Theta_{i..}}(\{Y_*\}) &= K_{i..} \prod_{a=1}^m m_a^{\Theta_{i..}}(\overline{\{Y_a\}})
\end{aligned} \tag{4.34}$$

2. une source spécialisée caractérise un triplet $m_j^{\Theta_{i..}}(\{Y_j\}), m_j^{\Theta_{i..}}(\overline{\{Y_j\}}), m_j^{\Theta_{i..}}(\Theta_{i..})$

avec $K_{i,\cdot} = \prod_{a=1}^n K_{i,a}$ la normalisation des m combinaisons, et :

$$\begin{aligned}
m^{\Theta_{i,\cdot}}(\emptyset) &= \left[\prod_{a=1}^{m-1} K_{i,a} \right] m_m^{\Theta_{i,\cdot}}(\{Y_m\}) \sum_{a=1}^{m-1} \left(m_a^{\Theta_{i,\cdot}}(\{Y_a\}) \left[\prod_{\substack{b=1 \\ b \neq a}}^m -1 A_m \right] \right) \\
A_m &= m_m^{\Theta_{i,\cdot}}(\Theta_{i,m}) + m_m^{\Theta_{i,\cdot}}(\overline{\{Y_a\}}) \\
K_{i,\cdot} &= \frac{1}{1 - m^{\Theta_{i,\cdot}}(\emptyset)} \\
K_{i,\cdot} &= \frac{1}{\prod_{j=1}^m (1 - m_j^{\Theta_{i,\cdot}}(Y_j)) \left(1 + \sum_{j=1}^m \frac{m_j^{\Theta_{i,\cdot}}(Y_j)}{1 - m_j^{\Theta_{i,\cdot}}(Y_j)} \right)}
\end{aligned} \tag{4.35}$$

De ce fait, la confiance n'est attribuée qu'à des singletons, à l'ignorance et à l'apparition/disparition. La masse d'hypothèses composées étant transférée sur l'ignorance (en assimilant l'ignorance partielle à l'ignorance totale (cf. Section 4.3.2)), celle-ci a donc fortement tendance à croître au cours du temps.

[Royère *et al.*, 2000] proposent une combinaison plus complexe mais plus raffinée où une masse de croyance est calculée pour chaque proposition du référentiel $2^{\Theta_{i,\cdot}}$. Les masses finales de toutes les associations possibles (y compris pour les disjonctions) sont définies par :

$$\begin{aligned}
m^{\Theta_{i,\cdot}}(\{Y_j\}) &= m_j^{\Theta_{i,\cdot}}(\{Y_j\}) \prod_{\substack{a=1 \\ a \neq j}}^m (1 - m_a^{\Theta_{i,\cdot}}(\{Y_a\})) \\
m^{\Theta_{i,\cdot}}(\{Y_j, Y_l\}) &= m_j^{\Theta_{i,\cdot}}(\Theta_{i,\cdot}) m_l^{\Theta_{i,\cdot}}(\Theta_{i,\cdot}) \prod_{\substack{a=1 \\ a \neq j \\ a \neq l}}^m (m_a^{\Theta_{i,\cdot}}(\overline{\{Y_a\}}))
\end{aligned}$$

Pour toutes les combinaisons d'union de 2 à $m - 1$ hypothèses :

$$\begin{aligned}
m^{\Theta_{i..}}(\{Y_j, \dots, Y_l\}) &= m_j^{\Theta_{i..}}(\Theta_{i..}) \dots m_l^{\Theta_{i..}}(\Theta_{i..}) \prod_{\substack{a=1 \\ a \neq j \\ \dots \\ a \neq l}}^m \left(m_a^{\Theta_{i..}}(\overline{\{Y_a\}}) \right) \\
m^{\Theta_{i..}}(\{Y_*\}) &= \prod_{a=1}^m m_a^{\Theta_{i..}}(\overline{\{Y_a\}}) \\
m^{\Theta_{i..}}(\overline{\{Y_j\}}) &= m_j^{\Theta_{i..}}(\overline{\{Y_j\}}) \prod_{\substack{a=1 \\ a \neq j}}^m m_a^{\Theta_{i..}}(\Theta_{i..}) \\
m^{\Theta_{i..}}(\Theta_{i..}) &= \prod_{a=1}^m m_a^{\Theta_{i..}}(\Theta_{i..}) \\
m^{\Theta_{i..}}(\emptyset) &= 1 - \left[\prod_{a=1}^m \left(1 - m_a^{\Theta_{i..}}(\{Y_a\}) \right) + \sum_{a=1}^m m_a^{\Theta_{i..}}(\{Y_a\}) \prod_{\substack{b=1 \\ b \neq a}}^m \left(1 - m_b^{\Theta_{i..}}(\{Y_b\}) \right) \right]
\end{aligned} \tag{4.36}$$

L'équation (4.36) est appliquée pour chaque cible X_i afin d'obtenir la distribution de croyance $m^{\Theta_{i..}}(\cdot)$ sur l'ensemble des associations possibles (association de X_i à m pistes). Le même calcul est effectué pour chaque piste Y_j afin d'obtenir la distribution de croyance $m^{\Theta_{..j}}(\cdot)$ sur l'ensemble des associations possibles (association de Y_j à n cibles). La combinaison dans un monde ouvert étendu génère une masse sur l'hypothèse Y_* (X_*) et l'ensemble vide \emptyset , ce qui permet de prendre des décisions respectivement sur l'apparition/disparition de pistes et la détection de situations conflictuelles et ambiguës.

4.4.2 Niveau pignistique

Le niveau pignistique consiste à choisir, à l'issue de la combinaison des masses, les hypothèses d'association les plus certaines en fonction d'un des critères présentés dans la section 4.3.5. Cependant, ces associations constituant la "meilleure" solution au problème doivent vérifier les contraintes suivantes [Gruyer et Berge-Cherfaoui, 1999], [Mercier et al., 2011] :

- les associations multiples ne sont pas acceptées, une cible est associée à une seule piste au plus et *vice versa*,
- plusieurs nouvelles pistes peuvent apparaître,
- plusieurs pistes existantes peuvent disparaître.

TABLE 4.3 – Matrice de croyance $M_{i,..}^{cr}(X \rightarrow Y)$ [Gruyer et Berge-Cherfaoui, 1999]

$M_{i,..}^{cr}$	Y_1	\dots	Y_m	Y_*	$\Theta_{i,..}$	
X_1	$m^{\Theta_{1,..}}(Y_1)$		$m^{\Theta_{1,..}}(Y_m)$	$m^{\Theta_{1,..}}(Y_*)$	$m^{\Theta_{1,..}}(\Theta_{1,..})$	$\sum_{A \subseteq \Theta_{1,..}} m^{\Theta_{1,..}}(A) = 1$
X_2	$m^{\Theta_{2,..}}(Y_1)$		$m^{\Theta_{2,..}}(Y_m)$	$m^{\Theta_{2,..}}(Y_*)$	$m^{\Theta_{2,..}}(\Theta_{2,..})$	$\sum_{A \subseteq \Theta_{2,..}} m^{\Theta_{2,..}}(A) = 1$
\vdots						
X_n	$m^{\Theta_{n,..}}(Y_1)$		$m^{\Theta_{n,..}}(Y_m)$	$m^{\Theta_{n,..}}(Y_*)$	$m^{\Theta_{n,..}}(\Theta_{n,..})$	$\sum_{A \subseteq \Theta_{n,..}} m^{\Theta_{n,..}}(A) = 1$

TABLE 4.4 – Matrice de croyance $M_{.,j}^{cr}(X \leftarrow Y)$ [Gruyer et Berge-Cherfaoui, 1999]

$M_{.,j}^{cr}$	X_1	\dots	X_n	X_*	$\Theta_{.,j}$	
Y_1	$m^{\Theta_{.,1}}(X_1)$		$m^{\Theta_{.,1}}(X_n)$	$m^{\Theta_{.,1}}(X_*)$	$m^{\Theta_{.,1}}(\Theta_{.,1})$	$\sum_{A \subseteq \Theta_{.,1}} m^{\Theta_{.,1}}(A) = 1$
Y_2	$m^{\Theta_{.,2}}(X_1)$		$m^{\Theta_{.,2}}(X_n)$	$m^{\Theta_{.,2}}(X_*)$	$m^{\Theta_{.,2}}(\Theta_{.,2})$	$\sum_{A \subseteq \Theta_{.,2}} m^{\Theta_{.,2}}(A) = 1$
\vdots						
Y_m	$m^{\Theta_{.,m}}(X_1)$		$m^{\Theta_{.,m}}(X_n)$	$m^{\Theta_{.,m}}(X_*)$	$m^{\Theta_{.,m}}(\Theta_{.,m})$	$\sum_{A \subseteq \Theta_{.,m}} m^{\Theta_{.,m}}(A) = 1$

Dans ce contexte, il est nécessaire de prendre les décisions sur les hypothèses singletons (donc dans $\Theta_{i,..}$) et non pas sur les propositions d'union (dans $2^{\Theta_{i,..}}$). [Gruyer et Berge-Cherfaoui, 1999] générant uniquement des masses sur les singletons, la décision se base sur deux matrices de croyance : $M_{i,..}^{cr}$ et $M_{.,j}^{cr}$ définies par l'ensemble des masses issues de la combinaison (cf. (4.34)). La matrice $M_{i,..}^{cr}$, liée aux associations $X \rightarrow Y$ est illustrée par la table 4.3, elle définit le jeu de masses normalisé $m^{\Theta_{i,..}}(\cdot)$ suite à la redistribution de la masse du conflit $m^{\Theta_{i,..}}(\emptyset)$. De même, la table 4.4 présente la matrice $M_{.,j}^{cr}$, liée aux associations $X \leftarrow Y$.

Dans [Royère et al., 2000], la combinaison est basée sur (4.36), ce qui génère des masses sur les disjonctions. La transformation pignistique de smets (cf. Section 4.3.5.3) est utilisée pour redistribuer les masses des propositions union et ainsi permettre une décision sur $\Theta_{i,..}$. À partir de mesures de probabilité, deux nouvelles matrices $BetP_{i,..}$ et $BetP_{.,j}$ sont ainsi obtenues (cf. Tables 4.5 et 4.6). Contrairement à la transformation de Smets (cf. (4.28)) où la masse conflictuelle est redistribuée, [Royère, 2002] définit une nouvelle transformation qui conserve cette masse. Cette transformation non-normalisée a été par la suite exploitée pour détecter les associations ambiguës et contradictoires [Mourllion, 2006], [Daniel et Lauffenburger, 2012]. Elle est définie par :

$$\begin{cases} BetP_{i,..}(\{Y_j\}) &= \sum_{\substack{A \in 2^{\Theta_{i,..}} \\ Y_j \in A}} \frac{m^{\Theta_{i,..}}(A)}{|A|} \\ BetP_{i,..}(\emptyset) &= m^{\Theta_{i,..}}(\emptyset) \end{cases} \quad (4.37)$$

L'association crédale intégrée à l'algorithme SMR sera exclusivement basée sur ces deux matrices d'association (et donc sur les relations de combinaison (4.36)) du fait d'une meilleure prise en

TABLE 4.5 – Matrice pignistique $BetP_{i..}$ (associations cible-à-piste) [Royère, 2002]

$BetP_{i..}$	Y_1	...	Y_m	Y_*	\emptyset
X_1	$BetP_{1..}(Y_1)$...	$BetP_{1..}(Y_m)$	$BetP_{1..}(Y_*)$	$BetP_{1..}(\emptyset)$
X_2	$BetP_{2..}(Y_1)$...	$BetP_{2..}(Y_m)$	$BetP_{2..}(Y_*)$	$BetP_{2..}(\emptyset)$
\vdots	\vdots	...	\vdots	\vdots	\vdots
X_n	$BetP_{n..}(Y_1)$...	$BetP_{n..}(Y_m)$	$BetP_{n..}(Y_*)$	$BetP_{n..}(\emptyset)$

TABLE 4.6 – Matrice pignistique $BetP_{.,j}$ (associations piste-à-cible) [Royère, 2002]

$BetP_{.,j}$	X_1	...	X_n	X_*	\emptyset
Y_1	$BetP_{.,1}(X_1)$...	$BetP_{.,1}(X_n)$	$BetP_{.,1}(X_*)$	$BetP_{.,1}(\emptyset)$
Y_2	$BetP_{.,2}(X_1)$...	$BetP_{.,2}(X_n)$	$BetP_{.,2}(X_*)$	$BetP_{.,2}(\emptyset)$
\vdots	\vdots	\vdots	...	\vdots	\vdots
Y_m	$BetP_{.,m}(X_1)$...	$BetP_{.,m}(X_n)$	$BetP_{.,m}(X_*)$	$BetP_{.,m}(\emptyset)$

compte de l'ignorance partielle et du conflit.

À partir des matrices pignistiques, il convient à présent de définir une stratégie de sélection des “meilleures” associations pour répondre à la problématique initiale qui est la mise en correspondance entre les pistes connues et les cibles perçues. Cette étape est cruciale pour définir finalement les relations entre les cibles et les pistes. La section suivante présente un état de l'art des récents algorithmes de prise de décision pour l'association de données. Le choix retenu dans le cadre du suivi de panneaux à l'aide du système SMR est également largement commenté.

4.5 Algorithmes de prise de décision par MCT

Pour le problème d'association de données, la prise de décision détermine la solution optimale en se basant sur deux mesures de coût : les matrices $BetP_{i..}$ et $BetP_{.,j}$. Différentes stratégies de prise de décision ont été proposées selon les objectifs recherchés [Mourllion *et al.*, 2005], [Mercier *et al.*, 2011], [Daniel et Lauffenburger, 2012], et [Denoeux *et al.*, 2014]. On distingue deux catégories d'approches selon le type d'optimisation recherchée : globale ou locale. L'approche globale sélectionne les hypothèses d'association maximisant une croyance [Gruyer et Berge-Cherfaoui, 1999] ou une probabilité pignistique globale [Royère, 2002] obtenue à partir de chacune des deux matrices pignistiques. Dans ce cas, l'ensemble des solutions possibles doit être déterminé selon un critère global. Pour [Mercier *et al.*, 2011], la solution choisie est celle qui maximise le produit des probabilités pignistiques des associations constituant cette dernière. Cependant, le calcul des probabilités jointes (cf. Section 4.5.1) et la recherche du maximum engendrent une complexité exponentielle qui n'est pas souhaitée dans les applications temps réel. [Denoeux *et al.*, 2014] remédient à cet inconvénient en transformant le problème d'association en un problème de programmation linéaire où l'objectif est de minimiser le coût d'association. Néanmoins, la “meilleure” solution d'un point de vue

global peut contenir une où plusieurs associations suspicieuses. Par conséquent, d'autres auteurs privilégient une optimisation locale [Daniel et Lauffenburger, 2012]. Dans ce cas, la "meilleure" solution est identifiée par une recherche des maxima locaux des probabilités pignistiques. Cependant, en considérant les contraintes de mise en relation présentées précédemment (association mono-hypothèse), ces méthodes procurent toutes une seule association pour chaque objet. Il existe des situations ambiguës où plusieurs associations peuvent être envisagées pour un même objet. C'est pourquoi, [Mourllion *et al.*, 2005] proposent une approche MHT où les associations ayant une probabilité pignistique supérieure à un seuil prédéfini sont sélectionnées. Tous les algorithmes cités précédemment exploitent les deux matrices pignistiques pour déterminer les associations. Cependant, résoudre le problème d'association selon deux points de vues ($X \rightarrow Y$ et $X \leftarrow Y$) va forcément aboutir à des associations contradictoires, i.e. des relations piste-cible et cible-piste incompatibles. Face à cet inconvénient, [Lauffenburger *et al.*, 2013] proposent la fusion des deux matrices $BetP_{i..}$ et $BetP_{..j}$ en une seule dénommée matrice pignistique duale (cf. Section 4.5.6).

Dans cette section, nous présentons un état de l'art des récents algorithmes de prise de décision susceptible d'être intégrés à l'algorithme SMR.

4.5.1 Probabilité pignistique jointe

Une fois les probabilités pignistiques $BetP_{i..}$ calculées pour $i = 1 \dots n$ avec redistribution du conflit (cf. (4.28)), la décision choisie est celle qui maximise la probabilité pignistique jointe $BetP_{\prod_{i=1}^n}$ définie par [Mercier *et al.*, 2011] :

$$BetP_{\prod_{i=1}^n} = BetP_{i..}(Y_a) \times BetP_{i..}(Y_b) \times \dots \times BetP_{i..}(Y_l) \quad (4.38)$$

avec $1 \leq a, b, \dots, l \leq m$ et $a \neq b \neq \dots \neq l$.

À partir des $(m + 1)^n$ solutions possibles, l'algorithme de *Probabilité Pignistique Jointe* (PPJ) supprime, dans un premier temps, celles qui ne vérifient pas les contraintes d'association (cf. Algorithme 5). Ensuite, la probabilité pignistique jointe (cf. (4.38)) est calculée pour les solutions retenues. La prise de décision consiste à sélectionner la solution (ensemble d'associations $X \rightarrow Y$) ayant la plus grande $BetP_{\prod_{i=1}^n}$. Le même algorithme est appliqué sur les $BetP_{..j}$ avec $j = 1 \dots m$, pour prendre une décision du point de vue des m pistes. Il est utile de préciser que l'algorithme PPJ satisfait un critère global, représenté par la probabilité pignistique jointe.

4.5.2 Probabilité pignistique jointe classée

[Mercier *et al.*, 2011] ont constaté que lorsqu'un certain nombre de probabilités pignistiques sont nulles, il n'est pas nécessaire de calculer la probabilité jointe pour toutes les solutions possibles

Algorithme 5 Probabilité Pignistique Jointe (PPJ)**Donnée :** Matrice pignistique $BetP_{i..}$,**Résultat :** “Meilleures” associations $\{Associations\}$ $\{Associations\} \leftarrow \{\}$ Rechercher l'ensemble des $(m + 1)^n$ solutions possibles**Pour** chaque $solution(l)$ possible **faire****Si** $solution(l)$ vérifie les contraintes d'association **alors**Calculer la probabilité pignistique jointe, $BetP_{\prod_{i=1}^n}(l)$ **Sinon**Supprimer $solution(l)$ **Fin Si****Fin Pour** $d_{max} \leftarrow \arg \max_l (BetP_{\prod_{i=1}^n}(l))$ $\{Associations\} \leftarrow solution(d_{max})$

ce qui permet de réduire le temps de calcul. Pour cela, les auteurs proposent de classer par ordre décroissant du nombre de probabilités nulles les lignes de la matrice $BetP_{i..}$. Ce classement permet de déterminer facilement les solutions valides, celles ne contenant pas d'association avec une ou plusieurs probabilités pignistiques nulles. Les solutions valides doivent aussi vérifier les contraintes d'association. Par la suite, la prise de décision est similaire à celle de l'algorithme PPJ (cf. Algorithme 6) basée sur le maximum de $BetP_{\prod_{i=1}^n}$.

Algorithme 6 Probabilité Pignistique Jointe Classée (PPJC)**Donnée :** Matrice pignistique $BetP_{i..}$,**Résultat :** “Meilleures” associations $\{Associations\}$ $\{Associations\} \leftarrow \{\}$ Classer les lignes de $BetP_{i..}$, en fonction du nombre décroissant des probabilités pignistiques nulles**Pour** chaque $solution(l)$ valide **faire****Si** $solution(l)$ vérifie les contraintes d'association **alors**Calculer la probabilité pignistique jointe, $BetP_{\prod_{i=1}^n}(l)$ **Sinon**Supprimer $solution(l)$ **Fin Si****Fin Pour** $l_{max} \leftarrow \arg \max_l (BetP_{\prod_{i=1}^n}(l))$ $\{Associations\} \leftarrow solution(l_{max})$

Le fait d'optimiser la prise de décision selon un critère global ne garantit pas l'exclusion d'associations fausses ou douteuses. En effet, la meilleure solution déterminée peut contenir des associations erronées conduisant à la génération de pistes fantômes ou à la perte de pistes existantes. La figure 4.6 illustre un exemple de scénario avec 3 objets perçus ($\{X_1, X_2, X_3\}$) et 2 objets connus ($\{Y_1, Y_2\}$). La matrice pignistique correspondante à l'association $X \rightarrow Y$ est donnée par la table 4.7. À partir de cette dernière, le maximum de $BetP_{\prod_{i=1}^3}$ est égal à $0.87 \times 0.35 \times 0.42 = 0.128$ et correspond aux

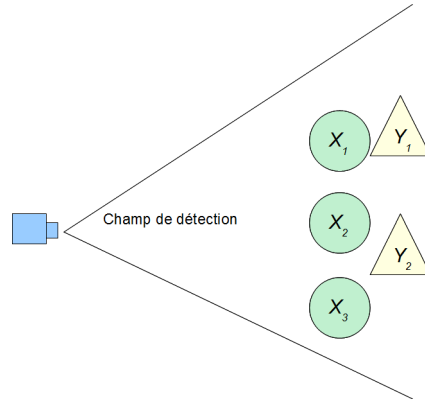


FIGURE 4.6 – Exemple de scénario avec 3 objets perçus et 2 objets connus.

TABLE 4.7 – Matrice pignistique $BetP_{i..}$ du scénario illustré par la figure 4.6

$BetP_{i..}$	Y_1	Y_2	Y_*
X_1	0.87	0.13	0.10
X_2	0.35	0.35	0.30
X_3	0.10	0.48	0.42

associations suivantes :

$$\begin{array}{ccc}
 X_1 & X_2 & X_3 \\
 \downarrow & \downarrow & \downarrow \\
 Y_1 & Y_2 & Y_*
 \end{array}$$

Cependant, les deux associations : $X_2 \rightarrow Y_2$ et $X_3 \rightarrow Y_*$ sont douteuses puisqu'elles ne sont pas optimales localement. En effet, d'après la matrice $BetP_{i..}$, il est tout à fait possible d'associer X_3 à Y_2 et de considérer X_2 comme un nouvel objet, ce qui reflète une situation ambiguë. [Mourllion *et al.*, 2005] définissent l'ambiguïté comme un problème d'association intra-matriciel où sur une même ligne de la matrice $BetP_{i..}$ ($BetP_{.,j}$) il peut y avoir plusieurs probabilités pignistiques maximales. À partir de ce constat, [Daniel et Lauffenburger, 2012] ont proposé une optimisation locale au problème d'association.

4.5.3 Probabilité pignistique locale

L'algorithme de *Probabilité Pignistique Locale* (PPL) sélectionne simplement de manière successive les n (m) associations $X \rightarrow Y$ ($X \leftarrow Y$) maximisant la probabilité pignistique $BetP_{i..}$ ($BetP_{.,j}$) [Daniel et Lauffenburger, 2012]. Lorsqu'une association est retenue, on procède à la suppression des pistes et cibles correspondantes dans la matrice $BetP_{i..}$ ($BetP_{.,j}$) afin de respecter les contraintes d'association (cf. Algorithme 7). Contrairement aux PPJ et PPJC, PPL se base sur des probabilités pignistiques non-normalisées, laissant apparaître le conflit (cf. (4.37)). Le principal avantage de PPL est sa faible complexité par rapport aux précédents algorithmes. En effet, cette approche ne nécessite pas de calcul supplémentaire contrairement aux autres approches crédales.

De ce fait, le PPL est plus adapté aux applications dont les contraintes temporelles sont fortes telles que les systèmes TSR [Boumediene et al., 2014a]. Par ailleurs, cet optimiseur local demeure moins conservatif et réduit les associations contradictoires et ambigus.

Algorithme 7 Probabilité Pignistique Locale (PPL)

Donnée : Matrice pignistique $BetP_{i..}$,

Résultat : “Meilleures” associations $\{Associations\}$

$\{Associations\} \leftarrow \{\}$

Pour chaque ligne de $BetP_{i..}$, **faire**

$(i_{max}, j_{max}) = \arg \max_{(i,j)} (BetP_{i..}(i, j))$

$\{Associations\} \leftarrow (i_{max}, j_{max})$

Supprimer la ligne i_{max} et la colonne j_{max} de $BetP_{i..}$.

Fin Pour

4.5.4 Probabilité pignistique classée par gradient

Sur le même principe d’optimisation locale, [Daniel et Lauffenburger, 2012] proposent un algorithme similaire à PPL, dénommé *Probabilité Pignistique Classée par Gradient* (PPCG). Avant la prise de décision, l’algorithme classe les lignes de la matrice $BetP_{i..}$ ($BetP_{i.,j}$) par ordre décroissant de la variation de probabilité pignistique définie par le gradient suivant :

$$\max_{1 \leq j \leq m} (BetP_{i..}(i, j)) - \min_{1 \leq j \leq m} (BetP_{i..}(i, j)) \quad (4.39)$$

Le gradient révèle l’incertitude liée à la prise de décision considérée. Par exemple, une ligne ayant un gradient élevé révèle une grande confiance dans la prise de décision qui consiste à sélectionner l’association ayant la plus grande probabilité. Par contre, un gradient faible indique la présence de plusieurs associations caractérisées par des probabilités voisines, c’est-à-dire une situation ambiguë et dans ce cas l’association choisie est incertaine.

Algorithme 8 Probabilité Pignistique Classée par Gradient (PPCG)

Donnée : Matrice pignistique $BetP_{i..}$,

Résultat : “Meilleures” associations $\{Associations\}$

$\{Associations\} \leftarrow \{\}$

Classer les lignes de $BetP_{i..}$ en fonction de $\max_{1 \leq j \leq M} (BetP_{i..}(i, j)) - \min_{1 \leq j \leq M} (BetP_{i..}(i, j))$

Pour chaque ligne de $BetP_{i..}$, **faire**

$(j_{max}) = \arg \max_j (BetP_{i..}(i, j))$

$\{Associations\} \leftarrow (i, j_{max})$

Supprimer la ligne i et la colonne j_{max} de $BetP_{i..}$.

Fin Pour

4.5.5 Seuil

Quelle soit basée sur une optimisation globale ou locale, la prise de décision n'est pas toujours une tâche simple. En effet, il est possible d'avoir des situations ambiguës où plusieurs associations locales peuvent être envisagées pour une cible ou une piste donnée. Cette ambiguïté est aussi observée par la présence de conflit ($BetP_{i..}(\emptyset) \neq 0$) et peut conduire à des associations multiples ainsi qu'à des non-associations. Face à ce problème, [Mourllion *et al.*, 2005] proposent la méthode *Seuil* qui prend en compte le conflit dans la décision. Un seuil pondéré s_i est défini pour chaque ligne de la matrice $BetP_{i..}$ ($BetP_{..j}$) :

$$s_i = k_s(1 - BetP_{i..}(\emptyset)) \quad (4.40)$$

avec k_s le poids de la pondération.

La méthode sélectionne, dans chaque ligne de $BetP_{i..}$ ($BetP_{..j}$), les associations dont la probabilité pignistique est supérieure au seuil s_i correspondant (cf. Algorithme 9). Cette approche ne respecte pas les contraintes d'association. En effet, si le conflit est trop élevé, aucune décision n'est prise. De plus, la méthode peut générer des associations multiples si plusieurs probabilités pignistiques sont supérieures au seuil considéré.

Algorithme 9 Seuil

Donnée : Matrice pignistique $BetP_{i..}$, et k_s

Résultat : “Meilleures” associations $\{Associations\}$

$\{Associations\} \leftarrow \{\}$

Pour chaque ligne de $BetP_{i..}$ **faire**

$s_i = k_s(1 - BetP_{i..}(\emptyset))$

Pour chaque colonne j de $BetP_{i..}$ **faire**

Si $BetP_{i..}(i, j) \geq s_i$ **alors**

$\{Associations\} \leftarrow (i, j)$

Fin Si

Fin Pour

Fin Pour

Pour gérer au mieux les apparitions et disparitions d'objets, la théorie des croyances traite le problème d'association selon deux points de vues : $X \longrightarrow Y$ et $X \longleftarrow Y$. Cette résolution bidirectionnelle nécessite la prise de décision (à l'aide de l'un des algorithmes présentés précédemment) sur les deux matrices pignistiques : $BetP_{i..}$ et $BetP_{..j}$. Néanmoins, cette approche bidirectionnelle peut aboutir à des associations contradictoires dues à la présence de conflit. [Mourllion *et al.*, 2005] définissent cette contradiction comme un problème inter-matriciel puisqu'elle se produit lorsque les deux matrices pignistiques indiquent deux ou plusieurs associations différentes pour la même cible ou la même piste. [Lauffenburger *et al.*, 2013] remédient à ce problème de contradiction par la fusion des deux matrices $BetP_{i..}$ et $BetP_{..j}$ en une seule dénommée *matrice pignistique duale* sur laquelle se base la méthode de décision duale.

4.5.6 Décision duale

[Lauffenburger *et al.*, 2013] partent du constat qu'une décision basée sur 2 critères distincts ($BetP_{i,\cdot}$ et $BetP_{\cdot,j}$) va irrémédiablement aboutir à des associations contradictoires et/ou suspicieuses. La décision doit donc être prise à partir d'une seule matrice pignistique au lieu de deux. C'est pourquoi, ils proposent la matrice pignistique duale M_{YX}^{Duale} définie par le produit de Hadamard des deux matrices $BetP_{i,\cdot}$ et $BetP_{\cdot,j}$ après une transposition de la première (cf. (4.41)). Les probabilités pignistiques des hypothèses : X_* et Y_* , ainsi que la masse conflictuelle (\emptyset) ne sont pas prises en considération dans (4.41). Le produit terme à terme diminue la probabilité des associations contradictoires. En effet, une association certaine est caractérisée par des probabilités pignistiques élevées dans les deux matrices (même association dans les deux sens) et leur produit le sera donc également. Dans le cas contraire, pour deux associations contradictoires, la probabilité duale obtenue est faible ce qui ne permet pas sa sélection.

$$M_{YX}^{Duale} = (BetP_{i,\cdot})^T \circledast BetP_{\cdot,j} \quad (4.41)$$

En plus, de la probabilité duale, la méthode de la décision duale de [Lauffenburger *et al.*, 2013] tient compte du conflit contrairement aux approches globales et locales décrites précédemment. Pour cela, l'information concernant le conflit est intégrée à la matrice M_{YX}^{Duale} par l'ajout d'une ligne et d'une colonne supplémentaires. Tout d'abord, une matrice duale du conflit M_{YX}^{Conf} est calculée en combinant les valeurs pignistiques du conflit $BetP_{i,\cdot}()$ et $BetP_{\cdot,j}()$ (cf. (4.42)). Chaque élément (j, i) de M_{YX}^{Conf} représente le conflit généré par l'association $Y_j \longleftrightarrow X_i$ avec les autres associations $Y_j \longleftrightarrow X_k$ où $k = 1, \dots, n$ et $k \neq i$. Par conséquent, le conflit dual lié à l'association de chaque piste (Y_j) et de chaque cible (X_i) est respectivement obtenu en additionnant les lignes et les colonnes de la matrice M_{YX}^{Conf} , ce qui permet d'obtenir deux vecteurs V_{YX}^{Conf} et V_{XY}^{Conf} (cf. (4.43)). La table 4.8 illustre la matrice pignistique duale obtenue à partir de laquelle les décisions sont prises.

$$M_{YX}^{conf} = BetP_{\cdot,j}() \cdot (BetP_{i,\cdot}())^T \quad (4.42)$$

$$\begin{cases} V_{YX}^{Conf}(j) = \sum_{i=1}^n M_{YX}^{Conf}(j, i) \\ V_{XY}^{Conf}(i) = \sum_{j=1}^m M_{XY}^{Conf}(j, i) \end{cases} \quad (4.43)$$

[Lauffenburger *et al.*, 2013] proposent l'algorithme de décision dual pour générer les associations $X \longleftrightarrow Y$ (cf. Algorithme 10) à partir de la matrice duale ainsi obtenue. Dans un premier temps, la probabilité pignistique maximum est détectée sur chaque ligne j de la matrice. Ensuite, la décision d'association de Y_j à $X_{i_{max}}$ est prise en fonction de la valeur du conflit dual $V_{YX}^{Conf}(j)$. Si le conflit est supérieur à la probabilité pignistique ($M_{YX}^{Dual}(j, i_{max})$), cela révèle la présence d'associations potentiellement contradictoires. Dans ce cas, Y_j n'est pas associé ($NA = \text{Non Associé}$). Dans le cas contraire, ($M_{YX}^{Dual}(j, i_{max}) > V_{YX}^{Conf}(j)$), la probabilité duale est comparée à un *Seuil* pour

TABLE 4.8 – Matrice pignistique duale

$M_{YX}^{Dual}(\cdot)$	X_1	...	X_n	V_{YX}^{Conf}
Y_1	$M_{YX}^{Dual}(1, 1)$...	$M_{YX}^{Dual}(1, n)$	$V_{YX}^{Conf}(1)$
Y_2	$M_{YX}^{Dual}(2, 1)$...	$M_{YX}^{Dual}(2, n)$	$V_{YX}^{Conf}(2)$
\vdots	\vdots	\vdots	\vdots	\vdots
Y_m	$M_{YX}^{Dual}(m, 1)$...	$M_{YX}^{Dual}(m, n)$	$V_{YX}^{Conf}(m)$
V_{XY}^{Conf}	$V_{XY}^{Conf}(1)$...	$V_{XY}^{Conf}(n)$	0

déterminer si Y_j a disparue ($Y_j \longleftrightarrow X_*$) ou pas ($Y_j \longleftrightarrow X_{i_{max}}$). Finalement, les cibles non associées à l'issue de ce processus peuvent être interprétées comme de nouvelles apparitions selon la valeur du conflit dual $V_{XY}^{Conf}(i)$.

Algorithme 10 Décision Duale

Donnée : Matrice duale M_{YX}^{Dual} et *Seuil*

Résultat : “Meilleures” associations $\{Associations\}$

$\{Associations\} \leftarrow \{\}$

Pour chaque ligne de $M_{YX}^{Dual}(j, i)$ **faire**

$i_{max} = \arg \max_i M_{YX}^{Dual}(j, i)$

Si $V_{YX}^{Conf}(j) > M_{YX}^{Dual}(j, i_{max})$ **alors**

$\{Associations\} \leftarrow (Y_j, NA)$

Sinon

Si $M_{YX}^{Dual}(j, i_{max}) \geq \text{Seuil}$ **alors**

$\{Associations\} \leftarrow (Y_j, X_i)$

Sinon

$\{Associations\} \leftarrow (Y_j, *)$

Fin Si

Fin Si

Fin Pour

Pour X_i non associé **faire**

Si $V_{XY}^{Conf}(i) < \text{Seuil}$ **alors**

$\{Associations\} \leftarrow (X_i, *)$

Sinon

$\{Associations\} \leftarrow (X_i, NA)$

Fin Si

Fin Pour

Le principal avantage de la décision duale est le fait qu'elle dépend d'une seule mesure de décision éliminant le problème d'associations contradictoires des méthodes globales et locales. De plus, tenir compte du conflit lors de la prise de décision rend l'approche duale prudente. En effet, l'algorithme dual reste indécis dans le cas où le conflit est supérieur aux probabilités d'association ce qui évite les associations suspicieuses. Cependant, cette stratégie n'est pas adaptée pour les algorithmes de suivi mono-hypothèse où une décision d'association doit être prise à chaque instant pour chaque piste ou cible. L'exemple réel de la figure 4.7 illustre très bien ce cas de figure. En effet, à partir de la matrice duale correspondante illustrée par la table 4.9, l'algorithme n'aboutit qu'à des décisions

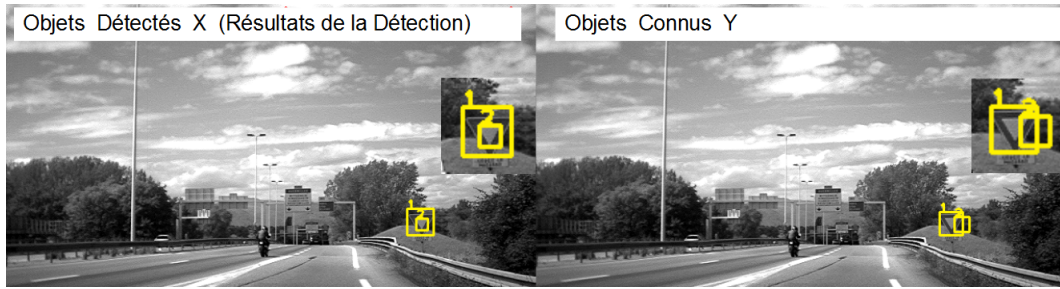


FIGURE 4.7 – Exemple réel d'association avec 2 objets perçus et 2 objets connus.

TABLE 4.9 – Matrice duale correspondante à la scène illustrée par la figure 4.7

M_{YX}^{Dual}	X_1	X_2	V_{YX}^{Conf}
Y_1	0.04	0.02	0.97
Y_2	0.01	0.03	0.61
V_{XY}^{Conf}	0.75	0.83	0

de non-association (NA) et dans ce cas il est possible d'envisager une propagation de plusieurs associations pour un même objets (i.e. $Y_1 \rightarrow X_1$ et $Y_1 \rightarrow X_2$) :

$$\begin{array}{cccccc}
 Y_1 & Y_2 & X_1 & X_2 & & Y_1 & Y_2 \\
 \updownarrow & \updownarrow & \updownarrow & \updownarrow & \Rightarrow & \swarrow \searrow & \swarrow \searrow \\
 NA & NA & NA & NA & & X_1 X_2 & X_1 X_2
 \end{array}$$

4.5.7 Conclusion

Dans cette section, nous avons présenté les principales approches de prise de décision. La majorité se basent sur les deux matrices pignistiques $BetP_{i..}$ et $BetP_{.,j}$ afin de résoudre le problème d'association selon deux points de vues : $X \rightarrow Y$ et $X \leftarrow Y$. Les approches globales sélectionnent la “meilleure” solution selon un critère global (cf. Sections 4.5.1 et 4.5.2). Cependant, même si la décision est globalement la meilleure, elle peut contenir des associations non-optimales. C'est pourquoi, d'autres approches se basent sur un critère local (cf. Sections 4.5.3 et 4.5.4). Ces approches s'appuient directement sur les matrices $BetP_{i..}$ et $BetP_{.,j}$ sans calcul supplémentaire et sont donc moins gourmandes ce qui suscite leur intérêt pour les applications temps réel telles que les TSR.

Par ailleurs, d'autres auteurs ont proposé des approches multi-hypothèses crédales (cf. Section 4.5.5). Ces dernières peuvent générer des associations multiples ainsi que des non-associations ce qui est tout à fait adapté au suivi multi-hypothèses. Cependant, cela ne présentent aucun intérêt pour les applications mono-hypothèse telles que les TSR où le maintien de plusieurs associations pour un même panneau n'apporte rien de plus au système et ne fait qu'augmenter sa complexité de calcul.

Le fait de dépendre de deux mesures de décision, $BetP_{i,\cdot}$ et $BetP_{\cdot,j}$, risque de générer des associations contradictoires. C'est pourquoi, d'autres auteurs ont proposé la fusion de ces deux mesures en une seule tout en tenant compte de la présence du conflit (cf. Section 4.5.6). Ainsi, l'approche paraît avantageuse pour les applications TSR puisqu'elle gère les situations ambiguës et contradictoires. Néanmoins, en présence de fort conflit, l'approche est prudente et ne prend pas de décision (non-association) ce qui risque de nuire à la gestion des apparitions et disparitions des panneaux.

À partir de ce constat, nous avons adopté une approche locale, en particulier l'algorithme PPL, pour l'étape d'association de données de l'algorithme SMR. Ce choix est justifié par la faible complexité de cette approche, ainsi qu'à sa capacité à éviter les associations douteuses qui risquent de nuire au processus de suivi. Cependant, il est tout à fait possible d'envisager l'utilisation de d'autres approches. Dans la section suivante, nous présentons un benchmark des algorithmes de décision sur des scénarios simulés ainsi que sur des données réelles concernant le suivi de panneaux.

4.6 Evaluation des algorithmes d'association

Dans un premier temps, afin de justifier davantage notre choix pour une approche d'association locale pour l'algorithme SMR, nous discutons une évaluation des algorithmes de prise de décision présentée par [Daniel et Lauffenburger, 2012]. Cette évaluation se base sur un scénario simulé pour comparer les performances des algorithmes PPL, PPCG, PPJ et PPJC. Par la suite, nous comparons l'algorithme SMR avec d'autres solutions et nous évaluons son apport au système de détection. Finalement, nous étendons l'évaluation sur des données réelles et nous mesurons l'influence du choix de l'algorithme de prise de décision sur les performances de l'algorithme SMR. Les résultats présentés dans cette section ont fait l'objet de deux communications et d'une publication internationales [Boumediene et al., 2014a], [Lauffenburger et al., 2014], et [Boumediene et al., 2014b].

4.6.1 Résultats de simulation

L'évaluation des algorithmes de prise de décision présentée par [Daniel et Lauffenburger, 2012] consiste à analyser et comparer leurs résultats d'association sur un exemple de la littérature (l'exemple n°5 dans [Mercier et al., 2011]). Ce scénario simulé considère 4 objets connus ($Y_j, j = 1, \dots, 4$) et 3 objets perçus ($X_i, i = 1, \dots, 3$). Les cadres de discernements correspondant sont définis par :

$$\begin{cases} \Theta_{i,\cdot} = & Y_1, Y_2, Y_3, Y_4, Y_* \\ \Theta_{\cdot,j} = & X_1, X_2, X_3, X_* \end{cases} \quad (4.44)$$

Les jeux de masses initiaux sont donnés par la table 4.10. À ce niveau, on constate déjà quelques situations conflictuelles entre les sources spécialisées ($Src_{i,j}$). En effet, la source $Src_{1,1}$ accorde une

TABLE 4.10 – Scénario simulé de [Mercier *et al.*, 2011] : 3 objets perçus vs. 4 objets connus

$$\begin{array}{cccc}
 Src_{1,1} \begin{cases} m_1^{\Theta_{1..}}(Y_1) = 0.80 \\ m_1^{\Theta_{1..}}(\overline{Y_1}) = 0.00 \\ m_1^{\Theta_{1..}}(\Theta) = 0.20 \end{cases} & Src_{1,2} \begin{cases} m_2^{\Theta_{1..}}(Y_2) = 0.00 \\ m_2^{\Theta_{1..}}(\overline{Y_2}) = 0.99 \\ m_2^{\Theta_{1..}}(\Theta) = 0.01 \end{cases} & Src_{1,3} \begin{cases} m_3^{\Theta_{1..}}(Y_3) = 0.00 \\ m_3^{\Theta_{1..}}(\overline{Y_3}) = 0.97 \\ m_3^{\Theta_{1..}}(\Theta) = 0.03 \end{cases} & Src_{1,4} \begin{cases} m_4^{\Theta_{1..}}(Y_4) = 0.00 \\ m_4^{\Theta_{1..}}(\overline{Y_4}) = 0.99 \\ m_4^{\Theta_{1..}}(\Theta) = 0.01 \end{cases} \\
 Src_{2,1} \begin{cases} m_1^{\Theta_{2..}}(Y_1) = 0.57 \\ m_1^{\Theta_{2..}}(\overline{Y_1}) = 0.00 \\ m_1^{\Theta_{2..}}(\Theta) = 0.43 \end{cases} & Src_{2,2} \begin{cases} m_2^{\Theta_{2..}}(Y_2) = 0.57 \\ m_2^{\Theta_{2..}}(\overline{Y_2}) = 0.00 \\ m_2^{\Theta_{2..}}(\Theta) = 0.43 \end{cases} & Src_{2,3} \begin{cases} m_3^{\Theta_{2..}}(Y_3) = 0.00 \\ m_3^{\Theta_{2..}}(\overline{Y_3}) = 0.52 \\ m_3^{\Theta_{2..}}(\Theta) = 0.48 \end{cases} & Src_{2,4} \begin{cases} m_4^{\Theta_{2..}}(Y_4) = 0.00 \\ m_4^{\Theta_{2..}}(\overline{Y_4}) = 0.99 \\ m_4^{\Theta_{2..}}(\Theta) = 0.01 \end{cases} \\
 Src_{3,1} \begin{cases} m_1^{\Theta_{3..}}(Y_1) = 0.00 \\ m_1^{\Theta_{3..}}(\overline{Y_1}) = 0.99 \\ m_1^{\Theta_{3..}}(\Theta) = 0.01 \end{cases} & Src_{3,2} \begin{cases} m_2^{\Theta_{3..}}(Y_2) = 0.61 \\ m_2^{\Theta_{3..}}(\overline{Y_2}) = 0.00 \\ m_2^{\Theta_{3..}}(\Theta) = 0.39 \end{cases} & Src_{3,3} \begin{cases} m_3^{\Theta_{3..}}(Y_3) = 0.00 \\ m_3^{\Theta_{3..}}(\overline{Y_3}) = 0.52 \\ m_3^{\Theta_{3..}}(\Theta) = 0.48 \end{cases} & Src_{3,4} \begin{cases} m_4^{\Theta_{3..}}(Y_4) = 0.00 \\ m_4^{\Theta_{3..}}(\overline{Y_4}) = 0.99 \\ m_4^{\Theta_{3..}}(\Theta) = 0.01 \end{cases}
 \end{array}$$

confiance élevée (0.80) à l'association de X_1 avec Y_1 , en même temps la source $Src_{2,1}$ manifeste une possibilité d'association entre X_2 et Y_1 (0.57). L'association de X_2 est aussi sujet à ambiguïté, $Src_{2,1}$ et $Src_{2,2}$ manifestant une confiance égale pour l'association avec Y_1 et Y_2 . Une autre situation ambiguë est générée par les sources $Src_{2,2}$ et $Src_{3,2}$ où deux associations sont possibles pour Y_2 : avec X_2 (0.57) ou avec X_3 (0.61). Finalement, il semble que Y_3 et Y_4 disparaîtront puisque toutes les associations sont rejetées.

Après la combinaison des jeux de masses, les matrices pignistiques des associations $X \rightarrow Y$ et $X \leftarrow Y$ sont données par les tables 4.11 et 4.12. Il est utile de préciser que les situations ambiguës et conflictuelles ne disparaissent pas forcément après la combinaison des masses. En effet, à partir de la matrice $BetP_{i..}$, l'association de X_2 est toujours ambiguë puisque deux probabilités pignistiques égales sont obtenues pour l'association de X_2 avec Y_1 (0.30) et Y_2 (0.30). À partir de la matrice $BetP_{.j}$, on observe une ambiguïté pour l'association de Y_2 et la présence de conflit considérable (0.35). Par conséquent, les associations les plus probables selon les deux points de vues : $X \rightarrow Y$ et $X \leftarrow Y$ sont :

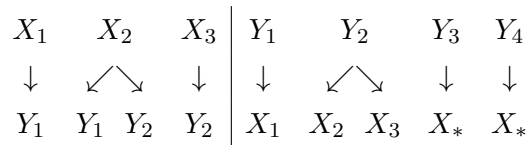


TABLE 4.11 – Probabilités pignistiques des associations $X \rightarrow Y$

$BetP_{i..}(\cdot)$	Y_1	Y_2	Y_3	Y_4	Y_*	\emptyset
X_1	0.90	0.00	0.00	0.00	0.10	0.00
X_2	0.30	0.30	0.02	0.00	0.05	0.33
X_3	0.00	0.77	0.06	0.00	0.16	0.00

TABLE 4.12 – Probabilités pignistiques des associations $X \leftarrow Y$

$BetP_{.j}(\cdot)$	X_1	X_2	X_3	X_*	\emptyset
Y_1	0.37	0.14	0.00	0.03	0.46
Y_2	0.00	0.28	0.32	0.06	0.35
Y_3	0.01	0.20	0.20	0.59	0.00
Y_4	0.00	0.00	0.00	0.99	0.00

TABLE 4.13 – Résultats de la prise de décision

Algorithme	$X_i \longrightarrow Y_j$	$X_i \longleftarrow Y_j$	Temps moyen de calcul	Remarques
<i>PPJ</i>	$X_1 \longrightarrow Y_1$ $X_2 \longrightarrow Y_2$ $X_3 \longrightarrow Y_*$	$X_1 \longleftarrow Y_1$ $X_3 \longleftarrow Y_2$ $X_* \longleftarrow Y_3$ $X_* \longleftarrow Y_4$	25 ms	Disparition de Y_3 et Y_4 Contradictions concernant X_3 et Y_2
<i>PPJC</i>	$X_1 \longrightarrow Y_1$ $X_2 \longrightarrow Y_2$ $X_3 \longrightarrow Y_*$	$X_1 \longleftarrow Y_1$ $X_3 \longleftarrow Y_2$ $X_* \longleftarrow Y_3$ $X_* \longleftarrow Y_4$	20 ms	Disparition de Y_3 et Y_4 Contradictions concernant X_3 et Y_2
<i>PPL</i>	$X_1 \longrightarrow Y_1$ $X_2 \longrightarrow Y_*$ $X_3 \longrightarrow Y_2$	$X_1 \longleftarrow Y_1$ $X_3 \longleftarrow Y_2$ $X_* \longleftarrow Y_3$ $X_* \longleftarrow Y_4$	7 ms	Apparition de X_2 Disparition de Y_3 et Y_4
<i>PPCG</i>	$X_1 \longrightarrow Y_1$ $X_2 \longrightarrow Y_*$ $X_3 \longrightarrow Y_2$	$X_1 \longleftarrow Y_1$ $X_3 \longleftarrow Y_2$ $X_* \longleftarrow Y_3$ $X_* \longleftarrow Y_4$	7 ms	Apparition de X_2 Disparition de Y_3 et Y_4

[Daniel et Lauffenburger, 2012] ont appliqué les algorithmes : PPJ, PPJC, PPL, et PPCG sur ce scénario simulé et ont comparé les résultats obtenus. La table 4.13 résume ces derniers ainsi que le temps de calcul. Les algorithmes PPJ et PPJC (approches globales) procurent des résultats similaires à l'exception du temps de traitement où PPJC réduit légèrement la complexité du calcul en éliminant les solutions contenant des associations nulles. Cependant, les deux algorithmes génèrent des associations contradictoires. En effet, si X_1 est associé avec Y_1 dans les deux sens ($X_1 \longleftrightarrow Y_1$), Y_2 est associé avec X_2 et X_3 respectivement pour l'association $X \longrightarrow Y$ et $X \longleftarrow Y$. Quant aux algorithmes PPL et PPCG, ils ont déterminé, pour ce scénario, des résultats concordants (associations similaires selon $X \longrightarrow Y$ et *vice versa*), ce qui constitue un avantage lors du suivi d'objets. De plus, les algorithmes PPL et PPCG se caractérisent par une faible complexité³ par rapport à PPJ et PPJC. En effet, d'après la simulation, les algorithmes PPL et PPCG (7 ms) sont trois fois plus rapides que les algorithmes PPJ et PPJC (20 ms), ce qui les rend plus adaptés aux applications dont les contraintes temporelles sont fortes comme les systèmes TSR.

Il est utile de préciser que pour cette comparaison, le conflit n'a pas été redistribué lors de la transformation pignistique. L'algorithme PPCG détecte ces situations conflictuelles par le calcul de la variation des probabilités pignistiques. La dernière colonne des matrices pignistiques (cf. Table 4.14) représente le gradient correspondant à chaque ligne. En effet, les décisions certaines comme l'apparition de Y_4 et l'association de X_1 avec Y_1 présentent un gradient élevé, alors que l'association de Y_2 et de X_2 , caractérisées par la présence de conflit, manifeste un gradient faible.

Par le biais de cette simulation, [Daniel et Lauffenburger, 2012] ont montré les limites des algorithmes de prise de décision en présence du conflit. D'après les résultats obtenus, les approches locales semblent plus robustes *vis-a-vis* du conflit et génèrent moins de décisions contradictoires contrairement aux approches globales. En effet, lors de la prise de décision, ces dernières se basent

3. Le temps moyen dans la table 4.13 est obtenu après l'exécution de 20 simulations sous Matlab, sur un processeur Core 2 Quad computer (2.33 GHz) avec 4 Go de RAM.

sur le produit des probabilités définies sur $[0, 1]$ et tend vers zéro proportionnellement au nombre d'associations considérées. De plus, les approches locales sont très attractives pour les applications temps réel vu leur faible complexité.

TABLE 4.14 – Probabilités pignistiques avec la classification du gradient

$BetP_{i, \cdot}(\cdot)$	Y_1	Y_2	Y_3	Y_4	Y_*	\emptyset	Gradient
X_1	0.90	0.00	0.00	0.00	0.10	0.00	0.90
X_3	0.00	0.77	0.06	0.00	0.16	0.00	0.77
X_2	0.30	0.30	0.02	0.00	0.05	0.33	0.30

$BetP_{\cdot, j}(\cdot)$	X_1	X_2	X_3	X_*	\emptyset	Gradient
Y_4	0.00	0.00	0.00	0.99	0.00	0.99
Y_3	0.01	0.20	0.20	0.59	0.00	0.58
Y_1	0.37	0.14	0.00	0.03	0.46	0.37
Y_2	0.00	0.28	0.32	0.06	0.35	0.32

4.6.2 Suivi de panneaux par association crédale : résultats expérimentaux

Après avoir démontré le potentiel des approches locales pour les applications temps réel, nous avons intégré l'algorithme PPL à l'étape d'association de données du SMR. Les résultats expérimentaux de validation ont été publiés dans la revue IEEE Transactions on Intelligent Transportation Systems [Boumediene *et al.*, 2014b]. Dans cette section, nous présentons les performances de l'algorithme SMR pour la réduction de FP et nous évaluons l'impact du choix de l'algorithme de prise de décision sur les performances d'un TSR [Lauffenburger *et al.*, 2014]. Les résultats, présentés ici, ont été obtenus durant des tests menés avec un véhicule instrumenté du laboratoire MIPS (cf. Section 3.3.1).

4.6.2.1 Configuration du système

Le système complet (détention et suivi temporel) est implémenté en C++ avec la bibliothèque OpenCV 2.4. Il est exécuté sur un processeur Intel core *i7 2.20 GHz* avec 8 GB de mémoire vive, sous Windows Seven 64 bits. Dans ces conditions, la répartition du temps de calcul du système temps réel complet est autour de 98% pour la détection de panneaux et seulement 2% pour leur suivi à l'aide de l'algorithme SMR. Le fonctionnement du processus de suivi nécessite l'initialisation d'un ensemble de paramètres : $T_{confidence} = 0.85$ (seuil de confiance pour l'analyse des pistes ($C(j)$)), $\alpha = 0.9$ (facteur d'affaiblissement des masses), $m_{max} = 10$ (nombre maximum de FKs utilisés), $(\sigma_x, \sigma_y, \sigma_s) = (2, 2, 3)$ (écarts types du modèle dynamique).

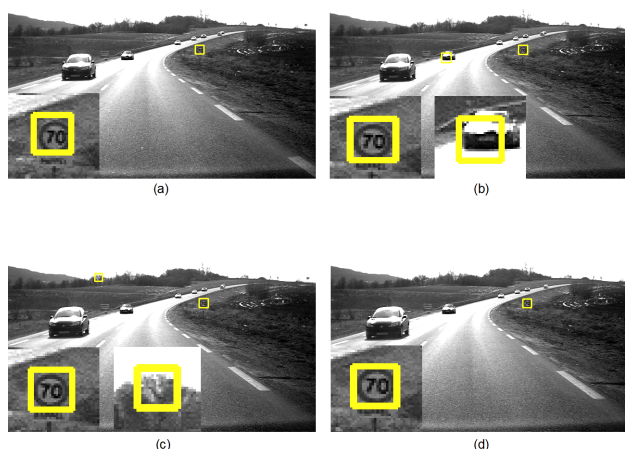


FIGURE 4.8 – Résultats de la détection sans SMR sur plusieurs images successives. (a) et (d) : détection de panneau correcte. (b) et (c) : présence de FP qui peuvent être supprimés par le SMR.

4.6.2.2 Performances de l’algorithme SMR

Le principal intérêt de notre approche de suivi est la réduction de FP. La figure 4.8 illustre un exemple de résultat brut issu de la détection, sans intégration de l’information temporelle. Sur cette succession d’images, des FP sont observées (cf. Figure 4.8 (b) et (c)). Toutefois, ces derniers sont caractérisés par une apparition brusque et une durée de vie courte que le SMR exploite afin de les supprimer.

Dans un premier temps, nous avons cherché à comparer notre méthode avec des méthodes récemment publiées. Pour une comparaison objective, nous avons choisi des approches reprenant la même structure, i.e. suivi-*après*-détection. Pour ce faire, nous nous sommes appuyés sur l’état de l’art proposé par [Møgelmoose *et al.*, 2012]. Les principales propriétés des algorithmes sélectionnés sont données dans la table 4.15 selon la nomenclature introduite dans [Møgelmoose *et al.*, 2012] à savoir : le type de panneaux détectés, l’aspect temps réel (> 5 images par seconde), l’invariance (notée *Inv.*) à la rotation, le type d’approche de détection (par modèle (forme géométrique) ou par apprentissage), et le type d’images utilisées : faible résolution pour les vidéos et haute résolution pour les images (prise instantanée). La table 4.16 synthétise leurs performances selon plusieurs critères : longueur des séquences d’évaluation (nombre d’images), le nombre de ROIs positives (contenant un panneau), les meilleurs taux de détection (TDPI et TDPP) et le nombre de FP obtenus, ainsi que les taux moyens dans le cas où l’approche concernée permet la détection de plusieurs types de panneaux.

Selon [Møgelmoose *et al.*, 2012], une comparaison en absolue est non pertinente du fait que différentes données et mesures sont utilisées pour évaluer chacun des algorithmes. Par exemple, dans [Deguchi *et al.*, 2011] et [Meuter *et al.*, 2011] deux mesures sont utilisées pour évaluer les FP : la Valeur Prédictive Positive ($VPP = VP/(VP+FP)$) et le Taux de Faux Positifs ($TFP = FP/\text{nombre}$

TABLE 4.15 – Principales propriétés des algorithmes suivi-*après*-détection sélectionnés [Møgelmoose *et al.*, 2012]

Référence	Type de panneaux détectés	Temps-Réel	Inv. Rotation	Modèle ou Apprentissage	Type des images de test
[Gu <i>et al.</i> , 2011]	Circulaire rouge et bleue	Oui	Oui	Les deux	Faible-res
[Deguchi <i>et al.</i> , 2011]	Circulaire rouge	Non	Oui	Apprentissage	Faible-res
[Meuter <i>et al.</i> , 2011]	Circulaire	Oui	Oui	Modèle	Faible-res
SMR	Circulaire, triangulaire, et cédez-le-passage	Oui	Oui	Modèle	Faible-res

TABLE 4.16 – Comparaison entre le SMR et des algorithmes suivi-*après*-détection [Møgelmoose *et al.*, 2012]

Référence	Longueur des séquences d'évaluation	Nombre de ROIs positives	Meilleur TDPI (%)	FP pour le meilleur TDPI	TDPI moyen	FP moyen
[Gu <i>et al.</i> , 2011]	2134 images	3298	94.03	FPPI : 3.41	-	-
[Deguchi <i>et al.</i> , 2011]	2967 images	4886	90.1	VPP : 85.6%	-	-
[Meuter <i>et al.</i> , 2011]	Images de 34 vidéos	Non-spécifié	87.12	TFP : 0.14%	-	-
SMR	18000 images	2500	89.92	FPPI : 0.16	74.19	FPPI : 0.13

d'exemples négatifs dans la base). À la différence de FPPI, VPP dépend du nombre d'images où les panneaux sont présents. Par conséquent, il est difficile de comparer l'algorithme SMR avec [Meuter *et al.*, 2011] puisque les auteurs ne donnent aucune information à propos des 34 séquences vidéo utilisées. Concernant le TFP, cette mesure est rarement utilisée pour l'évaluation de la détection puisque le nombre d'exemples négatifs, contenus dans une séquence vidéo, n'a aucune sens [Møgelmoose *et al.*, 2012]. En dépit de ces différentes mesures, l'algorithme SMR présente des performances intéressante avec un faible FPPI (0.16) pour un TDPI élevé (89.92%), contrairement à [Gu *et al.*, 2011] qui procure le meilleur TDPI (94.03%) mais avec un FPPI très élevé (3.41). Le TDPI moyen et FPPI moyen (plus précis puisqu'ils évaluent les résultats concernant plusieurs types de panneaux) mettent en évidence les bons résultats du SMR.

Pour évaluer l'apport de l'algorithme SMR au système de détection, nous avons comparé les mesures FPPI, TDPI, et TDPP sans (noté \overline{SMR}) et avec la présence du SMR. Les résultats présentés dans la table 4.17 démontrent clairement la réduction de 40% – 45% du FPPI par le SMR tout en maintenant pratiquement les mêmes taux de TDPI et TDPP [Boumediene *et al.*, 2014a], [Boumediene *et al.*, 2014b]. Cela élimine une quantité importante de données inutiles qui n'est plus transmise à l'étape de *Reconnaissance*, ce qui réduit le temps de traitement global du système TSR. La figure 4.9 présente les courbes ROC (Receiver Operating Characteristic) qui donnent plus de détails à propos de l'influence de l'algorithme SMR sur l'étape de *Détection*. Ces courbes représentent le TDPI en fonction du FPPI, en considérant différentes configurations pour le système avec et sans l'algorithme SMR. Pour la détection de panneaux circulaires, les courbes ROC illustrent clairement l'amélioration apportée par l'algorithme SMR où des TDPI équivalent à ceux du système sans SMR (\overline{SMR}) sont obtenus avec de faibles valeurs de FPPI (cf. Figure 4.9 (a)), particulièrement lorsque

TABLE 4.17 – Performances pour les séquences 1 et 2

	Séquence 1		Séquence 2	
	\overline{SMR}	SMR	\overline{SMR}	SMR
Panneau circulaire				
$TDPP$ (%)	87	87	95	95
$TDPI$ (%)	87.95	86.55(-2%)	92.81	93.3(+0.5%)
$FPPI$	0.28	0.15(-46%)	0.27	0.17(-37%)
Panneau triangulaire				
$TDPP$ (%)	91	91	100	100
$TDPI$ (%)	64.15	55.37(-14%)	69.54	61.56(-11%)
$FPPI$	0.26	0.14(-46%)	0.11	0.07(-37%)

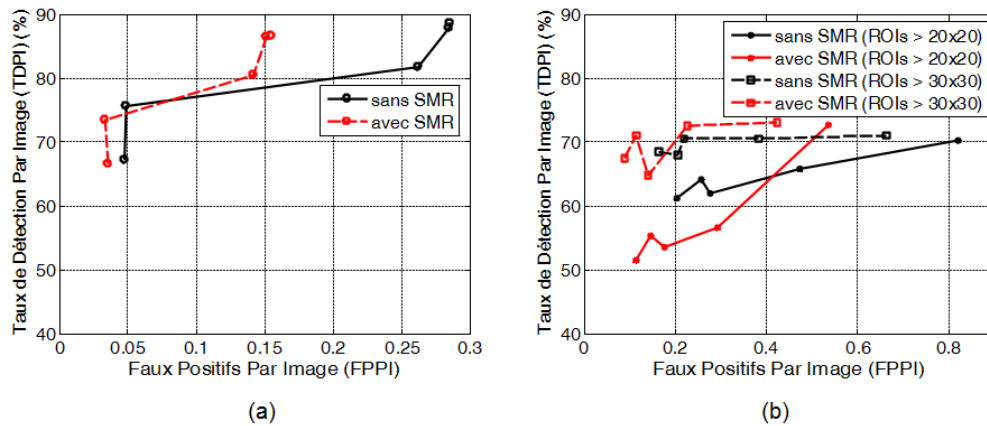


FIGURE 4.9 – Comparaison entre les courbes ROC (Receiver Operating Characteristic) du système sans et avec SMR pour les panneaux circulaires (a) et triangulaires (b).

TDPI est élevé ($> 80\%$). Pour la détection de panneaux triangulaires, la réduction de FPPI est aussi importante ($\sim 40\%$) mais en ayant en contrepartie un TDPI (71%) inférieur à celui des panneaux circulaires (87%), comme illustré dans la figure 4.9 (b). Cela est dû au fait que le détecteur de panneaux triangulaires n'est pas adapté aux panneaux de faible résolution ($< 30 \times 30$ pixels). Pour ces panneaux, la détection n'est pas stable, il y a une alternance entre détection et non-détection tout au long des images, ce qui fait que l'algorithme SMR les considère comme des FP. Pour des panneaux plus large à 30×30 pixels, l'algorithme SMR améliore les performances en maintenant le TDPI et en réduisant de manière importante le FPPI (cf. Figure 4.9 (b)).

La table 4.18 présente le temps de calcul moyen par image en millisecondes (ms), sans optimisation du code, pour le détecteur et le SMR sur les deux séquences. La détection de panneaux triangulaires et circulaires nécessite en moyenne $40 ms$ pour le traitement d'une seule image. Le temps de traitement de l'algorithme SMR dépend du nombre de panneaux détectés (n) et de panneaux suivis (m). En effet, le temps de traitement augmente avec le nombre d'associations possibles. Avec $n = m = 4$, l'algorithme SMR nécessite $\sim 0.70 ms$ pour prendre des décisions sur les associations,

TABLE 4.18 – Temps de calcul moyen

	Temps d'exécution (ms)
Détection (triangulaire et circulaire)	40
<i>SMR</i> ($n = 1, m = 1$)	0.25
<i>SMR</i> ($n = 2, m = 2$)	0.30
<i>SMR</i> ($n = 3, m = 3$)	0.56
<i>SMR</i> ($n = 4, m = 4$)	0.65
<i>SMR</i> ($n = 5, m = 5$)	1.56

TABLE 4.19 – Evaluation de l'apport des algorithmes de décision : PPCJ, PPL, PPCG, et la décision duale ainsi que de la méthode GNN sur la détection de panneaux circulaires sur la séquence vidéo 1.

	PPCJ	PPL	PPCG	Décision Duale	GNN
<i>TDPI</i> (%)	86.2	85.8	85.92	83.44	85.6
<i>FPPI</i>	0.1497	0.1483	0.1477	0.1309	0.1508

ce qui représente moins de 2% du temps de calcul global. Par conséquent, l'algorithme SMR procure de très bonnes performances en réduisant les FP avec une faible complexité de calcul.

4.6.2.3 Impact de l'association crédale sur l'algorithme SMR

Après avoir évalué les performances de l'algorithme SMR, nous avons mesuré l'impact du choix de l'algorithme de prise de décision sur le TSR. Pour cela, nous avons comparé l'influence de 4 algorithmes de prise de décision : PPCJ, PPL, PPCG, et la décision duale sur l'algorithme SMR [Lauffenburger *et al.*, 2014]. L'expérience consiste à utiliser un de ces algorithmes pour la sélection des associations $X \longleftrightarrow Y$ et de comparer leur impact (TDPI et FPPI) sur le système complet (détection + suivi) pour la détection de panneaux circulaires. La table 4.19 résume les performances obtenues sur la séquence vidéo 1. Nous constatons que tous les algorithmes d'associations procurent des résultats similaires (TDPI=86% et FPPI=0.15) à l'exception de la décision duale qui permet seulement un TDPI de 83% avec un FPPI de 0.13. Cette légère baisse est dû au fait que l'algorithme dual reste indécis lorsque le conflit est important ce qui le rend inadapté pour le suivi mono-hypothèse (cf. Section 4.5.6). Pour les autres algorithmes, il est impossible de les différencier à partir des paramètres TDPI et FPPI. La seule différence observée est au niveau du temps de traitement (cf. Table 4.20). Les algorithmes basés sur une optimisation locale (PPL et PPCG) sont les plus rapides. Avec $n = m = 5$, PPL nécessite seulement 1.23 ms contre 28.16 ms pour le PPCJ ce qui consolide notre choix de l'approche locale pour le suivi de panneaux.

TABLE 4.20 – Temps de calcul moyen des algorithmes : PPCJ, PPL, PPCG, et la décision duale ainsi que de la méthode GNN.

	PPCJ	PPL	PPCG	Décision Duale	GNN
$(n = 3, m = 3)$	0.5	0.14	0.16	0.14	0.0013
$(n = 4, m = 4)$	3.55	0.39	0.39	0.39	0.0015
$(n = 5, m = 5)$	28.16	1.23	1.23	1.25	0.002

Finale­ment, nous avons comparé les méthodes crédales avec la méthode classique d’association GNN. Cette dernière procure des résultats similaires à ceux des algorithmes de décision avec 85.6% de TDPI et 0.15 de FPPI (cf. Table 4.19) pour des temps d’exécution assez attractifs (cf. Table 4.20) [Lauffenburger *et al.*, 2014]. Néanmoins, cette comparaison ne peut être pertinente puisque GNN est une approche d’association piste-à-cible contrairement aux méthodes crédales où le problème d’association de données est traité de façon bidirectionnelle.

4.7 Conclusion

Après avoir démontré l’intérêt du *Suivi temporel* pour les systèmes TSR (cf. Section 2.4), nous avons opté pour le développement d’une approche de suivi-*après*-détection. L’approche proposée consiste à intégrer une étape de *Suivi temporel* après la *Détection* des panneaux qui est imparfaite en situation réelle. En effet, les résultats de la *Détection* sont imprécis et incertains. L’idée consiste donc à fusionner ces résultats afin de minimiser l’influence des FP sur les performances du système TSR. Pour cela, un algorithme de suppression des fausses ROIs dénommé SMR composé de trois étapes est proposé : *Filtrage*, *Association de données*, et *Analyse des pistes*.

La première étape, le *Filtrage*, est basée sur un ensemble de FK dont l’objectif est la prédiction et l’estimation des positions des ROIs dans le plan image. Quant à la deuxième étape, *Association de données*, elle consiste à mettre en correspondance les objets connus (ROIs suivies) avec les objets perçus (ROIs détectées) afin de mettre à jour les FK et interpréter les événements de la scène. Cette étape est primordiale pour garantir le bon fonctionnement du TSR : les pistes suivies doivent correspondre à des panneaux réellement présents dans les images pour ne pas initier la reconnaissance sur des aberrations. C’est pourquoi, nous avons opté pour une association crédale basée sur la théorie des croyances. Cette dernière modélise les données imparfaites par une distribution de masses de croyance représentant la confiance accordée aux associations possibles entre les objets perçus et les objets connus. L’association crédale intégrée dans l’algorithme SMR est basée sur le Modèle de Croyance Transférable qui est une interprétation subjective et non-probabiliste de la théorie des croyances. Son modèle de raisonnement est basé sur deux niveaux : crédal et pignistique. Au niveau crédal, les masses de croyance sont utilisées pour représenter et combiner les différentes données issues de la *Détection*, tandis qu’au niveau pignistique, ces masses

sont transformées en mesures de probabilité (probabilité pignistique) pour prendre une décision sur les associations à retenir. Nous avons choisi l'algorithme PPL pour la sélection des meilleures associations. Cette solution a pour principal avantage de ne nécessiter aucun calcul supplémentaire une fois les probabilités d'association calculées. Elle retient, parmi les associations possibles, celles dont la probabilité est maximale. Finalement, la dernière étape de l'algorithme SMR reconnaît les FP en analysant les pistes des ROIs. Des expérimentation en conditions réelles ont montré que l'approche proposée permet une réduction significative des FP.

Conclusion générale

La sécurité routière est un problème de santé publique très préoccupant vu les conséquences des accidents de la route. Outre les préjudices physiques, les accidents entraînent un coût économique important aux victimes et à la société. Ces dernières décennies la sécurité routière a eu un engouement particulier auprès des constructeurs automobiles (BMW, Mercedes-Benz, Nissan, Ford, PSA Peugeot Citroën, etc.) pour le développement des systèmes avancés d'aide à la conduite communément appelés "ADAS" (*Advanced Driver Assistance System*). L'objectif est de garantir la sécurité des passagers du véhicules et des usagers de la route. Dans ce contexte, les ADAS assistent le conducteur durant son trajet en détectant les situations à risque afin d'éviter l'accident. Le système de reconnaissance de panneaux routiers (TSR- *Traffic Signs Recognition*) est l'un de ces ADAS qui tient informé le conducteur de la signalisation liée au trajet parcouru. Durant le processus de conduite le conducteur peut avoir des moments d'inattention ce qui peut présenter un danger. En effet, manquer un panneau de limitation de vitesse, à cause d'un appel téléphonique juste avant de négocier un virage, engendre une situation à risque qu'il ne faut pas négliger. Ainsi, avec un système TSR le conducteur est informé en temps-réel de la signalisation via des messages visuels et/ou sonores ce qui attire son attention sur son environnement routier.

Les systèmes TSR sont basés sur un capteur optique. La scène routière est perçue via une caméra embarquée, sur l'avant du véhicule, dont les images seront interprétées par le système en reconnaissant la signalisation présente. Pour cela, deux principales étapes sont nécessaires : *Détection* et *Reconnaissance* (appelée aussi classification). La première étape détecte la présence des panneaux dans les images. Ce processus consiste à déterminer les régions d'intérêt (ROI - *Region Of Interest*) susceptibles de contenir un panneau et cela en se basant sur leurs caractéristiques visuelles comme la couleur et la forme. Quant à la deuxième étape, elle reconnaît les panneaux détectés en identifiant leur pictogramme. Cependant, après la *Reconnaissance*, il est possible d'ajouter une étape de suivi des panneaux reconnus afin de ne pas les re-signaliser au conducteur. En effet, les panneaux apparaissent sur plusieurs images avant leur disparition du champ de vision de la caméra ce qui permet leur suivi. Ce processus de *Suivi temporel* est assuré par une étape d'association de données. Cette dernière met en relation les panneaux connus (reconnus précédemment) avec les panneaux perçus (reconnus à l'instant).

Dans cette thèse, nous nous sommes intéressés à l'aspect temporel qui demeure faiblement exploité dans les systèmes TSR. L'idée était de mieux exploiter cet aspect pour améliorer la *Détection*. En effet, la *Détection* est imparfaite en situation réelle, un certain nombre de fausses détections communément appelées "les *Faux Positifs*" (FP) peuvent être générées ce qui dégrade les performances d'un système TSR. L'idéal étant d'avoir un système ayant un fort taux de détection avec

le plus faible nombre possible de FP. C'est pourquoi, les objectifs de cette thèse ont été multiples : d'une part l'amélioration de la *Détection* et d'une autre part l'exploitation de l'information temporelle pour éliminer les FP.

La première partie des travaux était consacrée à l'étude de l'étape de *Détection*. Le chapitre 2 a présenté une classification des algorithmes de *Détection* en trois catégories : approches couleurs, approches contours, et approches descripteurs. Le chapitre 3 a été consacré à notre détecteur original de panneaux. L'idée est de sélectionner les positions des panneaux candidats par une détection de coins relatifs à la forme géométrique et aux pictogrammes des panneaux. Ensuite, un codage de contours est effectuée en fonction des orientations du gradient des pixels. Le résultat de ce codage est utilisé pour rechercher les formes géométriques, triangle et cercle, caractérisant les panneaux autour de chaque position candidate. Pour cela deux approches ont été proposées : l'algorithme RSLD (*Ransac Symmetric Lines Detection*) pour la *Détection* de panneaux triangulaires et une méthode de *template matching* pour la *Détection* de panneaux circulaires. Les résultats d'évaluation des approches ont été discutés dans la dernière partie du chapitre 3.

La seconde partie des travaux a été dédiée au *Suivi temporel* des panneaux. Tout d'abord, le chapitre 2 a présenté une classification des récents systèmes TSR intégrant le *Suivi temporel*. Deux classes d'approches sont définies selon la séquentialité des étapes : suivi-*après*-reconnaissance et suivi-*après*-détection. Ensuite les principes du processus de suivi, le filtrage et l'association de données, ont été présentés. Le chapitre 4 a été consacré à notre algorithme de Suivi Multi-ROIs (SMR). L'idée était d'exploiter la redondance de la signalisation pour remédier aux imperfections de la *Détection* : détections multiples pour le même panneau, échec de détection dû aux occultations temporaires, et les FP. Pour cela, nous avons fusionné les résultats de la *Détection*, qui sont souvent imprécis et incertains, ce qui a permis d'améliorer les performances du système TSR. Comme les données à fusionner sont imparfaites, nous avons opté pour la théorie des croyances. Cette dernière offre les mécanismes nécessaires pour la modélisation de la connaissance et la méconnaissance. Le chapitre 4 a présenté cette théorie et a décrit son implementation dans l'algorithme SMR. La dernière partie du chapitre 4 a discuté les résultats obtenus.

Bien que les contributions apportées par cette thèse présentent des résultats satisfaisants pour les systèmes TSR, plusieurs améliorations sont envisageables. Notamment afin d'intégrer ces systèmes dans les véhicules totalement automatiques dits "*autonomes*" où la sécurité doit être irréprochable. En effet, ces véhicules "*autonomes*" ou sans conducteur doivent appréhender leur environnement sans l'intervention de l'homme. C'est pourquoi, une perception 3D de l'environnement, exhaustive, dynamique, et robuste aux perturbations, constitue un élément majeur dans leur développement. Ainsi, plusieurs modules de perception étendue peuvent être développés en se basant principalement sur :

- La fusion des capteurs embarqués : caméra, Lidar, GPS, GSM, etc. En effet, chaque capteur a ses propres limites de perception qui peuvent être soulevées si on fusionne les données issues de

plusieurs capteurs. Ce module de fusion de capteurs peut être aisément intégré à notre système TSR basé sur la vision ce qui permet d'améliorer ses performances. En effet, la caméra fournit des informations sur l'environnement périphérique (i.e. le contexte locale) or le GPS permet une perception globale de l'environnement routier (attribues de la route parcourue). Par conséquent, leur fusion étend la perception de l'environnement permettant ainsi de minimiser les erreurs de détection et de reconnaissance des panneaux. Il est aussi envisageable de fusionner avec ces deux capteurs : une caméra 3D ce qui permet d'obtenir une information sur la profondeur de la scène. Cette information s'avère pertinente pour localiser les panneaux ainsi que les obstacles présents sur la voie de roulement comme les véhicules et les piétons.

- La fusion des ADAS tels que : TSR, détection d'obstacles, détection de marquages au sol, etc. Actuellement, les différents ADAS embarqués fonctionnent de façon indépendante et chacun d'entre eux détient sa propre interprétation de l'environnement. Nous estimons qu'il est primordial de développer des ADAS coopératifs au sein du même véhicule afin d'augmenter la perception. Par exemple, les informations fournies par le détecteur de marquages au sol permettent de localiser le véhicule sur la voie de roulement et détermine le type de marquage : continu ou discontinu. Cette information s'avère pertinente pour le système TSR puisqu'elle permet de déterminer par exemple si le panneau lié à la bretelle de sortie concerne le véhicule considéré ou pas. De même, les résultats du système TSR peuvent contribuer à la localisation des marquages. En effet, la position ainsi que la résolution des panneaux dans le plan image procure une information de profondeur utile pour une meilleure interprétation de l'environnement.

Bibliographie

- [Alshaqaqi, 2014] ALSHAQAQI, B. (2014). *Développement des systèmes de vision d'aide à la conduite embarquée sur le véhicule*. Thèse de doctorat, Université des Sciences et de la Technologie d'Oran - Mohamed Boudiaf.
- [American Honda Motor, 2006] AMERICAN HONDA MOTOR (2006). Collision Mitigation Brake System (CMBS). Rapport technique.
- [Ammoun, 2007] AMMOUN, S. (2007). *Contribution des communications intervéhiculaires pour la conception de systèmes avancés d'aide à la conduite*. Thèse de doctorat, École des Mines de Paris.
- [Balas et Sinha, 2003] BALAS, B. J. et SINHA, P. (2003). STICKS : Image-representation via non-local comparisons. *Journal of Vision*, 3(9):12–12.
- [Bar-Shalom et Tse, 1975] BAR-SHALOM, Y. et TSE, E. (1975). Tracking in a Cluttered Environment With Probabilistic Data Association. *Automatica*, 11(5):451–460.
- [Barnes et al., 2005] BARNES, N., LOY, G., SHAW, D. et ROBLES-KELLY, A. (2005). Regular polygon detection. In *International Conference on Computer Vision*, pages 778–785, Beijing, China.
- [Barnes et Zelinsky, 2004] BARNES, N. et ZELINSKY, A. (2004). Real-time radial symmetry for speed sign detection. In *IEEE Intelligent Vehicles Symposium*, pages 556–571, Parma, Italy.
- [Barnes et al., 2008] BARNES, N., ZELINSKY, a. et FLETCHER, L. (2008). Real-Time Speed Sign Detection Using the Radial Symmetry Detector. *IEEE Transactions on Intelligent Transportation Systems*, 9(2):322–332.
- [Baró et al., 2009] BARÓ, X., ESCALERA, S., VITRIÀ, J. et RADEVA, P. (2009). Traffic Sign Recognition using Evolutionary Adaboost detection and Forest-ECOC classification. *IEEE Transactions on Intelligent Transportation Systems*, 10(1):113–126.
- [Belaroussi et Tarel, 2009] BELAROUSSE, R. et TAREL, J.-P. (2009). Angle Vertex and Bisector Geometric Model for Triangular Road Sign Detection. In *IEEE Workshop on Applications of Computer Vision (WACV'09)*, pages 577–583, Snowbird, Utah, USA.
- [Belaroussi et Tarel, 2010] BELAROUSSE, R. et TAREL, J.-P. (2010). Détection des panneaux de signalisation routière par accumulation bivariée. *Traitement du signal*, 27(3):265–298.

- [Bernet, 2012] BERNET, S. (2012). *Analyse du regard pour le développement d'indicateurs du risque. Application à la formation et l'évaluation des conducteurs*. Thèse de doctorat, Université de Haute-Alsace.
- [Blackman et Popoli, 1999] BLACKMAN, S. et POPOLI, R. (1999). *Design and Analysis of Modern Tracking System*. Artech House Radar Library.
- [Blackman, 1986] BLACKMAN, S. S. (1986). *Multiple-target tracking with radar applications*. Artech House Radar Library.
- [Boumediene et al., 2013] BOUMEDIENE, M., CUDEL, C., BASSET, M. et OUAMRI, A. (2013). Triangular traffic signs detection based on RSLD algorithm. *Machine Vision and Applications*, 24(8):1721–1732.
- [Boumediene et al., 2014a] BOUMEDIENE, M., LAUFFENBURGER, J.-P., DANIEL, J. et CUDEL, C. (2014). Coupled Detection, Association and Tracking for Traffic Sign Recognition. *In IEEE Intelligent Vehicle Symposium*, Dearborn, Michigan, USA.
- [Boumediene et al., 2014b] BOUMEDIENE, M., LAUFFENBURGER, J.-P., DANIEL, J., CUDEL, C. et OUAMRI, A. (2014). Multi-ROI Association and Tracking With Belief Functions : Application to Traffic Sign Recognition. *IEEE Transactions on Intelligent Transportation Systems*, 15(6):2470–2479.
- [Broggi et al., 2007] BROGGI, A., CERRI, P., MEDICI, P., PORTA, P. P. et GHISIO, G. (2007). Real Time Road Signs Recognition. *In IEEE Intelligent Vehicles Symposium*, pages 981–986, Istanbul, Turkey.
- [Broggi et al., 2009] BROGGI, A., MEMBER, S., CERRI, P., GHIDONI, S., GRISLERI, P. et JUNG, H. G. (2009). A New Approach to Urban Pedestrian Detection for Automatic Braking. *IEEE Transactions on Intelligent Transportation Systems*, 10(4):594–605.
- [Burges, 1998] BURGES, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167.
- [Chen et al., 2014] CHEN, X., THARMARASA, R. et KIRUBARAJAN, T. (2014). Chapter 15 : Multitarget Multisensor Tracking. *In Academic Press Library in Signal Processing*, pages 759–812.
- [Cherfaoui, 2009] CHERFAOUI, V. (2009). *Contribution à la gestion des incertitudes en fusion multicapteurs : Application à la perception du contexte de conduite*. Thèse d'habilitation à diriger des recherches, Université de Technologie de Compiègne.
- [Cifuentes et al., 2012] CIFUENTES, C. G., STURZEL, M., JURIE, F. et BROSTOW, G. (2012). Motion Models that Only Work Sometimes. *In British Machine Vision Conference*, Surrey, UK.
- [CNPSR, 2012] CNPSR (2012). Statistiques annuelles des accidents de la route - 2012. Rapport technique, Centre National de Prévention et de Sécurité Routière.
- [Creusen et al., 2010] CREUSEN, I. M., WIJNHOFEN, R. G. J., HERBSCHLEB, E. et de WITH, P. H. N. (2010). Color exploitation in hog-based traffic sign detection. *In International Conference on Image Processing*, pages 2669–2672, Hong Kong, China.

- [Daimler AG, 2013] DAIMLER AG (2013). Driver Assistance Systems, Technology Guide. Rapport technique.
- [Dalal, 2005] DALAL, N. (2005). *Finding People in Images and Videos*. Thèse de doctorat, L’Institut National Polytechnique de Grenoble.
- [Dalal et Triggs, 2005] DALAL, N. et TRIGGS, B. (2005). Histograms of Oriented Gradients for Human Detection. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 886–893, San Diego, CA, USA.
- [Dallil et al., 2013] DALLIL, A., OUSSALAH, M. et OULDALI, A. (2013). Sensor Fusion and Target Tracking Using Evidential Data Association. *IEEE Sensors Journal*, 13(1):285–293.
- [Daniel et Lauffenburger, 2012] DANIEL, J. et LAUFFENBURGER, J.-P. (2012). Multi-Object Association Decision Algorithms with Belief Functions. *In International Conference on Information Fusion*, pages 669 – 676, Singapore.
- [Daniel et Lauffenburger, 2013] DANIEL, J. et LAUFFENBURGER, J.-P. (2013). Fusing navigation and vision information with the Transferable Belief Model : Application to an intelligent speed limit assistant. *Information Fusion*, 18:62–77.
- [Deguchi et al., 2011] DEGUCHI, D., SHIRASUNA, M., DOMAN, K., IDE, I. et MURASE, H. (2011). Intelligent Traffic Sign Detector : Adaptive Learning Based on Online Gathering of Training Samples. *In IEEE Intelligent Vehicles Symposium*, pages 72–77, Baden-Baden, Germany.
- [Dempster, 1967] DEMPSTER, A. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical theory of evidence*, 38:325–339.
- [Denoeux, 1995] DENOEU, T. (1995). A k-Nearest Neighbor Classification Rule Based on Dempster-Shafer Theory. *IEEE Transaction on Systems, Man and Cybernetics*, 25(5):804–813.
- [Denoeux et al., 2014] DENOEU, T., EL ZOGHBY, N., CHERFAOUI, V. et JOUGLET, A. (2014). Optimal Object Association in the Dempster–Shafer Framework. *IEEE Transactions on Cybernetics*, 44(12):2521–2531 .
- [Dubois et Prade, 1988] DUBOIS, D. et PRADE, H. (1988). Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4(3):244–264.
- [Ehrlich, 2009] EHRLICH, J. (2009). Technologies d’assistance a la conduite automobile. *Technique de l’ingénieur, l’expertise technique et scientifique de référence se3772*.
- [El Zoghby et al., 2013] EL ZOGHBY, N., CHERFAOUI, V. et DENOEU, T. (2013). Optimal object association from pairwise evidential mass functions. *In International Conference on Information Fusion*, pages 774–780, Istanbul, Turkey.
- [Fang et al., 2003] FANG, C.-Y., CHEN, S.-W. et FUH, C.-S. (2003). Road-Sign Detection and Tracking. *IEEE Transactions on Vehicular Technology*, 52(5):1329–1341.

- [Fischler et Bolles, 1981] FISCHLER, M. A. et BOLLES, R. C. (1981). Random Sample Consensus : A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24(6):381–395.
- [Fortmann *et al.*, 1980] FORTMANN, T. E., BAR-SHALOM, Y. et SCHEFFÉ, M. (1980). Multi-target tracking using joint probabilistic data association. *In IEEE Conference on Decision and Control including the Symposium on Adaptive Processes*, pages 807–812, Albuquerque, NM, USA.
- [Freund et Schapire, 1997] FREUND, Y. et SCHAPIRE, R. E. (1997). A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting. *Journal of Computer and System Sciences*, 55(1):119–139.
- [Garcia-Garrido *et al.*, 2011] GARCIA-GARRIDO, M., OCANA, M., LLORCA, D. et SOTELO, M. (2011). Robust Traffic Signs Detection by means of Vision and V2I Communications. *In IEEE Conference on Intelligent Transportation Systems*, pages 1003–1008, Washington, DC, USA.
- [Garcia-Garrido *et al.*, 2006] GARCIA-GARRIDO, M. A., SOTELO, M. A. et MARTÍN-GOROSTIZA, E. (2006). Fast Traffic Sign Detection and Recognition Under Changing Lighting Conditions. *In IEEE Conference on Intelligent Transportation Systems*, pages 811–816, Toronto, Canada.
- [Gil Jiménez *et al.*, 2008] GIL JIMÉNEZ, P., BASCÓN, S. M., MORENO, H. G., ARROYO, S. L. et FERRERAS, F. L. (2008). Traffic sign shape classification and localization based on the normalized FFT of the signature of blobs and 2D homographies. *Signal Processing*, 88(12):2943–2955.
- [Gómez-moreno *et al.*, 2010] GÓMEZ-MORENO, H., MALDONADO-BASCÓN, S., GIL-JIMÉNEZ, P. et LAFUENTE-ARROYO, S. (2010). Goal Evaluation of Segmentation Algorithms for Traffic Sign Recognition. *IEEE Transactions on Intelligent Transportation Systems*, 11(4):917–930.
- [González *et al.*, 2011] GONZÁLEZ, A., GARCÍA-GARRIDO, M. A., LLORCA, D. F., GAVILÁN, M., FERNÁNDEZ, J. P., ALCANTARILLA, P. F., PARRA, I., HERRANZ, F., BERGASA, L. M., SOTELO, M. A. et TORO, P. R. D. (2011). Automatic Traffic Signs and Panels Inspection System Using Computer Vision. *IEEE Transaction on Intelligent Transportation Systems*, 12(2):485–499.
- [Grigorescu et Petkov, 2003] GRIGORESCU, C. et PETKOV, N. (2003). Distance Sets for Shape Filters and Shape Recognition. *IEEE Transactions on Image Processing*, 12(10):1274–86.
- [Gruyer et Berge-Cherfaoui, 1999] GRUYER, D. et BERGE-CHERFAOUI, V. (1999). Multi-objects association in perception of dynamical situation. *In Conference on Uncertainty in Artificial Intelligence*, pages 255–262, Stockholm, Sweden.
- [Gu *et al.*, 2011] GU, Y., YENDO, T., PANAHPOUR TEHRANI, M., FUJII, T. et TANIMOTO, M. (2011). Traffic sign detection in dual-focal active camera system. *In IEEE Intelligent Vehicles Symposium*, pages 1054–1059, Baden-Baden, Germany.
- [Harris et Stephens, 1988] HARRIS, C. et STEPHENS, M. (1988). A combined corner and edge detector. *In Alvey Vision Conference*, pages 147–152, , Manchester, UK.
- [Houben, 2011] HOUBEN, S. (2011). A single target voting scheme for traffic sign detection. *In IEEE Intelligent Vehicles Symposium*, pages 124–129, Baden-Baden, Germany.

- [iTowns, 2009] iTOWNS. La Base de données Stereopolis, consultée le 30-10-2014 sur <http://www.itowns.fr/benchmarking.html>.
- [Jean-Christophe RIAT, 2005] JEAN-CHRISTOPHE RIAT (2005). Lane departure warning system developed by PSA Peugeot Citroën. Rapport technique.
- [Keller *et al.*, 2008] KELLER, C. G., SPRUNK, C., BAHLMANN, C., GIEBEL, J., BARATOFF, G. et FREIBURG, A.-I.-U. (2008). Real-time Recognition of U . S . Speed Signs. In *IEEE Intelligent Vehicles Symposium*, pages 518–523, Eindhoven, Netherlands.
- [Kohlas et Monney, 1995] KOHLAS, J. et MONNEY, P.-A. (1995). A Mathematical Theory of Hints. An Approach to the Dempster-Shafer Theory of Evidence. *Lecture Notes in Economics and Mathematical Systems*.
- [Lafuente-Arroyo *et al.*, 2005] LAFUENTE-ARROYO, S., GIL-JIMENEZ, P., MALDONADO-BASCON, R., LOPEZ-FERRERAS, F. et MALDONADO-BASCON, S. (2005). Traffic sign shape classification evaluation I : SVM using Distance to Borders. In *IEEE Intelligent Vehicles Symposium*, pages 557–562, Las Vegas, Nevada, USA.
- [Lafuente-Arroyo *et al.*, 2008] LAFUENTE-ARROYO, S., MALDONADO-BASCON, S., GIL-JIMENEZ, P. et GOMEZ-MORENO, H. (2008). An intra-image tracking algorithm for traffic sign recognition. In *IEEE International Conference on Vehicular Electronics and Safety*, pages 259–264, Columbus, OH, USA.
- [Landesa-Vzquez *et al.*, 2010] LANDESA-VZQUEZ, I., PARADA-LOIRA, F. et ALBA-CASTRO, J. L. (2010). Fast Real-Time Multiclass Traffic Sign Detection based on Novel Shape and Texture Descriptors. In *IEEE Conference on Intelligent Transportation Systems*, pages 1388–1395, Madeira Island, Portugal.
- [Lauffenburger *et al.*, 2014] LAUFFENBURGER, J.-P., DANIEL, J. et BOUMEDIENE, M. (2014). Traffic Sign Recognition : Benchmark of Credal Object Association Algorithms. In *International Conference on Information Fusion*, Salamanca, Spain.
- [Lauffenburger *et al.*, 2013] LAUFFENBURGER, J.-P., DANIEL, J. et SAIF, O. (2013). Multi-Object Association with Belief Functions : the Dual Decision Algorithm. In *IFAC Workshop on Advances in Control and Automation Theory for Transportation Applications*, Istanbul, Turkey.
- [Lienhart et Maydt, 2002] LIENHART, R. et MAYDT, J. (2002). An extended set of Haar-like features for rapid object detection. In *International Conference on Image Processing*, pages 900–903, Rochester, New York, USA.
- [Lin *et al.*, 2010] LIN, H., DU, P., ZHAO, W., ZHANG, L. et SUN, H. (2010). Image registration based on corner detection and affine transformation. In *International Congress on Image and Signal Processing*, pages 2184–2188, Yantai, China.
- [Liu *et al.*, 2002] LIU, H. A. N., LIU, D. et XIN, J. (2002). Real-time recognition of road traffic sign in motion image based. In *International Conference on Machine Learning and Cybernetics*, pages 83–86, Beijing, China.

- [Loy et Barnes, 2004] LOY, G. et BARNES, N. (2004). Fast Shape-based Road Sign Detection for a Driver Assistance System. *In International Conference on Intelligent Robots and Systems*, pages 70–75, Sendai, Japan.
- [Loy et Zelinsky, 2003] LOY, G. et ZELINSKY, A. (2003). Fast Radial Symmetry for Detecting Points of Interest. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):959–973.
- [Maldonado-bascón *et al.*, 2007] MALDONADO-BASCÓN, S., LAFUENTE-ARROYO, S., GIL-JIMÉNEZ, P., GÓMEZ-MORENO, H. et LÓPEZ-FERRERAS, F. (2007). Road-Sign Detection and Recognition Based on Support Vector Machines. *IEEE Transaction on Intelligent Transportation Systems*, 8(2):264–278.
- [Martin, 2005] MARTIN, A. (2005). La fusion d’informations. Rapport technique.
- [Mèche, 2013] MÈCHE, A. (2013). *Le filtre de Kalman et ses applications à la poursuite et à la détection de voies*. Thèse de doctorat, Université des Sciences et de la Technologie d’Oran - Mohamed Boudiaf.
- [Mercier *et al.*, 2011] MERCIER, D., LEFÈVRE, E. et JOLLY, D. (2011). Object association with belief functions, an application with vehicles. *Information Sciences*, 181(24):5485–5500.
- [Meuter *et al.*, 2008] MEUTER, M., KUMMERT, A. et MULLER-SCHNEIDERS, S. (2008). 3D Traffic Sign Tracking Using a Particle Filter. *In IEEE Conference on Intelligent Transportation Systems*, pages 168–173, Beijing , China.
- [Meuter *et al.*, 2011] MEUTER, M., NUNN, C., GÖRMER, S. M., MÜLLER-SCHNEIDERS, S. et KUMMERT, A. (2011). A Decision Fusion and Reasoning Module for a Traffic Sign Recognition System. *IEEE Transactions on Intelligent Transportation Systems*, 12(4):1126–1134.
- [Milanés *et al.*, 2012] MILANÉS, V., VILLAGRÁ, J., GODOY, J., SIMÓ, J., PÉREZ, J. et ONIEVA, E. (2012). An Intelligent V2I-Based Traffic Management System. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):49–58.
- [Møgelmoose *et al.*, 2012] MØGELMOOSE, A., TRIVEDI, M. M. et MOESLUND, T. B. (2012). Vision-Based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems : Perspectives and Survey. *IEEE Transactions on Intelligent Transportation Systems*, 13(4):1484–1497.
- [Mourllion, 2006] MOURLLION, B. (2006). *Extension d’un système de perception embarqué par communication : Application à la diminution du risque routier*. Thèse de doctorat, Paris-Sud XI.
- [Mourllion *et al.*, 2005] MOURLLION, B., GRUYER, D., ROY, C. et THÉROUDE, S. (2005). Multi-Hypotheses Tracking Algorithm Based on the Belief Theory. *In International Conference on Information Fusion*, Philadelphia, USA.
- [Moutarde *et al.*, 2007] MOUTARDE, F., BARGETON, A., HERBIN, A. et CHANUSSOT, L. (2007). Robust on-vehicle real-time visual detection of American and European speed limit signs, with

- a modular Traffic Signs Recognition system. *In IEEE Intelligent Vehicles Symposium*, pages 1122–1126, Istanbul, Turkey.
- [Negri, 2008] NEGRI, P. A. (2008). *Détection et reconnaissance d'objets structurés : application aux transports intelligents*. Thèse de doctorat, Pierre et Marie Curie.
- [Nguwi et Kouzani, 2008] NGUWI, Y.-Y. et KOUZANI, A. Z. (2008). Detection and classification of road signs in natural environments. *Neural Computing and Applications*, 17(3):265–289.
- [Nunn et al., 2008] NUNN, C., KUMMERT, A. et MULLER-SCHNEIDERS, S. (2008). A two stage detection module for traffic signs. *In IEEE International Conference on Vehicular Electronics and Safety*, pages 248–252, Columbus, OH, USA.
- [Overett et al., 2014] OVERETT, G., TYCHSEN-SMITH, L., PETERSSON, L., PETERSSON, N. et ANDERSSON, L. (2014). Creating robust high-throughput traffic sign detectors using centre-surround HOG statistics. *Machine Vision and Applications*, 25(3):713–726.
- [Paclíka et al., 2000] PACLÍKA, P., NOVOTIČOVÁ, J., PUDILB, P. et SOMOLB, P. (2000). Road Sign Classification using the Laplace Kernel Classifier. *Pattern Recognition Letters*, 21(13-14):1165–1173.
- [Piccioli et al., 1996] PICCIOLI, G., DE MICHELI, E., PARODI, P. et CAMPANI, M. (1996). Robust method for road sign detection and recognition. *Image and Vision Computing*, 14(3):209–223.
- [Piccioli et al., 1994] PICCIOLI, G., MICHELI, E. D. et CAMPANI, M. (1994). A robust method for road sign detection and recognition. *In European Conference on Computer Vision*, pages 493–500, Stockholm, Sweden.
- [Prisacariu et al., 2010] PRISACARIU, V. A., TIMOFTE, R., ZIMMERMANN, K., REID, I. et GOOL, L. V. (2010). Integrating Object Detection with 3D Tracking Towards a Better Driver Assistance System. *In International Conference on Pattern Recognition*, pages 3344–3347, Istanbul, Turkey.
- [Puthon, 2013] PUTHON, A.-S. (2013). *Détermination de vitesse limite par fusion de données vision et cartographiques temps-réel embarquées*. Thèse de doctorat, Ecole des mines ParisTech.
- [Qingsong et al., 2010] QINGSONG, X., JUAN, S. et TIAN, L. (2010). A detection and recognition method for prohibition traffic signs. *In International Conference on Image Analysis and Signal Processing*, pages 583–586, Zhejiang, China.
- [Reid, 1979] REID, D. B. (1979). An algorithm for tracking multiple targets. *IEEE Transactions on Automatic Control*, 24(6):843–854.
- [Rombaut, 1998] ROMBAUT, M. (1998). Decision in multi-obstacle matching process using the theory of belief. *In International Conference on Advances in Vehicle Control and Safety (AVCS)*, pages 63–68, Amiens, France.
- [Royère, 2002] ROYÈRE, C. (2002). *Contribution à la résolution du conflit dans le cadre de la théorie de l'évidence : Application à la perception et à la localisation des véhicules intelligents*. Thèse de doctorat, Université Technique de Compiègne.

- [Royère *et al.*, 2000] ROYÈRE, C., GRUYER, D. et CHERFAOUI, V. (2000). Data association with believe theory. In *International Conference on Information Fusion*, Paris, France.
- [Ruta *et al.*, 2010] RUTA, A., LI, Y. et LIU, X. (2010). Real-time traffic sign recognition from video by class-specific discriminative features. *Pattern Recognition*, 43(1):416–430.
- [Ruta *et al.*, 2011] RUTA, A., PORIKLI, F., WATANABE, S. et LI, Y. (2011). In-vehicle camera traffic sign detection and recognition. *Machine Vision and Applications*, 22(2):359–375.
- [Salameh, 2011] SALAMEH, N. (2011). *Conception d'un système d'alerte embarqué basé sur les communications entre véhicules*. Thèse de doctorat, Institut National des Sciences Appliquées de Rouen.
- [Sepulcre *et al.*, 2013] SEPULCRE, M., GOZALVEZ, J. et HERNANDEZ, J. (2013). Cooperative vehicle-to-vehicle active safety testing under challenging conditions. *Transportation Research Part C : Emerging Technologies*, 26:233–255.
- [Shafer, 1976] SHAFER, G. (1976). *A mathematical theory of evidence*. Princeton University Press.
- [Sivaraman et Trivedi, 2010] SIVARAMAN, S. et TRIVEDI, M. M. (2010). Improved Vision-Based Lane Tracker Performance Using Vehicle. In *IEEE Intelligent Vehicles Symposium*, pages 676–681, San Diego, CA, USA.
- [Sivaraman et Trivedi, 2013] SIVARAMAN, S. et TRIVEDI, M. M. (2013). Vehicle Detection by Independent Parts for Urban Driver Assistance. *IEEE Transactions on Intelligent Transportation Systems*, 14(4):1597–1608.
- [Smets, 1990] SMETS, P. (1990). The Combination of Evidence in the Transferable Belief Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):447–458.
- [Smets, 1993] SMETS, P. (1993). Belief Functions : The Disjunctive Rule of Combination and the Generalized Bayesian Theorem. *International Journal of Approximate Reasoning*, 9(1):1–35.
- [Smets, 2005] SMETS, P. (2005). Decision making in the TBM : the necessity of the pignistic transformation. *International Journal of Approximate Reasoning*, 38(2):133–147.
- [Smets, 2007] SMETS, P. (2007). Analyzing the combination of conflicting belief functions. *Information Fusion*, 8(4):387–412.
- [Smets et Kennes, 1994] SMETS, P. et KENNES, R. (1994). The Transferable Belief Model. *Artificial Intelligence*, 66(2):191–234.
- [Timofte *et al.*, 2014] TIMOFTE, R., ZIMMERMANN, K. et VAN GOOL, L. (2014). Multi-view traffic sign detection, recognition, and 3D localisation. *Machine Vision and Applications*, 25(3):633–647.
- [Viola et Jones, 2001a] VIOLA, P. et JONES, M. (2001). Rapid object detection using a boosted cascade of simple features. In *IEEE Conference on Computer Vision and Pattern Recognition*, Kauai, Hawaii, USA.

- [Viola et Jones, 2001b] VIOLA, P. et JONES, M. (2001). Robust Real-time Object Detection. *In International Workshop on Statistical and Computational Theories of Vision - Modeling, Learning, Computing, and Sampling*, Vancouver, Canada.
- [Šegvić et al., 2014] ŠEGVIĆ, S., BRKIĆ, K., KALAFATIĆ, Z. et PINZ, A. (2014). Exploiting temporal and spatial constraints in traffic sign detection from a moving vehicle. *Machine Vision and Applications*, 25(3):649–665.
- [Weng et al., 2006] WENG, S.-K., KUO, C.-M. et TU, S.-K. (2006). Video object tracking using adaptive Kalman filter. *Journal of Visual Communication and Image Representation*, 17(6): 1190–1208.
- [OMS, 2013] ORGANISATION MONDIALE DE LA SANTÉ (2013). Global Status Report On road Safety. Rapport technique.
- [Winckler, 2011] WINCKLER, A. (2011). BMW Eco Navigation. *In Navigation Strategies USA*, San Jose, USA.
- [Yager, 1987] YAGER, R. R. (1987). On the dempster-shafer framework and new combination rules. *Information Sciences*, 41(2):93–137.
- [Zaklouta, 2011] ZAKLOUTA, F. (2011). *Multiclass Object Recognition for Driving Assistance Systems and Video Surveillance*. Thèse de doctorat, École des Mines de Paris.
- [Zaklouta et Stanciulescu, 2014] ZAKLOUTA, F. et STANCIULESCU, B. (2014). Real-time traffic sign recognition in three stages. *Robotics and Autonomous Systems*, 62(1):16–24.