De l'importance des méthodes de suppression de fond pour la détection d'objets mobiles dans des vidéos acquises par des cameras fixes : Etat de l'art, Nouvelles Approches et Perspectives

Thierry Bouwmans

▶ **To cite this version:**

Thierry Bouwmans. De l'importance des méthodes de suppression de fond pour la détection d'objets mobiles dans des vidéos acquises par des cameras fixes : Etat de l'art, Nouvelles Approches et Perspectives. Traitement des images [eess.IV]. Université de La Rochelle, 2014. tel-01654493

HAL Id: tel-01654493
https://hal.science/tel-01654493

Submitted on 4 Dec 2017

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

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

Habilitation à Diriger des Recherches
de l'Université de La Rochelle

14 Novembre 2014

DE L'IMPORTANCE DES METHODES DE SUPPRESSION DE FOND POUR LA DETECTION D'OBJETS MOBILES DANS DES VIDEOS ACQUISES PAR DES CAMERAS FIXES

Etat de l'art, Nouvelles Approches et Perspectives

par

Thierry BOUWMANS

Composition du jury :

Rapporteurs :		
	Jenny Benois-Pineau	Professeure, Univ. de Bordeaux 2
	Laure Tougne	Professeure, Univ. de Lyon 2
	Francois Brémond	Dir. de Recherche, INRIA Sophia-Antipolis
Examinateurs :	Thierry Chateau	Professeur, Univ. de Clermond-Ferrand
	Michel Berthier	Professeur, Univ. de La Rochelle

Laboratoire de Mathématiques, Image et Applications | UPRES EA 3165

Remerciements

Je tiens tout d'abord à remercier vivement Jenny Benois-Pineau (Professeure à l'Université de Bordeaux 2), Laure Tougne (Professeure à Lyon 2) et François Brémond (Directeur de Recherche à l'INRIA Sophia-Antipolis à Nice) d'avoir accepté d'être rapporteurs de ce manuscrit et d'avoir participé à mon jury d'habilitation à diriger des recherches.
Je remercie tout aussi vivement, Thierry Chateau (Professeur à l'Université de Clermond-Ferrand) ainsi que Michel Berthier (Professeur à l'Université de La Rochelle) d'être membre de mon jury en tant qu'examinateurs.
Mes remerciements vont également à Bertrand Vachon, El-Hadi Zahzah et Carl Frélicot pour m'avoir permis de co-encadrer des thèses avec eux, ainsi qu'aux doctorants et stagiaires que j'ai encadrés. Je remercie tout particulièrement Fida El Baf, Diana Farcas, Cristina Marghes et Charles Guyon.
Je pense aussi à remercier mes amis et collègues depuis mon arrivée à l'Université de La Rochelle comme Laurent Mascarilla, Bernard Besserer, Renaud Peteri, Christophe Saint-Jean, Pierre Courtellemont et Mike Eboueya ainsi que tous les collégues du MIA et du L3i.
Je remercie également Antoine Vacavant, Laure Tougne et Thierry Chateau pour m'avoir proposé le talk invité au Workshop BMC de ACCV 2012. J'addresse aussi toute ma gratitude à Mubarak Shah en sa qualité d'éditeur en chef du journal "Machine Vision and Applications" ainsi qu'à Larry Davis (Univ. of Maryland, USA), Jordi Gonzàlez (Univ. Barcelona, Spain), Massimo Piccardi (Univ. Sydney, Australia), Caifeng Shan (Philips Research, The Netherlands), pour m'avoir accorder leur confiance dans ma direction du numéro spécial. Je remercie également Fatih Porikli (NICTA, Australie), Benjamin Höferlin (Univ. Stuttgart, Deutschland) et Antoine Vacavant pour m'avoir suivi dans l'édition du handbook chez CRC Press.
Je remercie aussi toutes les personnes qui m'ont fait confiance dans mes différentes activités pédagogiques et d'enseignements comme Pascal Estraillier pour le Master Jeux Vidéos et Média Interactifs, Yvan Daniel pour la licence professionelle Lettres Cultures et Nouveaux Média ainsi que tous les collégues du département informatique.
Enfin, un remerciement amical à Sylviane et Erlandrie, secrétaires, qui assurent avec gaïté et efficacité le quotidien administratif.

Avant-Propos

Ce manuscrit décrit mes activités d'enseignement et de recherche depuis mon arrivée à l'Université de La Rochelle et se compose de deux grandes parties.

La première partie présente mon Curriculum Vitae en version détaillée. Je présente ainsi des informations relatives à mon cursus universitaire. Ensuite, mes activités d'enseignement effectuées durant les différents quatriennaux sont présentées. Puis, je présente mes activités de recherche ainsi que les activités d'encadrement de thèse et de stage relatives à celles-ci. Je décris ensuite mes activités administratives collectives ainsi que mes activités dans des projets. Finalement, je présente la liste exhaustive de mes publications depuis ma nomination à l'Université de La Rochelle.

La seconde partie présente la partie scientifique des travaux de recherche que j'ai menés et encadrés au sein du Laboratoire d'Informatique et d'Imagerie Industrielle (L3i) puis du laboratoire Mathématiques, Image et Applications (MIA). Dans ces deux laboratoires, mes recherches ont portés principalement sur la détection d'objets mobiles dans des séquences vidéo acquises dans des conditions non contrôlées. Ces travaux se déclinent en quatre axes majeurs développés dans quatre chapitres. Dans un premier temps, les aspects scientifiques portent sur un état de l'art approfondi du domaine des méthodes de suppression de fond pour la détection des objets mobiles dans les vidéos permettant ainsi au lecteur d'avoir une vue structurée et compléte de ce domaine. Ensuite, nous présentons trois approches innovantes dans la modélisation de fond qui sont basées respectivement sur 1) les concepts flous, 2) les approches de réduction de dimension discriminatives et mixtes et 3) les approches par ACP robustes par décomposition en matrices de rang faible et parcimonieuse.

Table des matières

I Partie 1 : Synthèse des Activités

1 Curriculum Vitae 1
- 1.1 Etat Civil 1
- 1.2 Affectation 1
- 1.3 Formation Doctorale 1
- 1.4 Parcours Professionnel 1
- 1.5 Parcours Enseignement 2
- 1.6 Parcours Recherche 2
- 1.7 Responsabilité Administrative (ULR) 2

2 Activités d'Enseignement 5
- 2.1 Enseignement au sein de l'ULR 5
- 2.2 Enseignement au sein d'une Université Européenne 7

3 Activités de Recherche 8
- 3.1 Etat de l'art dans la détection d'objets mobiles 8
- 3.2 Apport des concepts flous dans la détection d'objets mobiles 11
- 3.3 Approche discriminative et mixte par réduction de dimension pour la modélisation de fond 12
- 3.4 Approche par ACP robuste basée sur la décomposition en une matrice de rang faible et une matrice parcimonieuse 13
- 3.5 Autres travaux de recherche 14

4 Activités d'Encadrement en Recherche 14
- 4.1 Thèses 16
- 4.2 Master 2/DEA 19
- 4.3 Stages de Master 1/Maîtrise 22
- 4.4 Stage LMD ERASMUS 22
- 4.5 Stage Etudiant Etranger 25

5 Activités Administratives Collectives 26
- 5.1 Etablissement 26
- 5.2 Département Informatique 26
- 5.3 Laboratoire L3i/Laboratoire MIA 26

6 Activité dans des Projets 26

7 Rayonnement Scientifique 31

8 Liste des Publications . . . 36
8.1 Editorial . . . 37
8.2 Chapitres de Livre . . . 37
8.3 Journaux Internationaux avec Comité de Lecture . . . 38
8.4 Conférences Internationales avec Comité de Sélection et Actes . . . 39
8.5 Publications issues de Collaborations Locales . . . 40
8.6 Publications issues de Collaborations Internationales . . . 41
8.7 Publication de Vulgarisation . . . 41
8.8 Publication de Diffusion sur Internet . . . 41
9 Témoignages . . . 42

II Partie 2 : Recherche Scientifique

1 Introduction Générale 45

2 De l'importance de la détection d'objets mobiles dans les vidéos : Un domaine complet de recherche scientifique. 49
1 Introduction . . . 50
1.1 De l'utilisation des méthodes de suppression de fond . . . 50
1.2 De la multidisciplinarité scientifique des méthodes de suppression de fond . . . 53
1.3 Evolution et analyse du domaine . . . 54
1.4 Contributions . . . 55
2 Cadre commun général . . . 56
2.1 Terminologie . . . 57
2.2 Hypothèses . . . 61
2.3 Situations critiques . . . 61
3 Suppression de fond : Les différentes étapes . . . 70
3.1 Modélisation de fond . . . 70
3.2 Initialisation du fond . . . 71
3.3 Maintenance du fond . . . 71
3.4 Classification fond/objets mobiles . . . 73
3.5 Taille de la sous-entité de l'image . . . 74
3.6 Choix de la caractéristique . . . 74
3.7 Autres points clés . . . 76
4 Les modèles de fond . . . 86
4.1 Les modèles actuels . . . 86
4.2 Les modèles prospectifs . . . 87
4.3 Les modèles "Autres" . . . 90
5 Les stratégies . . . 91

6 Les implémentations temps réel . . . 95
6.1 Implémentations GPU . . . 95
6.2 Implémentations embarquées . . . 95
6.3 Implémentations sur architecture dédiée . . . 95
7 Les capteurs . . . 96
8 Evaluation des performances . . . 97
8.1 Evaluation avec vérité terrain . . . 97
8.2 Evaluation sans vérité terrain . . . 99
9 Les bases de données . . . 100
9.1 Les bases de données traditionnelles . . . 101
9.2 Les bases de données récentes . . . 104
10 Conclusion . . . 108

3 Apport des concepts flous dans la détection d'objets mobiles 111
1 Introduction . . . 111
2 Approche par les concepts flous : Une vue d'ensemble . . . 112
2.1 Principe des systèmes basés sur les concepts flous . . . 113
2.2 Modélisation floue de fond . . . 114
2.3 Classification floue fond/objets mobiles . . . 115
2.4 Mise à jour adaptative floue du fond . . . 116
2.5 Caractéristiques floues . . . 117
2.6 Post-traitements flous . . . 117
3 Modélisation du fond par un mélange de Gaussiennes floues de Type-2 . . . 119
3.1 Préambule . . . 119
3.2 T2-FMOG . . . 121
3.3 Application à la modélisation du fond . . . 123
3.4 Influence des coefficients de flous . . . 127
3.5 Discussion . . . 128
4 Détection d'objets mobiles par l'intégrale de Choquet . . . 133
4.1 Opérateurs d'agrégation . . . 134
4.2 Intégrales floues : Sugeno-Choquet . . . 135
4.3 Application à la détection d'objets mobiles . . . 137
4.4 Détection d'objets mobiles par l'utilisation de l'intégrale de Choquet 140
4.5 Discussion . . . 141
5 Règle de mise à jour adaptative . . . 141
6 Expérimentations et évaluations . . . 143
6.1 Modèlisation du fond par mélange de Gaussiennes floues de Type-2 . 143
6.2 Classification floue des pixels en fond/objet par l'intégrale de Choquet 154
6.3 Mise à jour de l'image du fond . . . 160
6.4 Temps de calcul . . . 161
7 Conclusion . . . 162

4 Approche discriminative et mixte par réduction de dimension pour la modélisation de fond **165**
1 Introduction . . . 165
2 Approche reconstructive de réduction de dimension pour la modélisation de fond : Une vue d'ensemble . . . 167
2.1 Réduction de dimension par ACP (RD-ACP) . . . 167
2.2 Réduction de dimension par ACI (RD-ACI) . . . 170
2.3 Réduction de dimension par INMF (RD-INMF) . . . 172
2.4 Réduction de dimension par Incremental Rank-Tensor (RD-IRT) . . . 173
2.5 Discussion . . . 173
3 Modélisation de fond par une approche discriminative de réduction de dimension . . . 174
3.1 Approches discriminatives de réduction de dimension : Une vue d'ensemble . . . 174
3.2 Réduction de dimension par la méthode de Critère de Marge Maximale 174
3.3 Modélisation de fond via Incremental Maximum Margin Criterion . . . 177
3.4 Algorithme SL-IMMC pour la suppression de fond . . . 179
3.5 Illustration . . . 180
4 Approche mixte de réduction de dimension pour la modélisation de fond . . . 183
4.1 Approches mixtes de réduction de dimension . . . 183
4.2 Définition du problème . . . 183
4.3 Modélisation de fond . . . 183
4.4 Classification . . . 184
4.5 Mise à jour du fond . . . 185
5 Expérimentations et évaluations . . . 186
5.1 Evaluation en termes de détection . . . 186
5.2 Evaluation en termes de complexité et de temps de calcul . . . 189
6 Conclusion . . . 190

5 Approche par ACP robuste basée sur la décomposition en une matrice de rang faible et une matrice parcimonieuse **193**
1 Introduction . . . 193
2 Approche par décomposition en une matrice de rang faible et une matrice parcimonieuse : Une vue d'ensemble . . . 194
2.1 Approches ACP robustes . . . 195
2.2 Approches NMF robustes . . . 205
2.3 Apprentissages robustes de sous-espaces . . . 206
2.4 Suivis de sous-espaces . . . 206
2.5 Minimisations en rang faible . . . 207
2.6 Application à la détection d'objets mobiles . . . 208
2.7 Discussion . . . 209

3 ACP Robuste par Iteratively Reweighted Least Squares 221
3.1 Minimisation avec contraintes spatiales 221
3.2 Détection des objets mobiles par factorisation robuste de la matrice de rang faible 222
3.3 Discussion 223
4 ACP robuste polynomiale 224
5 Expérimentations et évaluations 225
5.1 Evaluation des méthodes RPCA, suivi de sous-espaces et minimisation en rang faible sur BMC 2012 225
5.2 Evaluation de la méthode RPCA-IRLS 232
5.3 Evaluation de la méthode ACP polynomiale 237
6 Conclusion 238

6 Conclusion Générale **240**
1 Synthèse des avancées proposées 240
1.1 Concepts flous 240
1.2 Réduction de dimension discriminative et mixte 241
1.3 Décomposition en matrices de rang faible et parcimonieuse 241
2 A la poursuite de la méthode idéale 242
2.1 L'initialisation du fond 242
2.2 La modélisation de fond 242
2.3 La sélection de caractéristiques 242
3 Conclusion 243

Index **290**

Première partie

Partie 1 : Synthèse des Activités

Synthèse des Activités d'Enseignement et de Recherche

1 Curriculum Vitae

1.1 Etat Civil

— **Prénom, Nom :** Thierry BOUWMANS
— **Date de Naissance :** 28 Juin 1968
— **Nationalité :** Français

1.2 Affectation

— **Statut :** Maître de Conférences, 61ième section
— **Nomination :** 01/02/2001 (ULR)
— **Titularisation :** 01/02/2002 (ULR)
— **Téléphone :** 05.46.45.72.02
— **Email :** tbouwman@univ-lr.fr
— **Site web personnel :** http://sites.google.com/site/thierrybouwmans/

1.3 Formation Doctorale

— T. Bouwmans, "Apport de la logique floue dans un systeme d'analyse de scène 3D par multivision", Doctorat de l'Université du Littoral Côte d'Opale (ULCO), mention très honorable, 28/04/1997.

1.4 Parcours Professionnel

— 2002- en cours : MCF Titulaire (Section 61) - Université de La Rochelle (ULR)
— 01/02/2001-2002 : MCF Stagiaire (Section 61) - Université de La Rochelle (ULR)
— 01/09/2000-31/01/2001 : ATER Section 61 (6 mois) - Université de La Rochelle (ULR)
— 01/03/2000-31/08/2000 : ATER Section 61 (6 mois) - Université de La Rochelle (ULR)

1.5 Parcours Enseignement

— 01/02/2001- en cours : Service statutaire - Université de La Rochelle (ULR)
— 01/09/2000-30/01/2001 : ATER Section 61 (100%) - Université de La Rochelle (ULR)
— 01/02/2000-30/08/2000 : ATER Section 61 (50%) - Université de La Rochelle (ULR)
— 01/09/1997-30/08/1998 : Contractuel PRAG - l'Université du Littoral Côte d'Opale (ULCO)
— 01/09/1993-30/08/1997 : Vacataire - Université du Littoral Côte d'Opale (ULCO)

1.6 Parcours Recherche

— 01/01/2007- en cours : MCF - Laboratoire Mathématique Image et Applications (MIA, E.A. 3165) de l'ULR
— 01/02/2001-30/01/2006 : MCF - Laboratoire Laboratoire d'Informatique, Image et Interactions (L3I, E.A. 2118) de l'ULR
— 01/09/2000-30/01/2001 : ATER Section 61 (100%) - laboratoire L3I à l'Université de La Rochelle (ULR)
— 01/02/2000-30/08/2000 : ATER Section 61 (50%) - Laboratoire L3I à l'Université de La Rochelle (ULR)
— 01/09/1997-30/08/1998 : Recherche post-doctorale - Laboratoire LASL à l'Université du Littoral Côte d'Opale (ULCO)
— 01/09/1993-30/08/1997 : Doctorant - Laboratoire LASL à l'Université du Littoral Côte d'Opale (ULCO)

1.7 Responsabilité Administrative (ULR)

— Directeur des études du la Licence IUP Génie Informatique (2002-2003, 1 an)
— Directeur des études du DESS Jeux Vidéo et Média Interactifs sur La Rochelle (2001-2004, 3 ans)
— Responsable administratif et pédagogique de la Coloration Image dans le MASTER JVMI (2004-2008, 4 ans)
— Responsable de la formation en informatique du DEUG Lettres Modernes (2001-2004, 3 ans)
— Responsable de la formation en informatique de la Licence Professionnelle Lettres, Cultures et Nouveaux Média (2002- en cours, 11 ans)

Mes responsabilités administratives et pédagogiques en informatique sont donc relatives à trois formations : Informatique (IUP Génie Informatique et Licence Informatique), Jeux Vidéos et Media Interactifs et Lettres Modernes. Je présente ci-dessous les différentes tâches que j'ai effectué dans chaque responsabilité.

IUP "Génie Informatique" (2002-2003) : Dans le cadre de cette formation, j'ai assuré la charge de directeur des études qui consistait à faire les emplois du temps, la constitution des groupes de TP, la déclarations des heures des intervenants et d'organiser les jurys.

Licence "Informatique" (2013) : J'ai assuré la charge de responsable pédagogique de la Licence Informatique pendant 6 mois. Cette responsabilité a consisté à assurer le bon fonctionnement de la licence, les présidences de jury des semetres pairs et le recrutement pour l'année suivante.

DESS Jeux Vidéos et Media Interactifs (DESS JVMI) (2001-2004) : Mes activités au niveau du DESS JVMI ont consisté à participer à l'élaboration du contenu pédagogique de la formation en collaboration avec les différents partenaires ainsi qu'à la conception du fonctionnement. Le découpage entre les diff ?rents partenaires est le suivant : Univ. La Rochelle (Coloration Visuelle, Game Design, Chef de Projet), Univ. Poitiers (Ergonomie), CNAM Paris (Programmation, Coloration Sonore) et le CNBDI à Angoulême pour la localisation des enseignements. Cette mise en place s'est effectuée au cours de réunions mensuelles sur les différents sites durant la période 2000-2001. Entre le 01/09/ 2001 et le 30/06/2004, j'ai assumé la responsabilité pédagogique et administrative de cette formation conjointement avec Pascal Estraillier (PR) pour l'Université de La Rochelle. J'ai assumé ainsi la responsabilité des 3 UEs (30h Cours) du tronc commun relatives à l'Image (Traitement du Signal et de l'Information, Traitement des Images, Traitement de Séquences et de Vidéo) et la responsabilité de l'option Informatique et Traitement des Images (75h Cours, 75h TP). Dans ce cadre, j'ai élaboré le contenu de ces enseignements avec des intervenants universitaires de La Rochelle ainsi qu'avec des intervenants issus de l'industrie du jeu. Les différentes tâches administratives et pédagogiques réalisées ont été les suivantes :

- Recrutement sur dossier des étudiants de l'option "Informatique et Traitement des Images".
- Participation à la semaine d'intégration du DESS au CNBDI d'Angoulême.
- Réunions mensuelles pédagogiques sur les différents sites partenaires.
- Réalisation de la maquette des enseignements pour l'Université de La Rochelle.
- Recrutement des intervenants professionnels pour l'option "Informatique et Traitement des Images".
- Etude financière du matériel spécifique à l'option "Informatique et Traitement des Images".

Master "Jeux Vidéos et Media Interactifs" (Master JVMI) (2004-2008) : Mes activités au niveau du Master JVMI ont consisté à l'élaboration du contenu pédagogique de la formation qui est passée de 1 an (DESS JVMI) à 2 ans. Cette mise en place s'est effectuée au cours de plusieurs réunions mensuelles au CNBDI d'Angoulême. Sur le quatriennal 2004-2008,

j'ai assumé la responsabilité pédagogique et administrative de la coloration visuelle sur les 2 années du Master. Dans ce cadre, j'ai assumé la responsabilité des 3 UEs du tronc commun relatives à la coloration visuelle (Bases de l'informatique, Acquisition et Traitement de l'Image Fixe et Animée, Synthèse des Images) et des 6 UEs spécifiques à la Coloration Visuelle (Production des Images et de l'Animation pour les Média Numériques, Rendu Temps-réel, Animation 2D/3D, Effets Spéciaux, Contraintes liées à la Synthèse en Temps Réel, Compléments en Synthèse des Images et du Son). Dans ce cadre, j'ai élaboré le contenu de ces enseignements avec des intervenants universitaires de La Rochelle ainsi qu'avec des intervenants issus de l'industrie du jeu.

Deug "Lettres Modernes" (LM) (2001-2004) : Dans le cadre de le formation Deug LM mise en place par Yvan Daniel (PR) à la FLASH (ULR), j'ai conçu les enseignements d'informatique d'usage pour la première année ainsi que pour la deuxième année car cette formation était dans son année d'ouverture. Ces enseignements avait pour but de permettre aux étudiants de savoir utiliser les outils de bureautique de la suite Microsoft Office (Word, Excel, PowerPoint), de pouvoir naviguer de façon pertinente sur internet et de pouvoir créer un site web (Dreamweaver, HTML). Dans ce cadre, j'ai conçu 3 UEs de 15hTP intitulées respectivement Bureautique Niveau 1 (Traitement de texte, Présentation), Bureautique Niveau 2 (Tableur), et Bureautique Niveau 3 (Création de Site Web). Pour assurer l'adéquation avec la formation Lettres Modernes, je me suis rapproché des enseignants de la FLASH pour intégrer de façon optimale l'informatique dans cette formation. D'un point de vue pratique, cette collaboration s'est concrétisé par un projet commun entre le module d'informatique d'usage et le module d'histoire des média dans l'UE intitulée Bureautique Niveau 3. Ce projet consistait à apprendre aux étudiants à mettre sous forme d'un site web des informations recueillies lors de reportages effectués dans le domaine de l'événementiel sur La Rochelle.

Licence Professionnelle "Lettres, Cultures et Nouveaux Media" (LP-LCNM)(2002- en cours) : Lors de sa création en 2002 par Yvan Daniel (PR) à la FLASH (ULR, j'ai conçu les 4 UE d'informatique intitulées respectivement Environnement Informatique (20hTP), Initiation au Graphisme (26hTD), Recherche et veille d'informations sur internet (20hTP) et Création et Organisation de Site Web (25hTP). Les différentes tâches administratives et pédagogiques réalisées ont été les suivantes :

- Participation au recrutement des étudiants sur dossier et entretien,
- Réunions pédagogiques,
- Réalisation de la maquette des enseignements pour la partie informatique,
- Recrutement des intervenants professionnels
- Etude financière du matériel spécifique à la partie informatique.

Intitulé UE	Niveau	Formation	Forme	Heures	Années
Deug, Licence, Maîtrise (2000-2004)					
Bureautique Niveau 1 (R)	D	Deug LM	TP	56h	2002-2004
Bureautique Niveau 2 (R)	D	Deug LM	TP	48h	2002-2004
LMD (2004-2008) (2008-2012) (2012-2016)					
Informatique Usage N1	L	Licence L1	C, TP	262h	2004-2014
Informatique Usage N2	L	Licence L1	C, TP	126h	2004-2014
Informatique Usage N3	L	Licence L1	TP	10h	2004-2007
Méthodologie Universitaire (R3)	L	Licence L1	TD	240h	2004-2014

TABLE 1 – Informatique Usage et Méthodologie Universitaire - R (Responsabilité de l'UE) - 1, 2 et 3 : quatriennal concerné - Les heures sont indiquées en équivalent TD et entre parenthèses nombre d'heures effectuées par an.

2 Activités d'Enseignement

2.1 Enseignement au sein de l'ULR

Cette section regroupe les enseignements dont j'ai eu la responsabilité ou dans lesquels j'ai été simplement intervenant depuis ma nomination. Ces enseignements ont été effectués sur une période qui correspond à un quadrienenal (2000-2004) en format Deug, Licence, Maîtrise, DEA, DESS et Doctorat, et à trois quatriennaux en format LMD (2004-2008, 2008-2012, 2012-en cours). Ces enseignements peuvent être classés en trois catégories :

1. Informatique Usage et Méthodologie Universitaire (UE Transversales).
2. Traitement, Analyse et Synthèse d Images (27ième et 61ième Sections).
3. Informatique et Programmation (27ième Section).

Les tableaux 1, 2 et 3 montrent le détail de ces enseignements par intitulé d'UE. Tous ces enseignements correspondent à une moyenne approximative 225h équivalent TD par an. La responsabilité d'UE consiste à concevoir les cours, les TD et les TP à assurer, à effectuer l'animation pédagogique, à réaliser et surveiller les examens de première et seconde session, ainsi que la prise en charge d'au moins un groupe de TD et au moins un groupe de TP. Les enseignements sont accompagnés d'un support de cours soit sous forme électronique déposé sur la plateforme Moodle de l'ULR, soit sous forme papier. La responsabilité dans une UE est signalée par (R).

Intitulé UE	Niveau	Formation	Forme	Heures	Années
Deug, Licence, Maîtrise (2000-2004)					
Acquisition et Traitement du Signal	L	Licence IUP GI	TP	64h	2000-2004
Images Numériques	M	Maîtrise IUP GI	TP	28h	2000-2003
Synthèse des Images	L	Licence IUP GI	C, TD, TP	223h	2000-2004
Synthèse des Images Avancée	M	Maîtrise IUP GI	TD, TP	70h	2000-2004
Etude de Synthèse des Images	M	Maîtrise IUP GI	TP	36h	2000-2004
Traitement de Séquences et de Vidéo (R)	M	DESS JVMI	C, TD, TP	45h	2001-2003
LMD (2004-2008) (2008-2012) (2012-2016)					
Traitement du Signal	L	Licence IMA (S5)	TP	4h	2004-2005
Vision 1	L	Licence IMA (S6)	TP	20h	2004-2005
Jeux Vidéo et Informatique	L	Licence IMA (S2)	C	30h (7,5hC)	2005-2008
Visualisation Scientifique	M	Master 2 IMAE	C, TD, TP	103h	2004-2008
Images et Séquences d'Images	M	Master 2 IMAE	C	18h (4,5hC)	2004-2008
Eléments de Synthèse d'Images (R1)	M	Master 1 JVMI	C, TD, TP	56h	2004-2009
Compléments en Synthèse d'Images (R1)	M	Master 1 JVMI	C, TP	17h	2004-2005
Rendu Temps Réel (R1)	M	Master 2 JVMI	C, TD, TP	65h (3hC, 3hTP)	2005-2014
Jeux Vidéo et Informatique	L	Licence Informatique (S2)	C	13h	2008-2009
Suivi d'Objets en Mouvement	M	Master 2 IMI	C, TD, TP	65h (4,5hC, 9hTP)	2008-2013
Synthèse des Images (R2)(R3)	L	Licence Informatique (S4,S6)	C, TD, TP	207h (9hC, 9hTD, 18hTP)	2008-2014

TABLE 2 – Traitement, Analyse et Synthèse des Images - R (Responsabilité de l'UE) - 1, 2 et 3 : quatriennal concerné - Les heures sont indiquées en équivalent TD et entre parenthèses nombre d'heures effectuées par an.

Intitulé UE	Niveau	Formation	Forme	Heures	Années
Deug, Licence, Maîtrise					
Intelligence Artificielle	M	Licence IUP GI	TD, TP	193h	2000-2004
LMD (2004-2008) (2008-2012) (2012-2016)					
Génie Logiciel	L	Licence Informatique	TP	20h (30hTP)	2012-2014
Environnement Informatique (R1)(R2)(R3)	L	LP LCNM	TP	200h (25hTD)	2006-2014
Recherche d'information et veille sur Internet (R1)(R2)	L	LP LCNM	TP	60h	2004-2008
Création et Organisation de Site Web	L	LP LCNM	TP	25h	2009-2014

TABLE 3 – Informatique et Programmation - R (Responsabilité de l'UE) - 1, 2 et 3 : quatriennal concerné - Les heures sont indiquées en équivalent TD et entre parenthèses nombre d'heures effectuées par an.

2.2 Enseignement au sein d'une Université Européenne

ERASMUS Intensive Program on Computing Aspects of Computer Games Development [1](CACGD 2010)

- **Lieu :** Univ. La Rochelle (France)
- **Partenaires Européens :** Glamorgan University (GB), Technological Educational Institute of Crete (Grèce)
- **Dates :** 28 Juin 2010 - 10 Juillet 2010.
- **Cours :** Animation, Rendu Temps Réel (6h Cours).
- **Effectifs :** 50 étudiants européens.
- **Langue employée :** Anglais.

L'objectif de ce cours était de donner une vue générale des techniques d'animation et des techniques de rendu temps-réel dans le cadre d'une formation de deux semaines sur le jeux vidéo et les média numériques.

1. http://perso.univ-lr.fr/mike/IP2010/index.html

3 Activités de Recherche

Suite à ma nomination à l'Université de La Rochelle, mon activité de recherche s'est effectuée au sein du Laboratoire d'Informatique et d'Imagerie Industrielle (L3i) pendant 7 ans. Depuis Janvier 2008, elle se réalise au laboratoire Mathématiques, Image et Applications (MIA). Dans ces deux laboratoires, mes recherches ont portés principalement sur la détection d'objets mobiles dans des séquences vidéo acquises dans des conditions non contrôlées. Cette activité a été menée conjointement avec d'autres membres du laboratoire soit dans le cadre de projets spécifiques (Aquathèque, Voile, Mertail, TOPVISION) soit dans le cadre de recherche fondamentale lors d'encadrement de thèse. Ces activités ont été menées conjointement avec les collègues suivants : P. Courtellemont (PR), B. Vachon (PR) et E. Zahzah (MCF-HDR). Ils m'ont permis d'encadrer scientifiquement des étudiants en thèse dans mon domaine de recherche (Section 4) dont les travaux ont donné lieu à des publications dans le domaine de la détection d'objets mobiles par suppression de fond. L'originalité de ces travaux se décline en quatres axes qui sont les suivants :

1. Un état de l'art complet dans le domaine de la soustraction de fond (Chapitre 1).
2. Une approche basée sur les concepts flous (Chapitre 2).
3. Une approche de réduction de dimension discriminative et mixte (Chapitre 3).
4. Une approche basée sur la décomposition de rang faible et parcimonieuse (Chapitre 4).

Mes travaux de recherche consistent donc à un état de l'art approfondi du domaine et à trois études majeures sur l'apport d'outils mathématiques pour la modélisation et la détection des objets mobiles dans les vidéos. Les démarches et motivations associées à ces différents travaux sont expliquées dans les sous-sections suivantes et l'aspect scientifique correspondant est développé dans la Partie 2. Parmi ces travaux, certains ont été améliorés par d'autres chercheurs comme expliqué le cas échéant dans chaque sous-section.

3.1 Etat de l'art dans la détection d'objets mobiles

Plusieurs états de l'art existent dans le domaine de la soustraction de fond. En 2000, Mc Ivor [259] a examiné 9 algorithmes de base permettant ainsi une première comparaison de ces modèles mais celle-ci se limitait à la description des algorithmes. En 2004, Piccardi [402] a fourni un état de l'art sur 7 méthodes et une première évaluation en termes de vitesse, de mémoire requise et de précision. Cet état de l'art permet aussi de comparer les algorithmes en termes de complexité algorithmique pour permettre de choisir l'algorithme le plus adapté à une application spécifique sur un matériel spécifique. En 2005, Cheung and Kamath [117] ont classé différentes méthodes en termes de techniques récursives et non-récursives. Suivant

cette classification, Elhabian et al. [169] ont élargi cet état de l'art en le complétant par un plus grand nombre de méthodes. Cependant, cette classification en termes de techniques récursives et non-récursives est plus adaptée pour l'étape de mise à jour du fond que pour celle de modélisation. En 2010, Cristani et al. [136] ont classé les algorithmes les plus utilisés en termes d'algorithmes utilisant un ou plusieurs capteurs mais cette classification n'est pas optimale parce que l'appartenance à l'une des catégories est plus basée sur le choix des capteurs plutôt que sur le modèle en lui-même. Dans ce contexte, j'ai proposé une classification originale des méthodes suivant le modèle mathématique utilisée générant ainsi différents états de l'art. Ma contribution a donc consisté en différents états de l'art portant soit sur toutes les catégories de modèles [CL-2][CL-3][J-1] soit sur une catégorie de modèles [CL-4][CL-5][CL-6][J-2][J-4] soit sur des sous-catégories de modèles [J-6][J-7][J-RE-1].

1- Sous-catégories

Les deux sous-catégories que j'ai étudiées particulièrement sont celles sujettes aux plus grandes nombres d'améliorations et sont les suivantes :

— **Mélange de Gaussiennes :** La modélisation de fond basée sur le mélange de Gaussiennes développée par Stauffer and Grimson [490] est le modèle le plus étudié avec plus de 250 articles recensés entre 1999 et 2013. Devant ce grand nombre d'articles, j'ai donc effectué un état de l'art complet et une classification des nombreuses améliorations de la version originale du mélange de Gaussiennes [490] pour en dégager les pistes principales. Cette classification catégorise les diverses améliorations en 1) améliorations intrinsèques concernant directement le modèle comme l'initialisation, la mise à jour du modèle, la classification des pixels ainsi que la taille et le type des caractéristiques employées, 2) améliorations extrinsèques concernant les stratégies mis en oeuvre pour appliquer le modèle comme l'utilisation des champs de Markov, les approches hiérarchiques et les approches multi-niveaux (pixel, région, image) et 3) les améliorations pour atteindre des performances temps-réel. Cet état de l'art a été effectué à l'époque (2008) sur plus de 170 articles [J-7].

— **Réduction de dimension :** Les méthodes de réduction de dimension sont également aussi largement utilisées pour modéliser le fond. La première méthode basée sur l'Analyse en Composantes Principales (ACP) a été proposée by Oliver et al. [387] et a également fait l'objet d'un grand nombre d'améliorations. D'autres méthodes de réduction de dimension de type Analyse en Composantes Indépendantes (ACI) ont été employées de façon plus marginale. J'ai donc effectué un état de l'art complet ainsi qu'une classification des différentes améliorations de la méthode de Oliver et al. [387]. De plus, j'ai évalué certaines d'entre elles sur la base de données Wallflower [521] en les comparant aux méthodes statistiques telles que (SG, MOG, KDE). Cet état de l'art a été effectué en 2009 sur plus de 60 articles [J-6]. De plus, cette étude a

mis en évidence des pistes pour améliorer la méthode de Oliver et al. [387] et a inspiré par exemple les travaux de Quivy et Kumazawa [316]. Finalement, ce travail a permis de dégager l'idée de tester et d'évaluer les méthodes dites discriminatives (voir Section 3.3) alors que seules les méthodes dites reconstructives avaient été employées jusqu'alors.

— **modèles basés sur les approches ACP robustes :** Les modèles utilisant les récentes avancées en ACP robuste basée sur la décomposition en une matrice de rang faible et une matrice parcimonieuse, ont été développés récemment suite aux travaux de Candès et al. [92]. Dans la plupart des articles, seuls des résultats visuels et qualitatifs étaient montrés et ne permettaient pas une comparaison quantitative rigoureuse de ces méthodes. J'ai donc initié un état de l'art et une évaluation quantitative de ces différentes méthodes sur la base de données BMC 2012 [530]. Cet état de l'art a été effectué en 2013 sur plus de 70 articles [CL-4][J-2].

2- Catégories

Parmis les trois catégories que j'ai étudiées particulièrement, deux sont parmis les plus récentes et une est la plus étudiée. Celles-ci sont dans l'ordre les suivantes :

— **Modèles flous :** Les challenges rencontrés en vidéo-surveillance comme les fonds dynamiques et les variations lumineuses engendrent imprécisions et incertitudes qui doivent être idéalement prises en compte dans toutes les étapes de la soustraction de fond. Dans ce contexte, plusieurs auteurs ont récemment introduit l'utilisation des concepts flous dans les étapes de la soustraction de fond. Etant l'un des précurseurs dans ce domaine lors de mes premiers travaux démarrés en 2006 avec la thèse de Fida El Baf (voir Section 3.2), j'ai effectué 6 ans plus tard un état de l'art dans le domaine sur plus de 50 articles sous la proposition de Lucia Maddalena (ICAR, Italie) [CL-5].

— **Modèles basés sur la décomposition en matrices de rang faible et parcimonieuse :** Ces différentes modèles peuvent se classer en trois grands groupes que sont les modèles d'ACP robustes (RPCA), ceux de suivi de sous-espaces (Subspace Tracking), et ceux de minimisation en rang faible (Low Rank Minimization). Nous avons étudié et évalué les méthodes de ces trois groupes sur les bases de données ChangeDetection.net [204] et BMC 2012 [530] dans [J-RE-1] portant sur plus de 160 publications.

— **Modèles statistiques :** Les modèles statistiques étant les plus utilisés, j'ai regroupé tous les travaux les concernant pour permettre d'avoir une vue globale de ces différentes approches dans un état de l'art publié en 2010 dans un chapitre de livre [CL-6] portant sur plus de 90 publications. Ensuite, cet état de l'art a été mis à jour en 2011

dans une version journal [J-4] portant sur plus de 360 publications.

3- Toutes les catégories

Finalement, j'ai développé un état de l'art qui donnent une vue globale des différents modèles utilisés dans plus de 390 publications ainsi que les différents challenges et points clés dans deux chapitres du Handbook on "Background Modeling and Foreground Detection for Video Surveillance : Traditional and Recent Approaches, Benchmarking and Evaluation" [CL-2][CL-3]. Une version synthétique de cet état de l'art est présentée dans une version journal [J-1] sur plus de 270 publications.

De plus, toutes les informations collectées lors de ces différents états de l'art sont disponibles sur le site que nous avons développé et que nous avons appelé "Background Subtraction Web Site[2]". Ce site web régulièrement mis à jour contient une liste la plus exhaustive possible des publications du domaine, soit plus de 900 publications classées suivant nos états de l'art, ainsi que les ressources disponibles comme les bases de données pour une évaluation rigoureuse et large, et les implémentations.

3.2 Apport des concepts flous dans la détection d'objets mobiles

Je m'intéresse depuis plusieurs années à l'apport des concepts flous dans le domaine de l'image. Les concepts flous offrent un cadre très adapté pour les problèmes d'imprécision et d'incertitude apparaissant dans les trois principales étapes qui composent la détection d'objets mobiles par suppression de fond. Pour l'étape de modélisation du fond, il est intéressant de prendre en compte l'incertitude apparaissant en présence de fonds dynamiques pour l'estimation des paramètres des modèles existants comme la moyenne et la variance du modèle de mélange de Gaussiennes [C-10][C-11][J-5]. Pour l'étape de classification des pixels en fond/objet mobile, l'intégrale floue de Choquet permet de fusionner des résultats obtenus par différentes caractéristiques comme celles de couleur et de texture pour être plus robuste aux variations lumineuses [C-12][C-13][C-14][C-15]. Pour l'étape de mise à jour de l'image du fond, les degrés d'appartenances du pixel à la classe "fond" et à la classe "objet mobile" peuvent être employés pour mettre à jour de façon adaptative l'image du fond en évitant ainsi de se baser une décision binaire de classification du pixel à la classe "fond" ou "objets mobiles" [C-12].

Aucune méthode n'employait les concepts flous dans ce domaine quand nous avons proposé le sujet de thèse de Fida El Baf. Ces travaux ont donc été précurseurs dans le domaine. Depuis, d'autres travaux ont été développés en parallèle [620] et d'autres sont basés sur nos travaux. On peut citer entre autre les travaux de Maddalena et Petrosino [353] qui ont comparé

2. http://sites.google.com/site/backgroundsubtraction/Home

leur méthode à notre méthode de modélisation du fond par mélange de Gaussiennes floues de Type-2. L'utilisation de l'intégrale de Choquet pour fusionner les caractéristiques de couleur et texture a été reprise par Ding et al. [157][158][314][160][159]. Azab et al. [33] ont étendu son application à la fusion de trois caractéristiques (Couleur, texture et contours). Balcilar et Sonmez [48] l'ont quant à eux amélioré avec un critère spatial. D'autre part, Kashani et al. [272][273] ont repris la règle de mise à jour floue adaptative que nous avons proposé. Récemment, Calvo-Gallego et al. [87] ont amélioré cette règle de mise à jour avec un procédé plus adaptatif toujours basé sur les degrés d'appartenance à la classe "Fond" ou à la classe "Objet mobile". Finalement, Lu et al. [341] ont repris à la fois l'intégrale floue de Choquet et la règle de mise à jour en remplaçant la modélisation du fond calculée par une moyenne dans mes travaux par une modélisation du fond à partir d'un histogramme pour être plus robuste aux fonds dynamiques dans le cadre de la détection de voitures.

Un autre signe de reconnaissance internationale de ces travaux précurseurs en est la proposition de Lucia Maddalena (ICAR, Italie) de rédiger un chapitre de livre que j'ai intitulé "Background Subtraction For Visual Surveillance : A Fuzzy Approach", dans le "Handbook on Soft Computing for Video Surveillance" [CL-5].

De plus, des étudiants étrangers en Master m'ont sollicité pour les encadrer dans l'amélioration des méthodes proposées par Fida El Baf. En effet, Zhenjie Zhao (Université de Nankai, Chine) a amélioré la méthode de modélisation par mélange de Gaussiennes floues de Type-2 [C-11] en y ajoutant des champs de Markov. Enfin, Roshanak Zaman et Vahid Khorramshahi (Islamic Azad University, Iran) ont amélioré la détection proposée avec l'intégrale de Choquet dans [C-15] en modifiant le calcul de la mesure de similarité. Ces deux collaborations ont donné lieu respectivement aux publications [C-24] et [C-26].

3.3 Approche discriminative et mixte par réduction de dimension pour la modélisation de fond

Cette section correspond à une nouvelle direction de recherche que je me suis donnée après la thèse de Fida El Baf. Elle a consisté dans un premier temps en un état de l'art des méthodes de réduction de dimension employées pour la modélisation de fond. Cette étude à mis en évidence l'utilisation prépondérante de l'Analyse en Composantes Principales (ACP) et l'application d'autres méthodes reconstructives comme l'Analyse en Composantes Indépendantes (ACI). L'avantage des méthodes reconstructives est qu'elles permettent de reconstruire de façon non supervisée une image de fond en considérant que les objets mobiles ont une contribution faible lors de l'initialisation, alors que cette hypothèse est seulement vérifiée si les objets mobiles à détecter sont petits par rapport à la taille de l'image ou loin de la caméra. Enfin, ces méthodes reconstructives s'appliquent plus à reconstruire de façon non supervisée une image de fond que de fournir une bonne classification des pixels dans l'étape de

détection. Par contre, les approches discriminatives fournissent une meilleure reconstruction car ces représentations sont supervisées. Ces méthodes sont certes dépendantes de l'application, mais sont par ailleurs plus efficaces d'un point de vue implémentation. De plus, elles permettent une meilleure classification que les méthodes reconstructives.

Nous nous sommes donc intéressés, dans un premier temps, à l'apport d'une approche discriminative [C-8][C-9][J-3] et, dans un second temps, à celle d'une approche mixte pour la détection d'objets mobiles dans des séquences vidéo [C-25]. Pour l'instant, ces deux travaux sont les seuls dans le domaine à être basés sur ce type d'approches.

3.4 Approche par ACP robuste basée sur la décomposition en une matrice de rang faible et une matrice parcimonieuse

Les récentes avancées en ACP robustes initiées par Candès et al. [92] ont montré que la décomposition d'une matrice A contenant une séquence vidéo en une matrice rang faible L et une matrice parcimonieuse S, tel que $A = L + S$ permet de séparer de façon visuelle les objets mobiles du fond. La matrice de rang faible L contient le fond et la matrice parcimonieuse S contient quant à elle les objets mobiles. Une multitude de méthode ACP robustes basées sur ce principe de décomposition ont été développées ainsi que des algorithmes de résolution du problème de minimisation inhérent à cette décomposition. Dans tous ces articles, seuls des résultats qualitatifs visuels sont montrés sans aucune validation quantitative basée sur le taux de fausses détections par exemple.

Tout d'abord, notre contribution a donc consisté à une évaluation quantitative systématique de ces méthodes dans le cadre de la détection d'objets mobiles [CL-4]. Plus particulièrement, nous nous sommes intéressés à identifier et valider des variantes plus adaptées aux contraintes de la détection d'objets mobiles [C-5][C-6][C-7]. Ensuite, nous avons développé une variante rapide de l'ACP robuste optimisée par l'algorithme IRLS (Iteratively Reweighted Least Squares) [C-2][C-3][C-4]. L'objectif a consisté à pallier une limitation de la méthode originale de décomposition par Principal Component Pursuit (PCP)[92] en prenant en compte l'aspect spatial pour obtenir une estimation robuste de la matrice L de rang faible et de la matrice parcimonieuse S. Notre contribution se résume comme suit : 1) L'addition de la contrainte spatiale dans le processus de minimisation et 2) l'optimisation à partir de l'algorithme IRLS pour pondérer les paramètres de la norme $||.||_{\alpha,\beta}$ dans la décomposition.

Ensuite, nous avons développé une méthode de modélisation de fond basée sur une ACP polynomiale. L'aspect polynomial permet de prendre en compte les effets non linéaires qui apparaissent localement dans le voisinage des intensités des pixels. Ces effets non-linéaires proviennent des variations lumineuses soudaines, des variations des paramètres de la caméra (auto-ajustements) et du bruit lors de l'acquisition. Cette méthode polynomiale permet en

plus de prendre en compte l'aspect additif de certaines variations et l'aspect multiplicatif d'autres variations. De plus, lors de la minimisation, nous utilisons comme précédemment la norme $||.||_{\alpha,\beta}$ et l'algorithme IRLS.

3.5 Autres travaux de recherche

Les travaux de recherche expliqués dans les paragraphes précédents sont les principaux menés depuis ma nomination. Naturellement, d'autres travaux plus appliqués ont été menés dans le cadre de projets qui étaient soit propres au laboratoire (Aqu@thèque [C-21][C-22], Voile), soit issus d'une collaboration avec un autre laboratoire local (Analyse des jets d'air turbulents [J-CL-1][C-CL-1]) ou d'une autre université (Restauration de la continuité écologique, Mertail [C-19]). Ces recherches plus appliquées sont détaillées dans la Section 6 dédiée aux projets.

4 Activités d'Encadrement en Recherche

Cette section regroupe tous les encadrements et directions de recherche dans le cadre de thèses, de stages de Master et de stage ERASMUS que j'ai effectués depuis mon arrivée à l'Université de La Rochelle en Février 2000. Ces différents encadrements se déclinent suivant les catégories suivantes :

- **Thèse de Doctorat :** Fida El Baf (2005-2009), Charles Guyon (2010-2014), Andrews Sobral (2013- en cours) et Caroline Pacheco (2013- en cours).
- **M2 Recherche/DEA :** Théophane Langlais (2001), Zahra Hamou-Mamar (2002), Pascal Subra (2004), Stéphane Bazeille (2005), Joubert Alexandra (2005) et Julien Kerbrat (2007).
- **Mini projets de M2 Recherche/DEA :** Renan Coudray (2002), Elodie Bichot (2002), Karima Chaïeb (2003), Jida Bassem (2004), Stéphane Bazeille (2005), Lynda Aroussi (2007) et Julien Kerbrat (2007).
- **Stage Licence/Master/Doctorat (ERASMUS) :** Ces stages sont de trois niveaux : ***1) Licence :*** Alexis Soumagne, Michaël Henkens, Sakis Kiosis et Stavros Karolis ; ***2) Master :*** Doria Penciuc, Dan Jurca, Dana Felix , Sorin Patru, Diana Farcas et Cristina Marghes (Roumanie) ; ***3) Doctorat :*** Regina Radlova (République Tchéque).
- **Stage Etudiant Etranger :** Roshanak Zaman (Iran), Vahid Khorramshahi (Iran), Cristina Marghes (Roumanie) et Zhenjie Zhao (Chine).

Le tableau 4 montre une vue générale de ces encadrements ainsi que les publications relatives. Les informations relatives à ces différents encadrements sont détaillées dans les sous sections suivantes.

Catégorie	**Etudiant**	**Publication**
Thèse (1 soutenue, 3 en cours)	Fida El Baf (2005-2009)	1 Chapitre de livre, 2 Journaux, 9 Conférences
	Charles Guyon (2010-2014)	1 Chapitre de livre, 6 Conférences
	Andrews Sobral (2013- en cours)	1 Chapitre de livre, 2 Conférences
	Caroline Pacheco (2013- en cours)	-
M2 Recherche/DEA (6)	Théophane Langlais (2001)	-
	Zahra Hamou-Mamar (2002)	-
	Pascal Subra (2004)	-
	Stéphane Bazeille (2005)	1 Conférence
	Joubert Alexandra (2005)	-
	Julien Kerbrat (2007)	-
Mini projets de M2 Recherche/DEA (7)	Renan Coudray (2002)	NP
	Elodie Bichot (2002)	NP
	Karima Chaïeb (2003)	NP
	Jida Bassem (2004)	NP
	Stéphane Bazeille (2005)	NP
	Lynda Aroussi (2007)	NP
	Julien Kerbrat (2007)	NP
Stage Licence/Master/Doctorat (11) ERASMUS	Alexis Soumagne (Belgique) (2003)	-
	Michaël Henkens (Belgique) (2007)	1 Journal, 1 Conférence
	Sakis Kiosis (Crète) (2008)	-
	Stavros Karolis (Crète)(2010)	-
	Doria Penciuc (Roumanie) (2006)	1 Conférence
	Dan Jurca (Roumanie) (2007)	1 Conférence
	Dana Felix (Roumanie) (2008)	-
	Sorin Patru(Roumanie) (2008)	-
	Diana Farcas (Roumanie)(2009)	1 Journal, 1 Conférence
	Cristina Marghes(Roumanie) (2009)	1 Journal, 1 Conférence
	Regina Radlova (République Tchéque) (2006)	1 Conférence
Stage Etudiant Etranger (4) Encadrement à Distance	Roshanak Zaman (Iran) (2010)	1 Conférence
	Vahid Khorramshahi (Iran) (2010)	1 Conférence
	Cristina Marghes(Roumanie) (2012)	1 Conférence
	Zhenjie Zhao (Chine) (2012)	1 Conférence
Total	29 (+2)	3 Chapitres de livre, 4 Journaux, 23 Conférences

TABLE 4 – Encadrements de thèse, de stage recherche et de stage ERASMUS. Pour les mini-projets de M2 Recherche/DEA, il est indiqué Non Pertinent (NP) dans la colonne "Publication" car ces projets sont trop courts pour aboutir à des publications et leur objectif principal est d'initier les étudiants à la recherche bibliographique et au plus de proposer des pistes nouvelles sans les développer.

4.1 Thèses

[Doct-1] Fida El Baf

- **Titre :** Apport des concepts flous dans la détection d'objets en mouvement par suppression de fond : Application sur des séquences vidéo acquises dans des conditions non controlées.
- **Date :** Soutenue le 29 Juin 2009, Mention très honorable.
- **Durée :** 4 ans.
- **Financement :** MRT.
- **Directeur de Thèse :** Bertrand Vachon (Professeur, 10%).
- **Encadrement Scientifique : Thierry Bouwmans** (90%).
- **Président de Jury :** Sylvie Philip, Professeur, ENSEA, Cercy-Pontoise.
- **Rapporteurs :** Jenny Benois-Pineau (Professeur, LABRI, Université Bordeaux 1), François Brémond (Directeur de Recherche, INRIA, Sophia Antipolis).
- **Situation :** Post Doc - UTC (2009-2010).
- **Publications Associées :** 1 Chapitre de livre [CL-6], 2 Journaux [J-5][J-7], 9 Conférences [C-10] à [C-18].

L'objectif de cette thèse consistait à élaborer une méthode de détection d'objets en mouvement par suppression de fond robuste aux changements dynamiques pouvant se produire dans une scène filmée par une caméra fixe dans le domaine visible ou infrarouge. Tout d'abord, nous avons proposé un modèle basé sur les mélanges de Gaussiennes appelé T2-FMOG dont l'originalité est la flexibilité et la capacité d'adaptation à des environnements divers (intérieur et extérieur). Cette flexibilité est liée au caractère flou que nous avons intégré dans le modèle des mélanges de gaussiennes classiques en s'appuyant sur la théorie des sous-ensembles flous de Type-2. Ceci permet de gérer l'incertitude liée aux paramètres de ce modèle qui n'est pas prise en compte par le modèle classique. Le modèle proposé est robuste aux fonds dynamiques, aux variations lumineuses et aux ombres et notamment dans le cas de vibration de la caméra, de mouvement d'arbres, d'écoulement d'eau dans le fond et d'ondulation de la surface d'eau. Nous avons ensuite proposé dans l'étape de la classification Objet d'intérêt/Fond, l'utilisation de l'intégrale floue de Choquet pour fusionner plusieurs caractéristiques optimisant ainsi l'information portée par chacune d'elles. Le mécanisme de la fusion par cet opérateur flou améliore sensiblement la qualité de la détection d'objets. Finalement, nous avons développé une règle de mise à jour du modèle du fond qui prend en considération le degré de contribution de chaque pixel dans la classification objets d'intérêt/fond. Cette information se trouve dans le résultat de la classification obtenu par l'intégrale de Choquet. Le mécanisme proposé permet au modèle d'intégrer les changements récents qui se produisent dans la scène d'une façon adaptative. Les algorithmes proposés ont été testés pour différentes applications. Les évaluations qualitatives et quantitatives réalisées ont montré de bonnes performances par comparaison aux autres méthodes habituellement utilisées.

[Doct-2] Charles Guyon

- **Titre :** Détection d'objets mobiles par ACP robustes : Nouvelles approches et perspectives pour des séquences vidéo acquises dans des conditions non controlées.
- **Date :** Soutenance prévue en 2014.
- **Durée :** 4 ans.
- **Financement :** MRT.
- **Directeur de thèse :** El-Hadi Zahzah (Maitre de Conférences-HDR, 50%)
- **Encadrement Scientifique : Thierry Bouwmans** (50%)
- **Président de jury :** -
- **Rapporteurs :** -
- **Situation :** Doctorat.
- **Publications Associées :** 1 Chapitre de livre [CL-4], 1 Journal [J-RE-2], 6 Conférences [C-2] à [C-7].

L'objectif de cette thèse s'inscrit dans les récentes avancées par ACP robuste basée sur la décomposition en une matrice de rang faible et une matrice parcimonieuse pour les évaluer et les adapter dans le cadre la détection d'objets mobiles dans des séquences vidéo. Tout d'abord, nous avons proposé un état de l'art et une évaluation complète des méthodes existantes sur les trois récentes bases de données que sont ChangeDetection.net, SABS et BMC 2012.

Nous avons ensuite proposé un algorithme de décomposition résolue par un algorithme de type IRLS (Iteratively Reweighted Least Squares). De plus, nous avons ajouté un terme de régularisation lors de l'étape de minimisation pour prendre en compte les contraintes spatiales et temporelles que présentent les objets mobiles. Le modèle proposé est ainsi robuste aux fonds dynamiques et aux variations lumineuses comme le montre les évaluations sur la base de données BMC 2012.

Finalement, nous avons développé une méthode d'ACP robuste polynomiale pour prendre en compte les effets non linéaires qui apparaissent localement dans le voisinage des intensités des pixels. Ces effets non-linéaires proviennent des variations lumineuses soudaines, des variations des paramètres de la caméra (auto-ajustements) et du bruit lors de l'acquisition. Cette méthode polynomiale permet en plus de prendre en compte l'aspect additif de certaines variations et l'aspect multiplicatif d'autres variations.

[Doct-3] Andrews Sobral

- **Titre :** Détection d'objets mobiles dans des séquence vidéos par décomposition tensorielle robuste
- **Date :** 01/09/2013- En Cours
- **Durée :** 3 ans
- **Financement :** Science sans frontiéres - Ciencia sem Fronteiras (CSF)

— **Directeur de thèse :** El-Hadi Zahzah (Maitre de Conférences-HDR, 50%)
— **Encadrement Scientifique : Thierry Bouwmans** (50%)
— **Publications Associées :** 1 Chapitre de livre [CL-1], 2 Conférences [C-1][C-23]

L'objectif de cette thèse s'inscrit dans la suite de la thèse de Charles Guyon. Dans les approches ACP robustes, la décomposition en matrice de rang faible et parcimonieuse, et l'exploitation qui en est faite ne permet pas de prendre en compte les contraintes spatiales et temporelles des objets mobiles. L'idée est donc d'utiliser une décomposition tensorielle pour développer une approche spatio-temporelle robuste. Ensuite, il sera développé une méthode incrémentale et temps-réel.

[Doct-4] Caroline Pacheco do E. Silva

— **Titre :** Sélection de caractéristiques pour la détection d'objets mobiles par soustraction de fond
— **Date :** 01/09/2013- En Cours
— **Durée :** 3 ans
— **Financement :** MRT
— **Directeur de thèse :** Carl Frélicot (PR, 50%)
— **Encadrement Scientifique : Thierry Bouwmans** (50%)

La couleur dans l'espace RGB est la caractéristique qui est la plus employée car elle est directement fournie par les caméras. Cependant, l'espace RGB comporte plusieurs inconvénients : les trois composantes sont dépendantes et sensibles aux variations lumineuses. Des espaces alternatifs comme YUV or YCrCb peuvent être employés. Bien que la couleur soit souvent discriminative pour détecter les objets, elle présente plusieurs limitations en présence de certains challenges comme les variations lumineuses, le camouflage et les ombres. Pour résoudre cela, certains auteurs ont employé d'autres caractéristiques comme le contour ou la texture en plus de la couleur. L'avantage des contours apparaît lors des variations lumineuses et dans le cas de détection "ghost" quand un objet à l'arrêt commence à se déplacer. Dans le cas des caractéristiques stéréoscopiques, deux caméras doivent être utilisées pour obtenir l'information de disparité ou de profondeur. Celles-ci sont efficaces quand il y a camouflage en couleur. L'information de mouvement qui est obtenue par flot optique permet quant à elle de gérer les fonds dynamiques. Par contre, son temps de calcul en limite son usage. Enfin, la texture est appropriée quand il y a des variations lumineuses et des ombres.

Dans un premier temps, l'objectif de cette thèse consiste donc à une étude approfondie de ces différentes caractéristiques. Ensuite, il sera développé une méthode de sélection de caractéristiques permettant d'employer les caractéristiques les plus adaptées pour une zone de l'image où apparait tel ou tel challenge.

4.2 Master 2/DEA

Depuis la création du DEA Image et Calculs à l'Université de La Rochelle en 2000, j'ai encadré les stagiaires de 3ième cycle suivants.

1- Stages de M2 Recherche/DEA

[M2-1] Théophane Langlais

- **Titre :** Détection automatique de lignes de visualisation sur les voiles .
- **Dates :** 01 Février 2001- 30 Juin 2001.
- **Encadrement : Thierry Bouwmans** (50%), Pierre Courtellemont (50%).
- **Type Activités :** Voile.
- **Contrat :** Société JPS.
- **Situation :** Chef d'entreprise PME.

[M2-2] Zahra Hamou-Mamar

- **Titre :** Reconstruction 3D de voile à partir de vues stéréoscopiques prises en soufflerie.
- **Dates :** 01 Février 2002 - 30 Juin 2002.
- **Encadrement : Thierry Bouwmans** (90%), Pierre Courtellemont (10%).
- **Type Activités :** Voile.
- **Contrat :** Société JPS.
- **Poursuite Etudes :** Doctorat - Université Blaise Pascal - Clermont II - 2008.

[M2-3] Pascal Subra

- **Titre :** Utilisation des informations contenues dans les blocs DCT pour la détection d'objets mobiles dans les vidéos.
- **Dates :** 01 Février 2004 - 30 Juin 2004.
- **Encadrement : Thierry Bouwmans** (50%), Pierre Courtellemont (50%)
- **Type Activités :** Détection d'objets mobiles.

[M2-4] Stéphane Bazeille

- **Titre :** Traitement et Analyse d'images sous-marines pour la détection de langoustines.
- **Dates :** 01 Février 2005 - 30 Juin 2005.
- **Encadrement :** Thierry Bouwmans (50%), Jean-Chrisophe Burie (50%).

— **Type Activités :** MERTAIL-TOPVISION.
— **Poursuite Etudes :** Doctorat - Université de Bretagne Occidentale - 2008.
— **Situation Actuelle :** Maître de Conférence à l'Université de Bretagne Occidentale.
— **Publications Associées :** 1 Conférence [C-19].

[M2-5] Joubert Alexandra

— **Titre :** Etude d'algorithmes efficaces de calcul de medianes temporelles pour la segmentation de sequences d'images.
— **Dates :** 01 Février 2005 - 30 Juin 2005.
— **Encadrement :** Laurent Mascarilla (40%), Christophe Saint-Jean (40%), **Thierry Bouwmans** (20%).
— **Type Activités :** Détection d'objets mobiles.
— **Situation Actuelle :** Ingénieur d'études - CNBDI - Angoulême.

[M2-6] Julien Kerbrat

— **Titre :** Sélection de variables pour la détection d'objets en mouvement par soustraction de fond
— **Dates :** 01 Février 2007 - 30 Juin 2007.
— **Encadrement : Thierry Bouwmans** (100%).
— **Type Activités :** Détection d'objets mobiles.

2- Mini projets de M2 Recherche/DEA

[P2-1] Renan Coudray

— **Titre :** Utilisation de l'information de mouvement dans la segmentation d'images.
— **Dates :** 01 Octobre 2001 - 30 Janvier 2002.
— **Encadrement : Thierry Bouwmans** (100%).
— **Type Activités :** Détection d'objets en mouvement - Aqu@thèque.
— **Poursuite Etudes :** Doctorat - Université de La Rochelle.
— **Situation Actuelle :** Ingenieur Recherche et Developpement.

[P2-2] Elodie Bichot

— **Titre :** Utilisation des contours actifs pour la segmentation de séquences vidéo.
— **Dates :** 01 Octobre 2001 - 30 Janvier 2002.
— **Encadrement : Thierry Bouwmans** (100%).
— **Type Activités :** Détection d'objets mobiles - Aqu@thèque.

— **Poursuite Etudes :** Doctorat - Université de La Rochelle - 2006.

[P2-3] Karima Chaïeb

— **Titre :** Méthodes d'auto-calibration dans les systèmes stéréoscopiques.
— **Dates :** 01 Octobre 2002 - 30 Janvier 2003.
— **Encadrement : Thierry Bouwmans** (100%).
— **Type Activités :** Voile.
— **Poursuite Etudes :** Doctorat - Université de Poitiers.

[P2-4] Jida Bassem

— **Titre :** Seuillage adaptatif pour la segmentation de séquences d'images acquises dans un environnement aquatique.
— **Dates :** 01 Octobre 2003 - 30 Janvier 2004.
— **Encadrement : Thierry Bouwmans** (100%).
— **Type Activités :** Détection d'objets mobiles.
— **Poursuite Etudes :** Doctorat - Université du Littoral Cote d'Opale - 2008

[P2-5] Stéphane Bazeille

— **Titre :** Traitement et Analyse d'images sous-marines.
— **Dates :** 01 Octobre 2004 - 30 Janvier 2005.
— **Encadrement : Thierry Bouwmans** (50%), Jean-Christophe Burie (50%).
— **Type Activités :** MERTAIL-TOPVISION.
— **Poursuite Etudes :** Doctorat - Université de Bretagne Occidentale - 2008.
— **Situation Actuelle :** Maître de Conférence à l'Université de Bretagne Occidentale.
— **Publication Associée :** 1 Conférence [C-19].

[P2-6] Lynda Aroussi

— **Titre :** Modélisation de fond à l'aide de filtres pour la détection d objets en mouvement.
— **Dates :** 01 Octobre 2006 - 30 Janvier 2007.
— **Encadrement : Thierry Bouwmans** (100%).
— **Type Activités :** Détection d'objets mobiles.
— **Poursuite Etudes :** Master Management de Projet.

[P2-7] Julien Kerbrat

— **Titre :** Modélisation de fond à l'aide de la méthode de classification dite de Support Vector Machine pour la détection d' objets en mouvement.
— **Dates :** 01 Octobre 2006 - 30 Janvier 2007.
— **Encadrement : Thierry Bouwmans** (100%)
— **Type Activités :** Détection d'objets mobiles.

4.3 Stages de Master 1/Maîtrise

[P1-1]

— **Titre :** Détection de poissons dans une séquence d'images.
— **Dates :** 2006-2007.
— **Encadrement : Thierry Bouwmans** (100%).
— **Type Activités :** Collaboration LEA.

[P1-2] : Alexandre Lebon

— **Titre :** Développement d'un outil informatique pour l'analyse d'images vidéo.
— **Dates :** 2006-2007.
— **Encadrement : Thierry Bouwmans** (50%), Christophe Saint-Jean (50%)
— **Type Activités :** Collaboration IFREMER.

4.4 Stage LMD ERASMUS

J'ai encadré également des étudiants bénéficiant de financement européen ERASMUS tant au niveau Licence qu'au niveau Master et au niveau Doctorat.

ERASMUS Niveau Licence

[ERL-1] Alexis Soumagne (Haute Ecole Rennequin Sualem, Belgique)

— **Titre :** Analyse et élaboration d'une application temps réel de reconnaissance automatique de poissons
— **Dates :** 01 Février 2003 - 30 Juin 2003.
— **Encadrement : Thierry Bouwmans** (70%), Laurent Mascarilla (20%), Pierre Courtellemont (10%)
— **Type Activités :** Détection d'objets mobiles

[ERL-2] Michaël Henkens (Haute Ecole Rennequin Sualem, Belgique)

— **Titre :** Analyse de la dynamique des tourbillons d'un jet d'air par Imagerie Laser Haute Cadence
— **Dates :** 01 Février 2007 - 30 Juin 2007.
— **Encadrement : Thierry Bouwmans** (90%), Amina Meslem (10%).
— **Type Activités :** Collaboration LEPTIAB.
— **Publications Associées :** 1 Journal [J-CL-1], 1 Conférence [C-CL-1].

[ERL-3] Sakis Kiosis (University of Heraklion, Crete, Grêce)

— **Titre :** Analyse de la dynamique des tourbillons de jets d'air par détection de contours
— **Dates :** 01 Février 2008 - 30 juin 2008.
— **Encadrement : Thierry Bouwmans** (90%), Amina Meslem (10%)
— **Type Activités :** Collaboration LEPTIAB.

[ERL-4] Stavros Karolis (University of Heraklion, Crete, Grêce)

— **Titre :** Dynamic Control of Adaptive Mixture of Gaussians Background Model.
— **Dates :** 01 Février 2010 - 30 Juin 2010.
— **Encadrement : Thierry Bouwmans** (100%).
— **Type Activités :** Détection d'objets mobiles.

ERASMUS Niveau Master

Tous les étudiants ERASMUS de niveau Master que j'ai encadré viennent de l'Université de Timisoara en Roumanie. Ils ont tous travaillé dans le cadre de mon activité sur la détection d'objets en mouvement.

[ERM-1] Doria Penciuc

— **Titre :** Implémentation et comparaison de différentes méthodes de détection d'objets en mouvements dans le cadre de la capture de mouvement sans marqueurs.
— **Dates :** 01 Mars 2006 - 30 Juin 2006.
— **Encadrement :** Thierry Bouwmans (100%).
— **Situation :** Développeur Java Web (USA)
— **Publications Associées :** 1 Conférence [C-18].

[ERM-2] Dan Jurca

— **Titre :** Implémentation d'une méthode de détection d'objets en mouvement utilisant l'intégrale floue
— **Dates :** 01 Mars 2007 - 30 Juin 2007.

— **Encadrement : Thierry Bouwmans** (50%), Fida El baf (50%).
— **Situation :** Ingénieur télécommunication indépendant (Roumanie)
— **Publications Associées :** 1 Conférence [C-14].

[ERM-3] Dana Felix

— **Titre :** Implementation et comparaison de méthodes de détection d'objets en mouvement par réduction de dimension (ACP).
— **Dates :** 01 Mars 2008 - 30 Juin 2008.
— **Situation :** Ingénieur Alcatel-Lucent (Autriche)
— **Encadrement : Thierry Bouwmans** (100%).

[ERM-4] Sorin Patru

— **Titre :** Détection d'objets en mouvements par Support Vector Regression
— **Dates :** 01 Mars 2008 - 30 Juin 2008.
— **Situation :** Ingénieur télécommunication (Roumanie)
— **Encadrement : Thierry Bouwmans** (100%).

[ERM-5] Diana Farcas

— **Titre :** Subspace Learning using Incremental Maximum Margin Criterion (IMMC) for moving objects detection in video sequences.
— **Dates :** 01 Mars 2009 - 30 Juin 2009.
— **Encadrement : Thierry Bouwmans** (100%)
— **Situation :** Ingénieur Alcatel-Lucent (Roumanie)
— **Publication :** 1 Journal [J-3], 1 Conférence [C-9].

[ERM-6] Cristina Marghes

— **Titre :** Subspace Learning using Incremental Rank-(R1,R2,R3) Tensor (IRT) for moving objects detection in video sequences.
— **Dates :** 01 Mars 2009 - 30 Juin 2009.
— **Encadrement : Thierry Bouwmans** (100%)
— **Situation :** Ingénieur Alcatel-Lucent (Roumanie, 2 ans), Ingénieur Développement Logiciel en Computer Vision (MOVIDIUS)
— **Publications Associées :** 1 Journal [J-3], 1 Conférence [C-8]

ERASMUS Niveau Doctorat

[ERR-1] Regina Radlova (Université de Plzen, République Tchèque)

- **Titre :** Etat de l'art sur les modèles utilisés dans la capture du mouvement.
- **Encadrement : Thierry Bouwmans** (50%), Bertrand Vachon (50%).
- **Dates :** 01 Septembre 2005 - 30 Janvier 2006.
- **Type Activités :** Capture de Mouvement.
- **Publication Associée :** 1 Conférence [C-20].

4.5 Stage Etudiant Etranger

J'ai encadré également des étudiants étrangers à distance. Ceux-ci m'ont sollicité pour des renseignements sur les méthodes developpées au MIA pour les améliorer et m'ont demandé de les encadrer durant leur stage.

[EE-1] Roshanak Zaman et Vahid Khorramshahi (Islamic Azad University, Iran)

- **Titre :** Foreground object detection using Choquet fuzzy integral in video images
- **Dates :** 01 Mars 2010 - 30 Juin 2010.
- **Encadrement : Thierry Bouwmans** (100%).
- **Situation :** Enseignant à l'Islamic Azad University.
- **Publications Associées :** 1 Conférence [C-26].

[EE-2] Cristina Marghes (Université de Timisoara, Roumanie)

- **Titre :** Mixed Subspace Learning for moving objects detection in video sequences
- **Dates :** 01 Mars 2011 - 30 Juin 2011.
- **Encadrement :** Radu Vasiu (50%), **Thierry Bouwmans** (50%).
- **Situation :** Ingénieur Alcatel-Lucent (2 ans), Ingénieur Développement Logiciel en Computer Vision (MOVIDIUS)
- **Publications Associées :** 1 Conférence [C-25].

[EE-3] Zhenjie Zhao (Université de Nankai, Chine)

- **Titre :** A Fuzzy Background Modeling Approach for Motion Detection in Dynamic Backgrounds
- **Dates :** 01 Mars 2012 - 30 Juin 2012.
- **Encadrement :** Xuebo Zhang (40%), Yongchun Fang (10%), **Thierry Bouwmans** (50%).
- **Situation :** Doctorant à l'Université de Nankai
- **Publications Associées :** 1 Conférence [C-24].

5 Activités Administratives Collectives

5.1 Etablissement

— Membre suppléant de la Commission de Spécialistes 27ième Section et 61ième Section (2002-2006).
— Membre de la commission de recrutement des ATER 27ième Section et 61ième Section (2013)

5.2 Département Informatique

Mes activités administratives et pédagogiques sont celles citées dans la Section "Responsabilité Administrative (ULR)".

5.3 Laboratoire L3i/Laboratoire MIA

— Membre élu du Conseil de Laboratoire du L3i (2002-2006).
— Responsable scientifique de la partie "Détection des poissons" dans le projet Aquathéque (2000-2002).
— Représentant du laboratoire MIA au Conseil du Département Informatique (2010- en cours).

6 Activité dans des Projets

[PROJ-1] Aqu@thèque

— **Projet - Laboratoire :** Aqu@thèque (L3i).
— **Dates :** 2000-2003.
— **Partenaires :** 2d3D Animations - Angoulême, E.C.O.L.E. de la mer (Espace de Culture Océane du Littoral et de l'Environnement, Association issue du partenariat entre la région Poitou-Charentes et l'Aquarium de La Rochelle), ANVAR Région Poitou-Charentes.
— **Moyens :** Contrat d'un montant de 370 000 Euros au bénéfice de la société 2d3D Animations ; thèse, stage de DRT. Aide au transfert technologique de l'ANVAR de 60 000 Euros au bénéfice du L3i. Contrat de transfert technologique avec la société 2d3D-Animations de 71 000 Euros.
— **Responsable L3i :** Pierre Courtellemont, Carl Frélicot, Laurent Mascarilla, **Thierry Bouwmans**.

FIGURE 1 – Projet Aquatheque

— **Descriptif :** Le projet Aqu@théque consiste à la conception d'une borne interactive destinée aux aquariums. Ce projet se divise en un systéme de reconnaissance et un système de modélisation comportementale des poissons. Le système de reconnaissance est composée d'un écran tactile qui montre en temps réel les images filmées par une caméra vidéo fixe orientée vers un bassin. Lorsque le visiteur sélectionne sur cet écran tactile (Figure 1), le système de reconnaissance identifie automatiquement et en temps réel l'espéce à laquelle le poisson sélectionné appartient et fournit un ensemble d'informations pédagogiques concernant cette espéce. De plus, une fois que le poisson est identifié, le système de modélisation comportementale propose sur un autre écran un choix d'espéces appartenant au même milieu que le poisson d'intérêt comme ces prédateurs et ces proies. Le visiteur choisit les espéces qu'il souhaite voir évoluer dans cet aquarium virtuel. L'environnement évolue alors à partir du choix et l'animation est ainsi déterminée par le comportement des espéces en présence.
— **Travail Scientifique :** Prototype d'un système de détection de poissons fondé sur la soustraction de fond à partir d'un flux vidéo issu d'une caméra filmant un bassin de l'Aquarium de la Rochelle.
— **Mon Rôle Scientifique :** Responsable scientifique de la partie détection des poissons dans les séquence vidéos, participation aux travaux de la thèse de Dahbia Semani.
— **Publications Associées :** 2 Conférences [C-21][C-22].

[PROJ-2] Voile

— **Projet - Laboratoire :** Voile (L3i).
— **Dates :** 2000-2005.
— **Partenaires :** Société JPS et BSG DEVELOPPEMENTS.
— **Moyens :** Contrat de DEA de Théophane Langlais (7600 euros), Contrat de thèse CIFRE de Théophane Langlais (15 000 Euros), Contrat de DEA de Zahra Zahra Hamou-Mamar (6000 Euros).
— **Responsable L3i :** Pierre Courtellemont (50%), **Thierry Bouwmans** (50%).

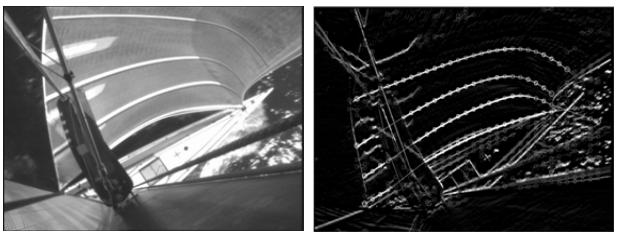


FIGURE 2 – Projet Voile

— **Descriptif :** Les voiles constituent un des éléments fondamentaux dans la performance des bateaux de course. Jusqu'a présent, il était cependant très difficile de les appréhender. En effet, la complexité des réglages d'un gréement et leur nature fortement dynamique compliquaient considérablement la tâche. L'objectif de cette recherche a donc été l'étude et le développement d'un système de vision permettant, en temps réel, de connaître avec exactitude la forme des voiles et la conception d'un système de vision stéréoscopique dédié permettant la reconstruction 3D de voiles, pour une comparaison temps réel d'un réglage de voile avec une situation de navigation déjà référencée. La Figure 2 montre l'extraction des lignes correspondant à des transitions entre des bandes blanches et noires posées sur la voile permettant ainsi de caractériser sa forme.
— **Travail Scientifique :** Prototype d'un système de détection de lignes dans les voiles et d'un système stéréoscopiques de reconstruction 3D de voiles.
— **Mon Rôle Scientifique :** Co-encadrant scientifique de la thèse de Théophane Langlais et du stage de Zahra Hamou-Mamar.
— **Publication Associée :** Projet Confidentiel !

[PROJ-3] Mertail

— **Projet -Laboratoire :** Mertail (L3i).
— **Dates :** 2004-2007.
— **Partenaires :** LEMH, laboratoire d'Ecologie et modèles pour l'halieutique, IFREMER, Nantes, France.
— **Responsable :** Verena Trenkel (IFREMER, Nantes).
— **Responsable L3i : Thierry Bouwmans** (50%), Jean-Christophe Burie (50%).
— **Descriptif :** La campagne MERTAIL[3] organisée par l'IFREMER avait pour objectif d'évaluer des outils technologiques permettant de mesurer les dimensions des organismes marins mobiles à partir d'un engin sous-marin. Les images de cette campagne représentent des fonds vaseux situés à environ 100 m de profondeur. On y distingue différentes espèces comme des crabes, des galatées, des langoustines, etc. Cependant la plupart du temps, seuls des trous sont visibles. Il peut s'agir de terriers de langoustines, de crabes ou de simples trous (Figure 3). L'objectif a donc consisté à développer

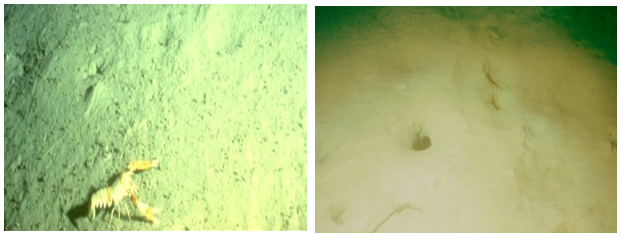


FIGURE 3 – Projet Mertail

une méthode de détection de terriers dans le cas complexe des images sous-marines. En effet, ces images sont souvent de qualité moyenne. D'une part l'éclairage artificiel n'est pas uniforme. D'autre part la présence de nombreuses particules en suspension atténue les couleurs et les contraste. De plus, une nuance bleutée apparaît sur beaucoup d'images. Afin de corriger la non uniformité de l'éclairage, un filtre homomorphique est appliqué en pré-traitement ainsi qu'une correction colorimétrique. Ensuite, la détection des terriers est effectuée.

— **Travail Scientifique :** Détection et comptage de langoustines et de leurs nids à partir de séquences vidéos sous- marines.
— **Publication Associée :** 1 Conférence [C-19].

[PROJ-4] Analyse des jets d'air turbulents

— **Projet - Laboratoire :** Analyse des jets d'air turbulents par traitements d'images. Application aux jets de climatisation du bâtiment et du véhicule de transport (MIA).
— **Dates :** 2007-2009.
— **Partenaires :** LEBTIAB - Laboratoire d'Étude des Phénomènes de Transfert et de l'Instantanéité : Agro-industrie et Bâtiment, Université de La Rochelle, France.
— **Moyens :** Contrat BQR (Bonus Qualité Recherche) - 7000 Euros.
— **Responsable : Thierry Bouwmans** (50%), Amina Meslem (MCF, LEPTIAB)(50%).
— **Descriptif :** Ce projet s'intéresse en à la compréhension et à l'amélioration du phénomène de mélange fluide primaire/fluide ambiant, par des moyens passifs utilisant des formes différentes de buses de soufflage. Un système PIV classique permet d'obtenir des coupes transversales et longitudinales des écoulements. L'objectif a donc été de concevoir une méthode de traitement et d'analyse de séquences d'images de jets turbulents de formes simples ou complexes en vue d'extraire des informations quantitatives relatives des dimensions du jet dans le plan longitudinal et dans le plan transversal. De plus, un outil logiciel utilisant ces méthodes de traitement de séquences d'images a permis à l'échéance du projet, une interprétation quantitative aisée et rapide de la capacité d'auto-induction des jets d'air. La Figure 4 montre, de gauche à droite, une image dans un plan longitudinal d'un jet turbulent, l'image segmentée ob-

3. http://www.ifremer.fr/sismer/FR/catal/campagne/campagne.htql?crno=4480090

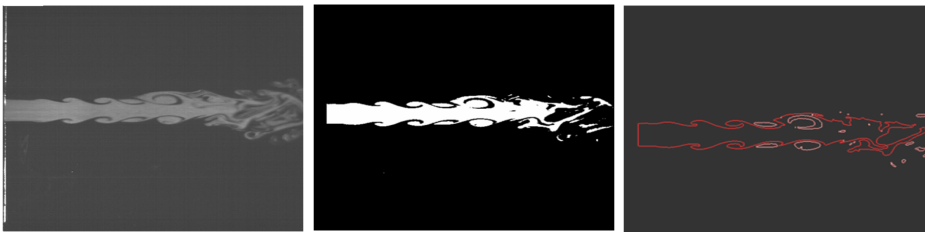


FIGURE 4 – Analyse des jets d'air turbulents par traitements d'images

tenue par seuillage et l'image contour à partir de laquelle les différentes hauteurs du jet peuvent être déterminées afin de le caractériser.

— **Travail Scientifique :** Développement de méthodes de traitement et d'analyse de séquences d'images de jets turbulents de formes simples ou complexes en vue d'extraire des informations quantitatives. Dans ce cadre, des mesures de variations spatio-temporelles de structures du jet circulaires ont été extraites. Cela a été intégré dans un outil logiciel appelé QuantiVisu.
— **Publications Associées :** 1 Journal [J-CL-1], 1 Conférence [C-CL-1].

[PROJ-5] Restauration de la continuité écologique

— **Projet - Laboratoire :** Restauration de la continuité écologique des cours d'eau à poissons migrateurs (MIA).
— **Dates :** 2007-2009.
— **Partenaires :** LEA - Laboratoire d'Etudes Aérodynamiques, UMR 6609, CNRS, Université de Poitiers, France.
— **Moyens :** CPER (Contrat Plan Etat Région) - 100 000 Euros.
— **Responsable :** Laurent David (Professeur, LEA, Poitiers).
— **Responsable MIA : Thierry Bouwmans**, Laurent Mascarilla, Carl Frélicot.
— **Travail Scientifique :** L'objectif du projet est de proposer des solutions de contrôle des écoulements dans les dispositifs de montaison à proximité des ouvrages pour favoriser le passage de toutes les espèces de poisson et de les valider avec des poissons à partir de tests effectués sur un modèle à l'échelle $\frac{1}{4}$ et sur une passe réelle. Afin de faciliter le traitement des données, la mise en place d'un logiciel de suivi des trajectoires des poissons et de comptage est indispensable. Dans ce cadre, ma contribution a consisté à développer une méthode de détection de poissons dans des séquences vidéos issue de passes à poissons localisées près d'ouvrages fluviaux (Figure 5). Ensuite, une méthode de suivi a été développée par les autres membres du projet.

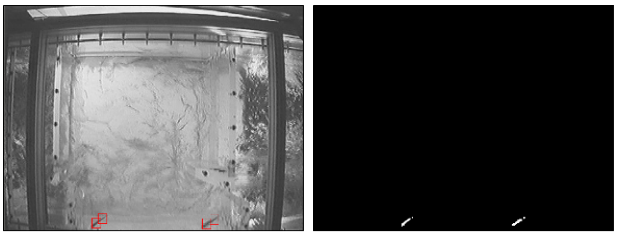


FIGURE 5 – Restauration de la continuité écologique

7 Rayonnement Scientifique

1- Conférence Invitée

— [CI-1] T. Bouwmans, "Background Modeling and Foreground Detection for Video Surveillance : Recent Advances and Future Directions", International Workshop on Background Models Challenge, ACCV 2012, Daejon, Korea, November 2012.

Cette présentation a consisté à présenter les récentes avancées dans le domaine de la détection d'objets mobiles par suppression de fond. Dans un premier temps, j'ai présenté les étapes clés ainsi qu'un état de l'art concis. J'ai ensuite décliné les nouvelles approches comme suit : les approches floues, les approches par les méthodes de réduction discriminative et mixte, et les approches par les récentes avancées en ACP robustes.

2- Activités Editoriales

— [AE-1] T. Bouwmans, L. Davis *(Univ. of Maryland, USA)*, J. Gonzàlez *(Univ. Barcelona, Spain)*, M. Piccardi *(Univ. Sydney, Australia)*, C. Shan *(Philips Research, The Netherlands)*, **Special Issue on "Background Modeling for Foreground Detection in real-world dynamic scenes"**, Machine Vision and Applications, July 2014.

— [AE-2] T. Bouwmans, F. Porikli *(MERL, USA)*, B. Höferlin *(Univ. Stuttgart, Deutschland)*, A. Vacavant *(Univ. Clermont-Ferrand, France)*, **Handbook on "Background Modeling and Foreground Detection for Video Surveillance : Traditional and Recent Approaches, Benchmarking and Evaluation"**, CRC Press, Taylor and Francis Group, 25 chapitres, June 2014.

Je suis le "Lead Guest Editor" de ces deux ouvrages dont je suis l'initiateur et le porteur. L'objectif du numéro spécial était de faire émerger de nouveaux travaux sur la détection d'objets mobiles par soustraction de fond. L'objectif du livre était de regrouper les travaux majeurs des équipes leader du domaine pour les chercheurs et les professionels en entreprise.

Pour chaque ouvrage, j'ai contacté les "co-guest editors" et les éditeurs (MVA et CRC Press) en soumettant la proposition d'ouvrage. Pour le numéro spécial, j'ai effectué l'affec-

tation des reviewers et le suivi jusqu'à la décision finale. Pour le livre, j'ai invité une partie des auteurs et j'ai assuré le suivi des 25 chapitres jusqu'à la publication. De plus, j'ai réalisé le site web[4] qui accompagne cet ouvrage et qui permet d'avoir accès aux résumés des 25 chapitres ainsi qu'à toutes les ressources citées dans le livre.

4. https://sites.google.com/site/backgroundmodeling/

3- Relecture pour des Revues Internationales

Nombre d'articles expertisés par années : 11 (2011), 15 (2012), 36 (2013), 55 (au 01/09/2014) pour les revues suivantes :

— IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)
— IEEE Transactions on Circuits and Systems for Video Technology (T-CSVT)
— IEEE Transactions on Image Processing (T-IP)
— IEEE Transactions on Systems, Man and Cybernetics (T-SMC)
— IEEE Transactions on Fuzzy Systems (T-FS)
— IEEE Transactions on Information Forensics and Security (T-IFS)
— IEEE Signal Processing Letters (SPL)

— SPRINGER International Journal of Computer Vision (IJCV)
— SPRINGER Machine Vision and Applications (MVA)
— SPRINGER Pattern Analysis and Applications (PAA)
— SPRINGER Signal, Image and Video Processing (SIVP)
— SPRINGER Soft Computing (SOCO)
— SPRINGER Neural Computing and Applications (NCA)

— ELSEVIER Image and Vision Computing (IVC)
— ELSEVIER Computer Science Review (COSREV)
— ELSEVIER Pattern Recognition Letters (PRL)
— ELSEVIER Fuzzy Sets and Systems (FSS)
— ELSEVIER Computer Vision and Image Understanding (CVIU)
— ELSEVIER Digital Signal Processing (DSP)
— ELSEVIER Signal Processing (SP)

— IET Computer Vision (IET-CV)

Expertises demandées : Modélisation de fond, suivi, concepts flous, etc..

4- Relecture pour des Conférences Internationales

2015

— IEEE Winter Conference on Applications of Computer Vision, WACV 2015, Waikoloa Beach, Hawaii, Janvier 2015.

2014

— International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2014, Seoul, Corée du Sud, 2014.
— International Conference on Image Processing, ICIP 2014, Paris, France, 2014.
— International Conference on Pattern Recognition, ICPR 2014, Stockholm, Suède, 2014.

2013

— International Conference on Image Analysis and Processing, ICIAP 2013, Naples, Italie, Septembre 2013.

2007

— International Conference on Computer Games : AI, Animation, Mobile, Educational and Serious Games, CGAMES 2007, La Rochelle, France, Octobre 2007.

Expertises demandées : Modélisation de fond, capture de mouvement, etc...

5- Area Chair pour des Conférences Internationales

— IEEE Winter Conference on Applications of Computer Vision, WACV 2015[5], Waikoloa Beach, Hawaii, Janvier 2015.

6- Comités de Programme de Workshops Internationaux

— ACM International Workshop on Analysis and Retrieval of Tracked Events and Motion in Imagery Streams, ARTEMIS 2010[6], Firenze, Italy, Octobre 2010.

7- Comités d'Organisation de Workshops Internationaux

— T. Bouwmans, E. Zahzah *(Univ. La Rochelle, France)*, B. Wohlberg *(Los Alamos University, USA)*, J. Wright *(Columbia University, USA)*, C. Castellanos Dominguez *(Univ. Manizales, Colombie)*, **"When robust subspace learning meet background modeling"**, pour être soumis à ICCV 2015.
— A. Vacavant *(Univ. d'Auvergne, France)*, L. Tougne *(Univ. Lyon 2, France)*, C. Spampinato *(Univ. Catania, Italie)*, L. Salgado *(Univ. Madrid, Espagne)*, L. Robinault *(Foxtream, France)*, T. Chateau *(Univ. Blaise Pascal, France)*, M. Camplani *(Univ. Madrid, Espagne)*, T. Bouwmans, **"Background Model Challenges"**, soumis à ECCV 2014 et rejeté, en attente de soumission à une autre conférence.

5. http://wacv2015.org/
6. http://iselab.cvc.uab.es/crew-artemis

8- Implication dans des Structures Nationales

— T. Bouwmans, T. Chateau *(Univ. Blaise Pascal, France)*, L. Tougne *(Univ. Lyon 2, France)*, A. Vacavant *(Univ. d'Auvergne, France)*, F. Brémond (INRIA, Sophia Antipolis, France), E. Zahzah, **"Détection d'objets mobiles par suppression de fond : Récentes avancées et perspectives"**, pour être soumis en Journées GDR ISIS en 2015.

9- Implication dans des Structures Régionales

— **PRIDES (Pôle Régional de Recherche en Images Données et Systèmes) :** Fédération de laboratoires de la région Poitou-Charente dont les activités se déclinent suivant les thèmes ci-dessous : Images, Données et Systèmes. Ma participation s'est effectuée au niveau du sous-thème Mouvements, Textures, Couleurs. Dans ce cadre, j'ai participé régulièrement aux réunions de travail organisées à Poitiers et La Rochelle. De façon plus formelle, mes travaux ont été exposés durant les manifestations suivantes : 1) Assemblée Générale en Juin 2006 : Exposé de Fida El Baf [38], 2) Assemblée Générale en Juillet 2008 : Exposé de Fida El Baf [41].

— **MIRES (Mathématiques et leurs Interactions, Images et information numérique, Réseaux et Sécurité) :** Fédération de laboratoires du PRES Limousin-Poitou-Charentes.

10- Organisation de Journées

— T. Bouwmans, E. Zahzah, "Détection d'objets mobiles par des méthodes de suppression de fond"', Laboratoire MIA, Univ. La Rochelle, 22 Mai 2014.

11- Participation à des jurys de thèses

J'ai participé à différents jurys de thèse en tant que :

— *Examinateur* de la thèse de Dahbia Semani, "Une méthode supervisée de sélection et discrimination avec rejet - Application à la synthèse d'un aquarium virtuel", Université de La Rochelle, Mai 2004. Encadrants : P. Courtellemont, C. Frélicot (Univ. La Rochelle).
— *Encadrant scientifique* de la thèse de Fida El Baf, "Apport des concepts flous dans la détection d'objets en mouvement par suppression de fond : Application sur des séquences vidéo acquises dans des conditions non contrôlées", Université de La Rochelle, Juin 2009. Encadrants : T. Bouwmans, B. Vachon (Univ. La Rochelle).
— *Membre invité* de la thèse de Nadeem Salamat, "Modélisation des relations spatiales entre objets en mouvement", Université de La Rochelle, Octobre 2011. Encadrant : E. Zahzah (Univ. La Rochelle).

	All	Since 2009
Citations	625	590
h-index	11	10
i10-index	12	11

TABLE 5 – Indices bibliométriques (Source Google Scholar au 01/09/2014)

12- Participation à des jurys de thèses hors Etablissement

— *Evaluateur pour la mention européenne* de la thèse de Noha Elfiky, "Compact, Adaptive and Discriminative Spatial Pyramids for Improved Object and Scene Classification Compact, Adaptive and Discriminative Spatial Pyramids for Improved Object and Scene Classification", Universitat Autonoma de Barcelona, Espagne, Mars 2012. Encadrant : J. Gonzàlez (CVC, Universitat Autonoma de Barcelona, Espagne).

— *Evaluateur extérieur* de la thèse de Munir Shah, "Adaptive Foreground-Background Segmentation of Complex Video Scenes", Univ. of Otago, Nouvelle-Zélande, Septembre 2013. Encadrants : J. Deng, B. Woodford (Univ. of Utago, NZ).

13- Participation à des jurys de thèses de master hors Etablissement

— *Evaluateur extérieur* de la thèse de Santiago Molina-Giraldo, "Video and Image Processing based on Kernel Representations", Universidad Nacional de Colombia, Manizales, Colombie, Mai 2014. Encadrants : G. Castellanos Dominguez, A. Alvarez Meza.

8 Liste des Publications

J'ai regroupé ici la liste de mes publications en les divisant en deux périodes : mes premiers travaux lors de mon arrivée au laboratoire L3i puis les travaux que j'ai effectué depuis mon arrivée au laboratoire MIA. Le tableau 6 montre le nombre de mes publications par type depuis mon arrivée à l'Université de La Rochelle. La liste de mes publications est fournie par catégories dans les sections suivantes, et est de plus disponible sur mon site web[7] ainsi que sur Academia[8], Research Gate[9], Science Stage[10] et Publication List[11]. De plus, toutes les mesures bibliométriques (Tableau 5) sont disponibles de façon détaillée sur mon profil Google Scholar .

Catégorie	L3i (2000-2006)	MIA (2007-2014)	Total
Editorial	-	2	2
Chapitre de livre	-	6	6
Journaux internationaux avec comité de lecture	-	7	7
Conférences internationales avec comité de sélection et actes	5	21	26
Total	5	36	41

TABLE 6 – Publications par type.

8.1 Editorial

— [E-1] T. Bouwmans, F. Porikli, B. Höferlin, A. Vacavant, "Preface", Handbook on "Background Modeling and Foreground Detection for Video Surveillance", CRC Press, Taylor and Francis Group, June 2014.

— [E-2] T. Bouwmans, J. Gonzàlez, C. Shan, M. Piccardi, L. Davis, "Special issue on Background Modeling for Foreground Detection in Real-World Dynamic Scenes", Special Issue on "Background Modeling for Foreground Detection in Real-World Dynamic Scenes", Machine Vision and Applications, July 2014.

8.2 Chapitres de Livre

— [CL-1] A Sobral, T. Bouwmans, "BGS Library : A Library Framework for Algorithm's Evaluation in Foreground/Background Segmentation", Handbook on "Background Modeling and Foreground Detection for Video Surveillance", Taylor and Francis Group, Chapter 23, June 2014.

— [CL-2] T. Bouwmans, "Recent Approaches in Background Modeling for Video Surveillance", Handbook on "Background Modeling and Foreground Detection for Video Surveillance", Taylor and Francis Group, Chapter 2, June 2014.

— [CL-3] T. Bouwmans, "Traditional Approaches in Background Modeling for Video Surveillance", Handbook on "Background Modeling and Foreground Detection for Video Surveillance", Taylor and Francis Group, Chapter 1, June 2014.

— [CL-4] C . Guyon, T. Bouwmans, E. Zahzah, "Robust Principal Component Analysis for Background Subtraction : Systematic Evaluation and Comparative Analysis", INTECH, Principal Component Analysis, Book 1, Chapter 12, page 223-238, March 2012.

— [CL-5] T. Bouwmans, "Background Subtraction For Visual Surveillance : A Fuzzy

7. http://sites.google.com/site/thierrybouwmans/
8. http://univ-larochelle.academia.edu/ThierryBOUWMANS
9. http://www.researchgate.net/profile/ThierryBOUWMANS
10. http://sciencestage.com/tbouwman
11. http://publicationslist.org/tbouwman

Approach", Handbook on "Soft Computing for Video Surveillance", Taylor and Francis Group, Chapter 5, March 2012.

— [CL-6] T. Bouwmans, F. El Baf, B. Vachon, "Statistical Background Modeling for Foreground Detection : A Survey", Handbook of Pattern Recognition and Computer Vision, World Scientific Publishing, Volume 4, Part 2, Chapter 3, pages 181-199, January 2010.

8.3 Journaux Internationaux avec Comité de Lecture

— [J-1] T. Bouwmans, "Traditional and Recent Approaches in Background Modeling for Foreground Detection : An Overview", Computer Science Review, Volume 11, pages 31-66, May 2014.

— [J-2] T. Bouwmans, E. Zahzah, "Robust PCA via Principal Component Pursuit : A Review for a Comparative Evaluation in Video Surveillance", Special Issue on Background Models Challenge, Computer Vision and Image Understanding, CVIU 2014, Volume 122, pages 22-34, May 2014.

— [J-3] D. Farcas, C. Marghes, T. Bouwmans, "Background Subtraction via Incremental Maximum Margin Criterion : A discriminative approach", Machine Vision and Applications, Volume 23, Issue 6, pages 1083-1101, October 2012.

— [J-4] T. Bouwmans, "Recent Advanced Statistical Background Modeling for Foreground Detection : A Systematic Survey", Recent Patents on Computer Science, Volume 4, No. 3, pages147-176, September 2011.

— [J-5] T. Bouwmans, F. El Baf, "Modeling of Dynamic Backgrounds by Type-2 Fuzzy Gaussians Mixture Models", MASAUM Journal of Basic and Applied Sciences, Volume 1, Issue 2, pages 265-277, September 2009.

— [J-6] T. Bouwmans, "Subspace Learning for Background Modeling : A Survey", Recent Patents on Computer Science, Volume 2, No. 3, pages 223-234, November 2009.

— [J-7] T. Bouwmans, F. El Baf, B. Vachon, "Background Modeling using Mixture of Gaussians for Foreground Detection : A Survey", Recent Patents on Computer Science, Volume 1, No. 3, pages 219-237, November 2008.

Journaux en rédaction

— [J-RE-1] A. Sobral, T. Bouwmans, E. Zahzah, "Foreground Detection via Decomposition in Low-Rank and Spase Matrices : A Review for a Comparative Evaluation with Large Scale Datasets", Computer Science Review, en rédaction, 2014.

— [J-RE-2] C. Guyon, T. Bouwmans, E. Zahzah, "Background Modeling via Nonlinear Polynomial RPCA with Spatial Constraint", Computer Vision and Image Understanding, CVIU 2014, en rédaction, 2014.

8.4 Conférences Internationales avec Comité de Sélection et Actes

— [C-1] A. Sobral, C. Baker, T. Bouwmans, E. Zahzah, "Incremental and Multi-feature Tensor Subspace Learning applied for Background Modeling and Subtraction"', International Conference on Image Analysis and Recognition, ICIAR 2014, October 2014.

— [C-2] C. Guyon, T. Bouwmans, E. Zahzah, "Foreground Detection via Robust Low Rank Matrix Decomposition including Spatio-Temporal Constraint", International Workshop on Background Model Challenges, ACCV 2012, pages 315-320, Daejeon, Korea, November 2012.

— [C-3] C. Guyon, T. Bouwmans, E. Zahzah, "Foreground Detection via Robust Low Rank Matrix Factorization including Spatial Constraint with Iterative Reweighted Regression", International Conference on Pattern Recognition, ICPR 2012, Tsukuba, Japan, November 2012.

— [C-4] C. Guyon, T. Bouwmans, E. Zahzah, "Moving Object Detection via Robust Low Rank Matrix Decomposition with IRLS scheme", International Symposium on Visual Computing, ISVC 2012, pages 665-674, Rethymnon, Crete, Greece, July 2012.

— [C-5] C. Guyon, T. Bouwmans, E. Zahzah, "Moving Object Detection by Robust PCA solved via a Linearized Symmetric Alternating Direction Method", International Symposium on Visual Computing, ISVC 2012, pages 427-436, Rethymnon, Crete, Greece, July 2012.

— [C-6] C. Guyon, T. Bouwmans, E. Zahzah, "Foreground Detection by Robust PCA solved via a Linearized Alternating Direction Method", International Conference on Image Analysis and Recognition, ICIAR 2012, pages 115-122, Aveiro, Portugal, June 2012.

— [C-7] C. Guyon, T. Bouwmans, E. Zahzah, "Foreground detection based on low-rank and block-sparse matrix decomposition", IEEE International Conference on Image Processing, ICIP 2012, Orlando, Florida, September 2012.

— [C-8] C. Marghes, T. Bouwmans, "Background Modeling via Incremental Maximum Margin Criterion", International Workshop on Subspace Methods, ACCV 2010 Workshop Subspace 2010, Queenstown, New Zealand, November 2010.

— [C-9] D. Farcas, T. Bouwmans, "Background Modeling via a Supervised Subspace Learning", International Conference on Image, Video Processing and Computer Vision, IVPCV 2010, pages 1-7, Orlando, USA , July 2010.

— [C-10] F. El Baf, T. Bouwmans, B. Vachon, "Fuzzy Statistical Modeling of Dynamic Backgrounds for Moving Object Detection in Infrared Videos", OTCBVS 2009, pages 60-65, Miami, Florida, June 2009.

— [C-11] F. El Baf, T. Bouwmans, B. Vachon, "Type-2 Fuzzy Mixture of Gaussians Model : Application to Background Modeling", International Symposium on Visual Computing, ISVC 2008, pages 772-781, Las Vegas, USA, December 2008.

— [C-12] F. El Baf, T. Bouwmans, B. Vachon, A Fuzzy Approach for Background Sub-

traction, IEEE International Conference on Image Processing, ICIP 2008, pages 2648-2651, San Diego, California, U.S.A, October 2008.

— [C-13]F. El Baf, T. Bouwmans, B. Vachon, "Fuzzy Foreground Detection for Infrared Videos", Joint IEEE International Workshop on Object Tracking and Classification in and Beyond the Visible Spectrum, OTCBVS 2008, pages 1-6, Anchorage, Alaska, USA, 27 June 2008.

— [C-14] F. El Baf, T. Bouwmans, B. Vachon, "Fuzzy Integral for Moving Object Detection", IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2008, pages 1729-1736, Hong-Kong, China, 1-6 June 2008.

— [C-15] F. El Baf, T. Bouwmans, B. Vachon, "Foreground Detection using the Choquet Integral", International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2008, pages 187-190, Klagenfurt, Austria, May 2008.

— [C-16] F. El Baf, T. Bouwmans, B. Vachon, "Comparison of Background Subtraction Methods for a Multimedia Learning Space", International Conference on Signal Processing and Multimedia, SIGMAP 2007, Barcelona, Spain, July 2007.

— [C-17] F. El Baf, T. Bouwmans, B. Vachon, "Comparison of Background Subtraction Methods for a Multimedia Application", International Conference on systems, Signals and Image Processing, IWSSIP 2007, pages 385-388, Maribor, Slovenia, June 2007.

— [C-18] D. Penciuc, F. El Baf, T. Bouwmans, "Comparison of Background Subtraction Methods for an Interactive Learning Space", NETTIES 2006, Timisoara, Roumania, September 2006.

— [C-19] S. Helan, J.C Burie, T. Bouwmans, S. Bazeille, "Object detection in underwater images", SEA TECH WEEK Caractérisation du Milieu Marin, CMM 2006, October2006.

— [C-20] R. Radlova, T. Bouwmans, B. Vachon, "Models used by vision-based motion capture", International Conference on Computer Graphics and Artificial Intelligence, 3IA 2006, pages 191-196, May 2006.

— [C-21] D. Semani, T. Bouwmans, C. Frélicot, P. Courtellemont, "Automatic fish recognition in interactive live videos", IVRCIA 2002, July 2002.

— [C-22] D. Semani, C. Saint-Jean, C. Frélicot, T. Bouwmans, P. Courtellemont, "Alive Fish Species Characterization for on line Video-based Recognition", SPR 2002, pages 689-698, August 2002.

8.5 Publications issues de Collaborations Locales

— [J-CL-1] I. Nastase, A. Meslem, T. Bouwmans, "Vortical structures analysis in jet flows using a classical 2D-PIV system and time resolved visualization image processing", Journal of Flow Visualization and Image Processing, JFVIP 2008, Volume 15, Issue 4, pages 275-300, 2008. (Rang A dans le domaine)

— [C-CL-1] I. Nastase, A. Meslem, T. Bouwmans, "Vortex dynamics analysis in jet flows using 2D planar PIV and high-speed laser tomography image processing", Internatio-

nal Symposium on Flow Visualization, Nice, France, 2008.

8.6 Publications issues de Collaborations Internationales

— [C-23] S. Javed, A. Sobral, S. Oh, T. Bouwmans, S. Jung, "OR-PCA with MRF for Robust Foreground Detection in Highly Dynamic Backgrounds", Asian Conference on Computer Vision, ACCV 2014, Singapore, November 2014.

— [C-24] Z. Zhao, T. Bouwmans, X. Zhang, Y. Fang, "A Fuzzy Background Modeling Approach for Motion Detection in Dynamic Backgrounds", International Conference on Multimedia and Signal Processing, Shanghai, China, December 2012.

— [C-25] C. Marghes, T. Bouwmans, R. Vasiu, "Background Modeling and Foreground Detection via a Reconstructive and Discriminative Subspace Learning Approach", International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV 2012, Las Vegas, USA, July 2012.

— [C-26] R. Zaman, V. Khorramshahi, T. Bouwmans, "Foreground Object Detection using Choquet Fuzzy integral in Video Images", Iranian Conference on Machine Vision and Image Processing, MVIP 2010, Isfahan, Iran, October 2010.

8.7 Publication de Vulgarisation

— [V-1] F. El Baf, T. Bouwmans, "Robust Fuzzy Statistical Modeling of Dynamic Backgrounds in IR Videos", SPIE Newsroom, Defense and Security, 13 November 2009. doi:10.1117/2.1200911.1842

8.8 Publication de Diffusion sur Internet

— [I-1] T. Bouwmans, http://sites.google.com/site/backgroundsubtraction/Home, site web sur le domaine de la modélisation de fond. 28 029 visites et 14 855 visiteurs depuis sa création 01/03/2010 (Source Google Analytics - 09/06/2014).

— [I-2] T. Bouwmans, http://sites.google.com/site/fuzzybackgroundmaintenance/, site web sur la publication ICIP 2008.

— [I-3] T. Bouwmans, http://sites.google.com/site/t2fmog/, site web sur la publication ISVC 2008.

— [I-4] T. Bouwmans, http://sites.google.com/site/fuzzyforegrounddetection/, site web sur la publication FUZZ'IEEE 2008.

— [I-5] T. Bouwmans, http://sites.google.com/site/immcforegrounddetection/, site web sur les publications IVPCV 2010, ACCV 2010 et MVA 2012.

— [I-6] T. Bouwmans, http://sites.google.com/site/rpcaforegrounddetection/, site web sur les publications ICIP 2012, ICIAR 2012, ISVC 2012 et ICPR 2012.

9 Témoignages

Plusieurs de mes travaux ont reçu des appréciations enthousiastes de la part de membres de la communauté du "background modeling" et j'ai regroupé certaines d'entre-elles dans la partie "comments" des articles sur mon site web pour ce qui concernent les publications et sur l'onglet "comments" du site "Background Subtraction Web Site" pour ce site.

Deuxième partie

Partie 2 : Recherche Scientifique

Chapitre 1

Introduction Générale

Après une thèse soutenue en 1997 à l'Université du Littoral Côte d'Opale, j'ai obtenu un poste d'ATER à l'Université de La Rochelle en Février 2000 et ensuite j'ai été nommé Maître de Conférences en Février 2001. J'y ai intégré le Laboratoire d'Informatique et d'Imagerie Industrielle (L3I - EA. 2118), aujourd'hui Laboratoire Informatique, Image, Interaction. A partir du 1er Janvier 2008, je suis devenu membre du Laboratoire de Mathématiques, Image et Applications MIA. Ce changement de laboratoire correspondait à un regroupement d'une partie des chercheurs du L3i (7 chercheurs et 2 doctorants) avec le laboratoire de mathématiques avec un thème commun : l'image. Cette mise en commun de ces compétences s'est faite dans le cadre du Pôle Régional de Recherche en Images, Données et Systèmes (PRIDES) qui est une Fédération des Laboratoires du Poitou-Charentes dans le domaine des Sciences et Technologies de l'Information et de la Communication. Je décris ci-dessous dans quel axe et quelle équipe de recherche mes travaux s'inscrivaient ou s'inscrivent pour chaque laboratoire :

— **Recherches au laboratoire L3i :** Comme mes recherches se sont effectuées au sein du L3i à mon affection, je vais donc décrire la structure du laboratoire L3i tel qu'il était lorsque j'en étais membre. Une partie des activités du L3I concernait les aspects analyse/modélisation/synthèse, c'était l'objet de la thématique Image et Séquences d'Images (ISI). Pour l'analyse d'Image et pour l'interprétation de comportements complexes, le travail de fusion/classification nécessaire se faisait dans la thématique Données, Formes, Interprétation (DoFIn). Finalement, la thématique Modèles, Comportements, Architectures (MoCA) contribuait à inscrire les travaux menés dans un contexte opérationnel et un cadre méthodologique (logiciel et matériel). Mes travaux se sont naturellement inscrits dans la première thématique ISI durant ma période au laboratoire L3i.

— **Recherches au laboratoire MIA :** Le MIA est structuré en deux équipes : "Mathématiques et Image" et "Mathématiques et Problèmes d'Evolution". L'équipe Mathématiques et Image, est structurée suivant les projets suivants : 1) Détection, analyse et

synthèse de textures dynamiques, 2) Traitement et analyse d'images multi-spectrales, 3) Classification et Reconnaissance de forme et 4) Géométrie de l'information. Mon activité de recherche s'effectue donc l'équipe " Mathématiques et Image" dans les projets 1), 2) et 3).

Comme indiqué dans mon curriculum vitae, mon parcours scientifique s'est effectué dans le domaine du Traitement et de l'Analyse d'Images. Durant ma thèse, j'ai abordé la contribution des concepts flous dans un système de multivision. A mon arrivée au L3i, j'ai participé au projet Aqu@thèque sur la détection des poissons dans des séquences vidéo. J'ai donc naturellement centré mon activité de recherche sur la détection d'objets mobiles dans les vidéos. La détection d'objets mobiles peut se faire soit par différence consécutive d'images, soit par des méthodes de soustraction de fond ou soit par des méthodes basées sur le flot optique. La méthode par différence consécutive d'images est simple mais peu robustes dans des environnements non-contrôlés. Les méthodes basées sur le flot optique sont relativement précises mais coûteuse en temps de calcul. Par contre, les méthodes par soustraction de fond offrent un bon compromis entre robustesse et performance.

Dans ce contexte, j'ai donc commencé par une première étude sur le domaine des méthodes de suppression de fond. Cette étude a mis en évidence un grand nombre de publications dans le domaine proposant des méthodes spécifiques pour des applications spécifiques. De plus, les méthodes les plus performantes de l'époque comme le mélange de Gaussiennes [490] et l'estimation de noyau [167] ne donnaient pas de résultats satisfaisants sur les séquences issues du projet Aqu@thèque [398][39][40]. En effet, celles-ci présentent des situations de fonds dynamiques plus variés et des variations lumineuses plus fortes que celles rencontrées dans d'autres applications comme la vidéo-surveillance. De plus, la forme des objets mobiles (ici en l'occurrence les poissons dans le projet Aqu@thèque), ont des formes et des couleurs bien plus variées que les objets mobiles issus d'activités humaines. Devant cette situation, j'ai engagé un état de l'art systématique du domaine tant sur certaines catégories de modèles que sous-catégories plus ou moins étudiées à l'époque pour en dégager des pistes nouvelles. J'ai ensuite engagé des recherches pour proposer l'utilisation d'autres outils mathématiques comme les concepts flous ainsi que les méthodes de réduction de dimension discriminatives et mixtes. Finalement, j'ai étudié particulièrement les recherches récentes sur les ACP robustes. Mon activité de recherche se décline donc ainsi en quatre chapitres qui correspond à chacune de ces recherches spécifiques et j'explique ci-dessous de façon concise le chemin que j'ai suivi pour chacune d'entre-elles. L'aspect motivation et rayonnement correspondant est également expliqué dans la Section 3 de la Partie 1 concernant l'aspect "Activités de Recherche".

Devant le grand nombre de publications dans ce domaine, ma recherche a donc d'abord consisté à effectuer un état de l'art complet dans le domaine dans un premier temps pour me former et aussi pour structurer le domaine en classant les différentes méthodes existantes suivant le modèle mathématique utilisé à chacune des étapes des méthodes de suppression de fond. J'ai ainsi effectué une analyse et une synthèse la plus exhaustive possible de l'activité du domaine reconnue par la communauté scientifique. De plus, cet état de l'art (Chapitre 2) a permis de dégager de nouvelles pistes pour développer des méthodes plus robustes aux challenges clés du domaine que sont les variations lumineuses et les fonds dynamiques.

Ces nouvelles pistes se déclinent en trois catégories qui correspondent aux encadrements de travaux de deux thèses et de deux stages de type "Recherche" de Master 2 effectués sous ma direction scientifique. Ces trois pistes développées dans les trois derniers chapitres de cette Partie 2 concernant l'aspect "Recherche Scientifique" sont listées ci-dessous. Dans un premier temps, j'ai donc étudié et évalué l'apport des concepts flous dans les différentes étapes de la détection d'objets mobiles dans le cadre de la thèse de Fida El Baf. L'objectif était de prendre en compte l'imprécision et l'incertitude au niveau de la modélisation du fond, de la classification fond/objet et de la mise à jour du fond en présence de fonds dynamiques et de variations lumineuses. Cette partie est développée dans le Chapitre 3.

Dans un deuxième temps, je me suis concentré sur les méthodes de réduction de dimension pour en évaluer leur adéquation à la détection d'objets mobiles. Ces travaux ont concerné une approche discriminative et une approche mixte dans le cadre des stages de Master ERASMUS de Diana Farcas et Cristina Marghes de l'Université de Timisoara en Roumanie. Cette partie est développée dans le Chapitre 4.

Dans un troisième temps, les recherches récentes sur les ACP robustes basées sur une décomposition en une matrice de rang faible et une matrice parcimonieuse ont attirées mon attention car quelques publications en montraient la possible application pour la séparation fond/objet mobiles. Ces travaux effectués dans le cadre de la thèse de Charles Guyon ont donc porté sur la mise en oeuvre de méthodes ACP robustes adaptées pour la détection d'objets mobiles dans le cas de variations lumineuses et de fonds dynamiques. Cette partie est développée dans le Chapitre 5. Finalement, nous présentons une conclusion sur ces différents travaux ainsi que les perspectives possibles de recherche concernant l'initialisation du fond, la modélisation de fond et la sélection de caractéristiques.

Chapitre 2

De l'importance de la détection d'objets mobiles dans les vidéos : Un domaine complet de recherche scientifique.

Dans ce chapitre, nous présentons tout d'abord l'importance des méthodes de suppression de fond pour la détection d'objets mobiles dans des vidéos expliquant ainsi notre motivation à en faire notre thématique de recherche. L'objectif de notre travail a donc consisté dans un premier temps en un état de l'art approfondi des différentes méthodes de suppression de fond, ce qui nous a permis d'aboutir à la détermination d'un cadre général pour ces méthodes, et également de dégager les hypothèses émises par ces méthodes ainsi que les situations critiques. Nous présentons aussi brièvement les méthodes d'évaluation de ces méthodes ainsi que les bases de données de séquences vidéo disponibles pour les tester. Ainsi, le lecteur trouvera dans ce chapitre une vue structurée et compléte du domaine.

Ce travail d'état de l'art a donné lieu à des publications de type "Survey" qui portent sur les méthodes de suppression de fond par les modèles statistiques (1 chapitre de livre HPRCV [76], 1 Journal RPCS [70]), par le mélange de Gaussiennes (1 Journal RPCS [75]), par les modèles de réduction de dimension (1 Journal RPCS [69]), par les modèles flous (1 chapitre de livre CRC Press [71]), et par les ACP robustes (1 chapitre de livre INTECH [227], 1 Journal CVIU [79]). Enfin, deux chapitres de livre dressent un état de l'art complet sur les méthodes de suppression de fond actuelles [73] et prospectives [72].

De plus, toutes les informations relatives à ces méthodes de suppression de fond sont disponibles sur le site que nous avons développé et que nous avons appelé "Background Subtraction Web Site[1]". Ce site web contient une liste exhaustive des publications dans le domaine, des ressources disponibles comme les bases de données et les implémentations. Dans chaque cas, la liste est régulièrement mise à jour et les méthodes sont classées conformément à la classification fournie dans ce chapitre et dans les publications relatives. L'idée de ce site est de diffuser ce cadre commun dans la communauté scientifique, et de faciliter ainsi

les recherches tant au niveau des chercheurs débutants dans le domaine que pour les experts amenés à examiner l'originalité et la pertinence d'une publication.

1 Introduction

Avec les nouvelles technologies et la puissance de calcul des ordinateurs, la détection d'objets mobiles en temps réel dans des vidéos acquises par des caméras fixes connait un grand essor depuis une quinzaine d'années. Ainsi, cette détection devient naturellement une alternative essentielle à des opérateurs humains ou à des capteurs mécaniques dans un grand nombre d'applications. Cette étape qui est bien souvent l'étape préliminaire à d'autres traitements nécessite la séparation de l'information qui change (objets mobiles) de l'information statique (fond). Pour effectuer cette étape, le processus le plus utilisé est la suppression de fond car elle offre un bon compromis performance et temps de calcul. Cette étape est donc nécessaire dans une multitude de domaines d'application comme la vidéo surveillance, la capture de mouvement et les applications nécessitant des interactions hommes-machine via caméra statique. Toutes ces applications montrent ainsi l'importance de cette étape qui est de plus l'étape préliminaire aux autres étapes comme le suivi d'objet, la reconnaissance d'objets ou l'analyse de comportement.

1.1 De l'utilisation des méthodes de suppression de fond

Les méthodes de suppression de fond présentent donc un moyen pratique pour la détection d'objets mobiles dans des vidéos dans le cas d'application où la caméra est fixe ou peu mobile.

1- Cas des applications avec caméras fixes

Les méthodes de suppression de fond sont essentiellement employées dans des applications où les caméras sont fixes. Les principaux domaines d'utilisation des méthodes de suppression de fond sont donc les suivants :

— **Vidéo surveillance intelligente d'activités humaines :** C'est l'application principale de la suppression de fond pour la détection d'objets mobiles ou abandonnés pour assurer la sécurité des zones surveillées, ou pour calculer des statistiques sur le trafic comme sur les routes [527], les aéroports [65] ou la surveillance maritime [67]. Les objets d'intérêt sont de nature très variée comme les véhicules, les avions, les navires, les personnes et les bagages. La surveillance peut être encore plus spécifique comme la surveillance du comportement d'un client dans un magasin [13][29][306]. La vidéo surveillance intelligente d'activités humaines est ainsi la plus demandeuse en terme

1. http://sites.google.com/site/backgroundsubtraction/Home

d'applications des méthodes de suppression de fond. Les industriels du domaine s'y intéressent fortement et s'y impliquent au niveau recherche comme Thales, Foxtream[2] (France), Teb-Online[3] (France), IntuVision[4] (USA) et Geeksys[5] (Brésil).

— **Vidéo surveillance intelligente d'activités animales :** La vidéo surveillance permet aussi l'observation des activités animales dans des zones protégées (rivière, océan, etc..) ou dans des zoos pour l'éthologie. Les objets d'intérêt sont aussi très divers comme des oiseaux [289][290][458], des abeilles [88][525][287], des poissons [487] et des animaux sauvages de la forêt [360][277].

— **Vidéo surveillance intelligente d'environnements naturels :** Dans le cadre de la protection et de la surveillance de l'environnement, il peut être utile de détecter certains objets mobiles comme par exemple la détection d'objets flottants (morceaux de bois) dans une rivière [14][15].

— **Capture de mouvement sans marqueur :** L'objectif est d'obtenir une capture de mouvement complète et précise d'une personne ou de plusieurs personnes à partir de vues obtenues par des caméras fixes [213]. La silhouette est généralement extraite par suppression de fond dans chaque vue. Ensuite, l'enveloppe convexe est obtenue en trois dimensions. Plus il y a de caméras plus l'enveloppe convexe est précise.

— **Interaction Homme-Machine :** Plusieurs applications requièrent une interaction Homme-Machine via une caméra fixe en temps réel comme les jeux-vidéos (Microsoft's Kinect) et les applications ludo-pédagogiques comme Aqu@theque [37].

— **Codage vidéo basé sur le contenu :** L'idée est de séparer les objets mobiles du fond pour ne transmettre qu'une fois l'image du fond et ensuite que les objets mobiles pour diminuer le flux vidéo. Le fond enregistré et les objets mobiles sont ainsi codés séparément. Ainsi, le codage vidéo basé sur le contenu a besoin de méthodes efficaces de détection d'objets mobiles dans des environnements présentant tant des fonds statiques que des fonds dynamiques [118][396].

— **Video matting :** La technique de *video matting* est très utilisée en cinéma ou télévision. Pratiquement, l'idée est de replacer des personnes filmées en premier plan dans un environnement différent. L'objectif est donc d'extraire les objets de premier plan avec leur coefficient *alpha matte* pour les mettre sur un fond différent. Les premières approches employaient souvent trois cartes (*trimaps*) labelisées manuellement par un

2. http://www.foxstream.fr/newsite/fr/
3. http://www.teb-online.com/fr/index.php
4. http://www.intuvisiontech.com/
5. http://www.geeksys.com.br/en/heatmaper

Applications	Sous-catégories	Objets d'intérêt
Vidéo surveillance intelligente d'activités humaines	Surveillance routière	Voitures, personnes [527]
	Surveillance d'aéroport	Avions, personnes [65]
	Surveillance maritime	Navires, personnes [67]
	Surveillance de magasins	personnes [13][29][306]
Vidéo surveillance intelligente d'activités animales	Surveillance d'oiseaux	Oiseaux [289][290][458]
	Surveillance d'insectes	Abeilles [88][525][287]
	Surveillance de poissons	Poissons [487]
Vidéo surveillance intelligente d'environnements naturels	Rivières, fleuves, etc...	Morceaux de bois [14][15]
Capture de mouvement sans marqueur	Enveloppe convexe	Personnes [213]
Interaction Homme Machine	Jeux vidéo	Personnes
	Ludo-applications	Personnes, animaux [37]
Codage vidéo basé sur le contenu	Contenu	Objets [118][396]
Video Matting	Video Matting	Personnes [189]

TABLE 2.1 – Vue d'ensemble des applications de la détection d'objets mobiles.

expert : objets de premier plan, arrière plan (fond) et région inconnue. Les techniques récentes emploient des méthodes de suppression de fond pour aider à ce processus [189].

2- Cas des applications avec des caméras mobiles

Les méthodes de suppression de fond peuvent être adaptées et utilisées dans le cadre d'applications où la caméra est mobile. Cela peut être le cas d'applications utilisant des caméras PTZ [184][215][34], des caméras des systèmes mobiles (smartphones) [460] ou des caméras embarquées (voitures, robots) [460]. Dans ce cas, différentes approches permettent alors d'utiliser les méthodes de suppression de fond développées dans le cas des applications avec des caméras :

1. Approches par compensation de mouvement de la caméra [460][171]
2. Approches par reconstruction d'une image de fond panoramique [526][494]
3. Approches géométriques [610][612][611]

De plus, certaines méthodes reconstruisent le fond en deux dimensions [460] et d'autres en trois dimensions [507]. Une autre solution [140] consiste à utiliser deux caméras dont 1) l'une est fixe, permettant d'avoir une image de fond de la scéne compléte en haute définition, et 2) l'autre est mobile (PTZ), permettant de se focaliser sur une partie de la scène. La détection des objets mobiles se fait dans la zone filmée par la caméra mobile et en utilisant l'image de fond HD fourni par la caméra HD sur cette zone [140]. Ces objets mobiles peuvent ainsi être suivi par la caméra PTZ en ayant toujours une image de fond disponible.

1.2 De la multidisciplinarité scientifique des méthodes de suppression de fond

Le domaine de la détection des objets mobiles par des méthodes de suppression de fond est un domaine qui demande des connaissances de plusieurs disciplines qui sont les suivantes :

— **Mathématique :** Les modèles mathématiques utilisés dans les méthodes de suppression de fond sont très variées comme les suivants :

1. des modèles simples (moyenne, médiane)
2. des modèles statistiques avec des distributions de formes différentes comme la distribution gaussienne, de Dirichlet [231][173] ou de Poisson [178].
3. des modèles de réduction de dimension de bases (PCA [387], ICA [579], NMF [83]) ou robuste par décomposition en matrices de rang faible et parcimonieuse (RPCA [92], RNMF [212], etc...)
4. des modèles flous basés sur les différents concepts des ensembles flous [46].

D'autres modèles peuvent être employés comme la thèorie de vraisemblance de Dempster-Shafer [372].

— **Image :** Les entités à comparer qui vont de la taille d'un pixel à la taille d'un cluster en passant par la taille d'un bloc ont besoin de caractéristiques basés sur des représentations d'espace couleur (RGB, YCrC, HSV, etc..) [293][431][270], de contours [261], de texture [245], de mouvement [642] et de stéréoscopie [202].

— **Informatique :** Les algorithmes sont testés et implémentés dans différents langages (Matlab, C, C++,...) en employant aussi des librairies "image" comme OpenCV. L'implémentation de ces algorithmes mettent en jeu aussi des régles de génie logiciel pour avoir des codes ré-utilisables. De plus pour diminuer le temps de calcul, certaines implémentations peuvent être réalisées en employant la carte graghique (GPU) et le langage employé dans ce cas est CUDA [179].

— **Electronique :** Pour atteindre des performances temps réel, il est parfois utile de concevoir des architectures dédiées qui peut être basée soit sur la technologie DSP [27], soit sur la technologie VLSI [524], soit sur la technologie FPGA [247].

Mon champ de recherche (et par conséquent celui de ce manuscrit) est quasiment totalement consacré aux aspects mathématiques, images et informatiques.

1.3 Evolution et analyse du domaine

Le domaine de recherche des méthodes de suppression de fond a connu des publications majeures dans les années 1999-2000 avec la publication des premières méthodes sophistiquées et basées sur le mélange des Gaussiennes (MOG) développée par Stauffer et Grimson [490], l'estimation de densité par la méthode du noyau (KDE) développée par Elgammal et al. [167] et l'analyse en composante principale (ACP) conçue par Oliver et al. [387] alors qu'avant seuls des modèles basiques basés sur la moyenne, la médiane ou l'histogramme étaient employés.

Une fois, ces méthodes parues et fonctionnant relativement bien dans les applications de l'époque, c'est-à-dire la vidéo surveillance et la capture de mouvement dans des environnements relativement simples, une myriade de variations de ces méthodes a été développée au fur et à mesure pour adapter et améliorer ces méthodes dans des environnements de plus en plus complexes. En effet, à l'heure actuelle, nous avons dénombré plus de 250 publications sur les MOG [75][70], plus de 70 publications sur les KDE [70] et plus de 30 sur l'ACP [69]. Chacune de ces publications améliorent la méthode originale de façon progressive dans des applications où les challenges sont de plus en plus complexes. Cette évolution très progressive marquée par un grand nombre de publications a donné à tort le sentiment que la question était soit quasiment résolue ou sur-étudiée alors qu'aucune méthode n'arrivait et n'arrive pas encore à l'heure actuelle à être vraiment robuste à un large spectre de challenges rencontrés dans les environnements extérieurs complexes lors de longues séquences comme dans les cas réels.

De plus, l'émergence de récentes applications autres que la vidéo surveillance et pour lesquelles les vidéos contiennent aussi des fonds dynamiques complexes, nécessite de nouveaux développements pour détecter de façon robuste des objets de formes et de couleurs beaucoup plus diverses que ceux rencontrés dans le cas de la vidéo surveillance classique. Ainsi, des méthodes plus efficaces et robustes que les MOG [490] et les KDE [167] sont nécessaires pour gérer à la fois la présence de fonds dynamiques et de variations lumineuses avec des caméras fixes ou mobiles. De plus, ces méthodes avancées doivent pouvoir fonctionner en temps réel et avec peu de mémoire notamment pour les systèmes embarqués et mobiles.

C'est dans ce contexte que se sont développées des méthodes prospectives basées sur des modéles statistiques avancées [51][248], des modéles flous [71], des modèles par ACP robustes via la décomposition en matrices de rang faible et parcimonieuse [79], des modèles parcimonieux [255][254][628][575] et des modèles à bases de tenseurs [501][502][522][312][318]. Ces méthodes sont encore à leur début et nécessitent pour la plupart des recherches pour être incrémentales et temps réel et être ainsi directement applicables.

Dans ce contexte, nous croyons que nous vivons une transition clé dans le domaine de la détection d'objets mobiles par les méthodes de suppression de fond tant dans la structuration du domaine que dans le développement de nouvelles méthodes ainsi que dans la validation plus compléte de ces méthodes. Notre travail s'inscrit dans cette logique et participe donc à cette dynamique de transition. De plus, l'activité actuelle montre que les autres chercheurs actifs du domaine agréent à cette analyse et la communauté scientifique se retrouve dans le cas de la vidéo surveillance classique autour de numéro spéciaux dans des journaux (Special Issue on "Background Modeling and Foreground Detection for Video Surveillance" dans le journal MVA, Special Issue on "Background Models Comparison" dans le journal CVIU), des conférences (AVSS par exemple) et des workshops (ChangeDetection.net[1] [204] en conjonction avec CVPR, Background Models Challenge[2] [530] en conjonction avec ACCV). Pour le cas de la vidéo surveillance intelligente d'activités animales, il existe deux workshops qui sont organisées en conjonction avec ICPR : Visual observation and analysis of Vertebrate And Insect Behavior (VAIB[3]) et Computer Vision for Analysis of Underwater Imagery (CVAUI[4]).

1.4 Contributions

Depuis le début des années 2000, la multitude d'applications a donc fait émerger un grand nombre de méthodes plus ou moins simples suivant le contexte environnemental (fonds simples ou fonds dynamiques) et suivant le niveau d'expertise en image du domaine d'application. Les méthodes développées dans le cadre de la vidéo-surveillance sont les plus robustes et complètes. Cependant, aucun algorithme traditionnel n'émergeait et ne permettait de pallier simultanément les principaux challenges rencontrés dans les vidéos. Cela était essentiellement dû aux raisons suivantes :

1. **Manque d'un cadre commun** car chaque méthode étant développée dans des contextes différents, les différentes étapes et points clés n'étaient pas systématiquement explicités et abordés de façon optimale avec le même formalisme.
2. **Manque de progrès scientifique** sur les modèles utilisés : En effet, il y a un très grand nombre d'amélioration des mélanges de Gaussiennes [490], et les autres modèles étaient souvent peu ou pas examinés.
3. **Manque d'une base de données commune** pour évaluer les algorithmes de façon systématique et qui soit suffisamment large pour couvrir tous les challenges possibles et possédant la vérité-terrain pour toutes les images de chaque séquence.

1. http://www.changedetection.net/
2. http://bmc.univ-bpclermont.fr
3. http://homepages.inf.ed.ac.uk/rbf/vaib14.html
4. http://cvaui.oceannetworks.ca/

Pratiquement, le domaine possédait un grand nombre de publications sur certains modèles (comme par exemple sur les modèles SG [562], MOG [490] et KDE [167]) et un grand nombre d'évaluations comparatives [59] [58][22] [384][152][186][185][232][392][52][246] [376][244]. Pour autant dans toutes ces évaluations comparatives, aucun algorithme particulier ne sortait du lot en termes de robustesse sur tous les challenges à cause des trois raisons citées précédemment. Remarquons aussi que l'objectif des évaluations comparatives est différent de celui des états de l'art car ces évaluations se focalisent sur la description de méthodes existantes et leur évaluation sur une même base de données, ne permettant pas ainsi de structurer le domaine et de dégager des pistes nouvelles.

Mes travaux s'inscrivent donc dans ce contexte historique et scientifique du domaine. Pour le point (1), je me suis donc attaché à dégager un cadre commun pour les méthodes de suppression de fond, à effectuer des états de l'art systématiques et complet pour donner une vue plus globale et précise du domaine (développée dans ce chapitre) dans le but de dégager des directions pour innover. Pour le point (2), je me suis concentré sur l'examination et l'évaluation de l'apport d'outils mathématiques différents pour proposer de nouvelles pistes développées. Celles-ci sont respectivement développées dans les chapitres 3, 4 et 5. Pour le point (3) sur les bases de données, la création récente des bases de données complètes que sont ChangeDetection.net [204] et BMC 2012 [530], chacune étant adossée à des workshops, offre maintenant la possibilité de comparer de façon efficace les différents algorithmes comme le montrent les évaluations comparatives récentes sur ces bases de données [204][530][486].

2 Cadre commun général

Devant la très grande diversité des méthodes de détection d'objets mobiles, il est difficile d'identifier la spécificité de chacune d'entre elles et donc de choisir celle qui convient le mieux à une application donnée. Pour cela, nous avons tout d'abord dégagé un cadre commun pour ces méthodes qui est basé sur les questions suivantes :

— Quelles sont les hypothèses émises ?
— Quelles sont les situations critiques qui peuvent apparaître ?
— Quelles sont les différentes méthodes et comment les classer ?
— Quelles sont les différentes étapes communes à ces différentes méthodes ?
— Comment évaluer leurs performances respectives ?
— Existe-t-il des séquences de référence test ?

Notre travail a donc consisté dans un premier temps à répondre à ces différentes questions dont les réponses sont présentées dans les sections suivantes.

2.1 Terminologie

Pour bien comprendre la problèmatique des méthodes de suppression de fond et de la détection des objets mobiles, il est indispensable d'introduire le paragdime implicite sur lesquel elles reposent ainsi que quelques définitions de termes couramment employés dans le domaine.

1- Notions de fond et d'objets d'intérêts

Les méthodes de suppression de fond sont basées sur le paragdime que la différence entre l'image du fond et l'image courante est due à la présence d'objets mobiles. Les pixels qui "ne changent pas" sont considérés comme appartenant au *"fond" (Background)* ou *"arrière-plan"*, et les pixels qui "changent" sont considérés comme appartenant à l'*"avant-plan" (Foreground*). Par extension dans la litterature, la notion d'*"avant plan"* est assimilé aux *"objets d'intérêts"*. En pratique, cela n'est pas forcément vérifié car les *"objets d'intérêts"* peuvent autant se trouver dans l'*"avant-plan"* que dans l'*"arrière-plan"* si ces termes sont pris dans leur signification cinématographique. Il en va de même pour les éléments qui constituent le *"fond"*.

Les *"objets d'intérêts"* peuvent être classés en deux sous-catégories que sont les *"objets stationnaires"* et les *"objets mobiles"*. Pour la détection, il existe trois types d'applications :

1. Certaines applications ne s'intéressent qu'aux *"objets stationnaires"*. On peut citer entre autres les applications qui concernent la détection d'objets abandonnés [409].

2. Certaines applications se focalisent que sur les *"objets mobiles"*. C'est le cas de la plupart des applications où seuls les objets mobiles sont intéressants comme pour la détection d'intrusion ou le comptage de véhicules sur une route.

3. D'autres applications nécessitent de détecter à la fois les *"objets stationnaires"* et les *"objets mobiles"*. On peut citer par exemple le cas de l'application Aqu@thèque [37] où il est nécessaire de pouvoir détecter à la fois les poissons qui sont en mouvement et ceux qui restent immobiles.

Pour le second cas, la notion d'*"avant-plan"* est assimilée aux *"objets mobiles"* et il en sera ainsi tout au long de ce manuscrit. La figure 2.1 montre la relation entre ces différents termes.

D'un point de vue pratique, il est implicitement admis que le "fond" est l'image de la scène filmée qu'on retrouve après élimination dans cette scène de 1) tous les objets mobiles se mouvant par eux-mêmes (comme des personnes ou des voitures) et/ou 2) les objets mobiles qui sont mus par une action extérieure (comme des morceaux de bois flottant sur une

rivière [14][15]), avec 1) et 2) qui ne font que passer dans cette scène dans un temps relativement court avec la notion de "court" relative à l'application. Les fonds peuvent se présenter sous deux formes que sont les *"fonds statiques"* et les *"fonds dynamiques"*. Les *"fonds dynamiques"* sont ceux où 1) des objets faisant partie du fond peuvent être déplacés ou insérés et/ou 2) des objets faisant partie du fond peuvent être en mouvement sous une action extérieure comme le vent. Par exemple, les mouvements des arbres ou les mouvements de l'eau à sa surface sont considérés comme "fond" et ne doivent donc pas être détectés car ceux-ci restent effectivement présents dans la scène. Par ailleurs, un objet faisant partie du fond qui est déplacé ou inséré ne doit donc pas non plus être détecté.

Les *objets mobiles* quant à eux présentent différentes tailles, formes et couleurs suivant les applications et aucune information *a priori* est émise quant à ces trois caractéristiques. Effet, les "objets mobiles" sont de nature très variés comme des véhicules (voitures, avions, navires), des personnes ou des animaux (oiseaux, abeilles, poissons et animaux sauvages de la forêt). De plus, tous ces "objets mobiles" peuvent se trouver plus ou moins proches de la caméra.

Un autre point clé est que, suivant l'application, certains *"objets mobiles"* peuvent passer temporairement dans la catégorie *"fond"*. Par exemple, une voiture stationnée ne doit pas être détectée et ne doit l'être que lors de son démarrage. Inversement, une voiture qui s'arrête doit être intégrée dans le fond au cours d'un certain temps qui dépend de l'application.

En pratique, des pixels qui "ne changent pas" par comparaison avec l'image courante peuvent parfois appartenir à un objet mobile si celui-ci à la même couleur que le fond par exemple. De même, des pixels qui "changent" peuvent appartenir au fond dans le cas d'un arbre qui bouge à cause du vent par exemple.

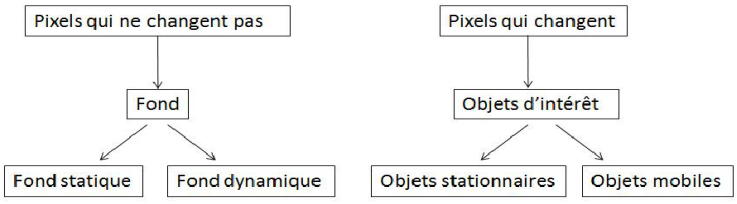


FIGURE 2.1 – Terminologie

Maintenant que les notions de "fond" et d'"objets mobiles" ont été explicitées, nous allons voir dans la section suivante les définitions associées aux différents étapes des méthodes de suppression de fond.

2- Définitions des différentes étapes

Dans ce manuscrit, nous définissons les différentes étapes des méthodes de suppression de fond comme suit :

1. **L'initialisation du fond** (*Background Initialization*) consiste à déterminer la première image de fond.
2. **La modélisation de fond** (*Background Modeling*, *Background Representation*) décrit le modèle utilisé pour représenter le fond.
3. **La maintenance du fond** (*Background Maintenance*) concerne le mécanisme de mise à jour du modèle de fond pour s'adapter aux changements qui apparaissent au cours du temps.
4. **La classification des pixels en fond/objets mobiles** (*Foreground Detection*) consiste comme son nom l'indique de classer les pixels dans la classe "fond" ou "objets mobiles".

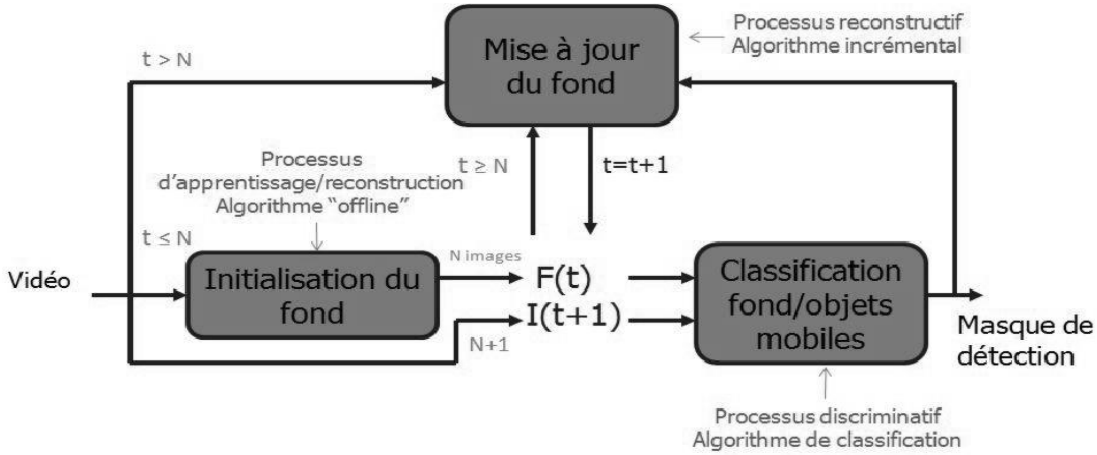


FIGURE 2.2 – Pipeline des méthodes de suppression de fond. N est le nombre d'images nécessaires à l'initialisation, $F(t)$ est l'image de fond à l'instant t et $I(t+1)$ est l'image courante à l'instant $t+1$

Ces différentes étapes (examinées de la Section 3.1 à la Section 3.4) mettent en jeu des méthodes qui ont des objectifs et des contraintes différentes. Par conséquent, elles nécessitent des algorithmes à caractère différent. L'initialisation du fond nécessite un algorithme d'apprentissage qui peut être effectué "hors-ligne" (*"off-line"*) en prenant les données dans leur ensemble (*"batch"*). Par contre, l'algorithme de mise à jour doit être effectué en ligne (*"on-line"*) en prenant les données au cours du temps (*incremental*). Les parties initialisation, modélisation et maintenance ont besoin d'un algorithme de type reconstructif alors que la partie classification nécessite quant à elle un algorithme de type discriminatif. La figure 2.2

montre le processus des méthodes de suppression de fond avec ses différentes étapes.

Il est important de voir également que deux images sont comparées et que pour cela, les méthodes comparent une sous-entité de l'entité "image" de l'image de fond avec la correspondante dans l'image courante. Cette sous-entité peut être de la taille d'un pixel, d'une région ou d'un cluster (voir Section 3.5). De plus, cette sous-entité est caractérisée par une caractéristique (*feature*) qui peut être de type couleur, contour, texture, stéréoscopique ou mouvement (voir Section 3.6).

3- Définition autour de la thèmatique

Dans la littérature, on trouve les expressions de "méthodes de suppression de fond" (*"Background Suppression"*) et de "méthodes de soustraction de fond" (*"Background Subtraction"*). Les méthodes de soustraction de fond sont un sous-ensemble des méthodes de suppression de fond (qui englobent d'autres méthodes) dans lequel les objets mobiles sont obtenus par soustraction avec l'image courante. Ceux-ci peuvent et sont donc obtenus de façon différente (comme par exemple par comparaison d'un ratio d'intensité) par les autres méthodes de l'ensemble des "méthodes de suppression de fond".

On trouve également le terme de "méthodes d'extraction de fond" (*Background Extraction"*) qui mettent en jeu le même mécanisme que les méthodes de suppression de fond mais donc l'objectif est centré sur l'obtention de l'image de fond au lieu de l'obtention des objets mobiles.

En travaillant dans la thématique des méthodes de suppression de fond, on rencontre aussi deux thématiques qui sont connexes à celle-ci et qui sont les suivantes :

— La détection de changement entre deux images (*Change Detection*) [422].
— La détection de mouvements saillants (*Salient Motion Detection*) [139].

Les méthodes de suppression de fond sont un cas particulier de la détection de changement entre deux images quand 1) une des images correspond au fond et l'autre image est une image courante comportant des objets mobiles, et 2) les changements entre les deux images sont supposés uniquement dus aux objets mobiles.

Par contre, la détection de mouvements saillants a pour objectif de trouver des régions ayant un sens sémantique dans l'image (humain par exemple) et de filtrer les zones sans intérêts. L'idée est dérivée de la vision humaine. Ainsi, la détection de mouvements saillants peut être considérée comme un cas particulier des méthodes de suppression de fond.

Maintenant que les définitions des différents termes ont été explicitées, nous allons voir dans la section suivante sous quelles hypothèses les méthodes de suppression de fond peuvent bien fonctionner.

2.2 Hypothèses

Pour que l'idée de base fonctionne dans la plupart des cas, cinq principales hypothèses doivent être respectées pour que les seuls changements qui apparaissent au niveau pixel correspondent à un objet mobile d'intérêt :

— ***Hypothèse 1*** : La caméra est fixe et ses caractéristiques le sont également.

— ***Hypothèse 2*** : La scène ne comporte pas de variations lumineuses.

— ***Hypothèse 3*** : Le fond est visible et statique, c'est-à-dire que les pixels ont une distribution unimodale et qu'il n'y a pas d'objet faisant partie du fond qui bouge sous une action extérieure, ou d'objet qui soit déplacé ou inséré dans la scène au cours du temps.

— ***Hypothèse 4*** : Le fond ne contient pas d'*"objets mobiles"* statiques.

— ***Hypothèse 5*** : Les "objets mobiles" peuvent être facilement détéctés en seuillant la différence entre l'image courante et l'image du fond.

En pratique, toutes ces hypothèses ne sont pas respectées à cause de différentes situations qui apparaissent tant dans des scènes en intérieur que dans des scènes en extèrieur. Toutes ces situations sont classés dans la Section 2.3.

2.3 Situations critiques

Sous les conditions idéales indiquées dans la Section 2.2, les méthodes de suppression de fond donnent des résultats satisfaisants. En pratique, plusieurs situations dites critiques apparaissent et pertubent le processus. En 1999, Toyama et al. [521] les ont recensées dans le domaine de la vidéosurveillance. En étendant cette liste sur d'autres domaines d'application, nous avons identifié plus de situations critiques présentées dans la liste suivante [70] :

— ***Bruit :*** Phénomène lié soit à la qualité du capteur caméra (Webcam), soit à la compression de la vidéo. Dans ce cas, de fausses détections apparaissent [280].

— ***Ajustements Automatiques des Paramètres de la Caméra :*** Cela concerne les ajustements automatiques comme l'auto-exposition, l'auto-focus et l'auto-white balance. Dans ce cas, beaucoup de fausses détections sont obtenues dûes au changement global sur toute l'image suite à un auto-réglage. Des solutions spécifiques existent dans chaque cas comme pour l'automatic gain control [260][298] et l'automatic white ba-

lance [615].

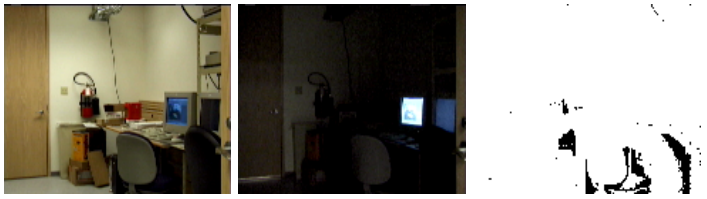
— ***Vibration de la Caméra :*** Des changements du champ visuel de la caméra dus au vent par exemple ont pour conséquence la détection de faux positifs correspondant au contour de l'environnement comme le montre la situation de la Figure 2.6.a. Ce challenge relève du modèle de fond, de la mise à jour ou d'un mécanisme basé sur le flot optique pour compenser le mouvement de la caméra.

— ***Variations Lumineuses Progressives (locales ou globales) :*** Le fond doit s'adapter à ce type de variations causées par exemple par le changement de luminosité au cours du jour comme le montre la Figure 2.3. Ce challenge concerne la mise à jour du modèle.

— ***Variations Lumineuses Soudaines (locales ou globales) :*** Ces variations sont illustrées par une lumière "ON-OFF" dans un environnment à l'intérieur (Figure 2.4) ou un passage nuageux rapide dans une scène filmée à l'extérieur. Ce challenge concerne la mise à jour du modèle ou d'un mécanisme de détection de changement de luminosité au niveau "image".

a) Eclairage faible b) Eclairage modéré. c) Eclairage fort d) Masque de détection

FIGURE 2.3 – De gauche à droite : a) La première image montre une scène d'intérieur avec un éclairage faible. b) La deuxième image montre une scène d'intérieur avec une éclairage modéré alors que la troisième c) montre la même scène avec un éclairage fort. d) La quatrième image montre le masque de détection obtenu avec les MOG [490]. Cette séquence appelée "Time of Day" vient de la base de données Wallflower [521].



a) Lumiére allumée b) Lumière éteinte c) Masque de détection

FIGURE 2.4 – De gauche à droite : a) La première image montre une scène d'intérieur avec l'éclairage allumée. b) La deuxième image montre la même scène d'intérieur avec l'éclairage éteint c) La troisième image montre le masque de détection obtenu avec les MOG [490]. Cette séquence appellée "Light Switch" vient de la base de données Wallflower [521].

— ***Fond Peu Visible ou pas :*** Cette situation se présente quand un objet d'intérêt occupe presque le même endroit durant la phase d'initialisation. Ce challenge est appelé "Bootstrapping" dans la littérature et concerne l'étape d'initialisation du fond [251].

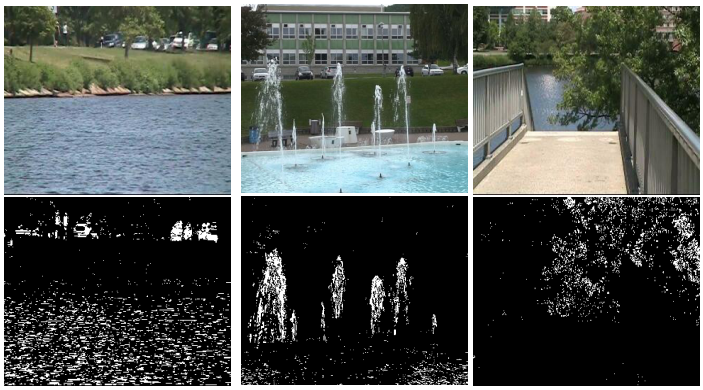
— ***Camouflage :*** Les pixels de l'objet d'intérêt et ceux du fond ont pour une caractéristique donnée la même valeur. Le cas le plus courant est le camouflage en couleur. Dans ce cas, l'objet d'intérêt peut ne pas être détecté entièrement. Figure 2.7.b illustre une situation de camouflage en couleur. Une personne entre dans une pièce avec une valise et la dépose sur le sol. La difficulté réside dans la similarité de couleur entre celle de la valise et du sol. Figure 2.9 et Figure 2.10 montrent respectivement des situations de camouflage en couleur et en profondeur. Ce challenge relève donc du choix de la caractéristique.

— ***Trous dans l'Objet d'intérêt :*** Ce challenge appelé Foreground Aperture [529] se produit quand un objet homogène bouge. Dans ce cas, les changements au niveau des pixels d'intérieur ne sont pas entièrement détectés comme le montre la Figure 2.7.c.

— ***Objet du Fond Déplacé :*** L'objet déplacé appartenant au fond ne doit pas être détecté comme un objet d'intérêt [626][515][535][283]. La Figure 2.6.b présente une scène où une personne déplace une boite. La nouvelle position de la boite doit être intégrée dans le fond rapidement. Ce challenge relève donc de l'étape de mise à jour du fond.

— ***Objet Inséré dans le Fond :*** L'objet doit être intégré dans le fond rapidement [626][515] [535][283] comme les différents objets laissés sur le sofa dans la Figure 2.6.c. Ce challenge concerne l'étape de mise à jour du fond.

— ***Fond dynamique ou multimodal*** : Des exemples de ce mouvement sont illustrés dans la Figure 2.5. On peut y voir trois types de mouvement : les feuilles ou les branches d'un arbre qui bougent à cause du vent, la surface de l'eau qui oscille à cause des



a) Surface de l'eau b) Fontaine c) Arbres qui bougent

FIGURE 2.5 – La première ligne présente des images originales de scènes contenant des fonds dynamiques : a) Boats Sequence , b) Fountain Sequence and c) Overpass Sequence. Ces trois séquences proviennent de la base de données ChangeDetection.net [204]. La seconde ligne montre les masques de détection obtenus avec les MOG [490].

vagues et une fontaine qui fait jaillir l'eau à une certaine hauteur. Ces mouvements qui se produisent dans le fond nécessitent un modèle de fond multimodal pour être bien gérés.

— ***Reprise du Mouvement d'un Objet d'intérêt (initialement immobile) :*** La région du fond qui était occupée par cet objet ainsi que sa nouvelle position sont détectées comme objets d'intérêt comme le montre la Figure 2.7.a. La position initiale de l'objet est appelée "ghost" dans la littérature [581]. Ce challenge concerne l'étape de mise à jour du fond [581].

— ***Ralentissement ou Arrêt d'un Objet d'intérêt :*** Si l'objet d'intérêt ralentit ou s'arrête, il faut ou pas l'intégrer dans le fond selon l'application [257][53]. Ce challenge concerne l'étape de mise à jour du fond [404][514][471][472].

— ***Taille des Objets d'intérêt :*** La taille des objets mobiles peut être problématique comme 1) dans le cas où les objets sont très éloignés de la caméra et sont donc très petits par rapport à la taille de l'image [170][563][411] ou 2) dans le cas où les objets occupent presque la totalité de l'image [177].

— ***Ombres :*** Les ombres sont difficilement distinguables et/ou séparables des objets d'intérêt comme le montre la Figure 2.8. La détection d'ombre est un domaine de recherche en lui-même [413][625][12][442][17].

Le tableau 2.2 montrent une vue concise de ces challenges ainsi que certaines solutions pour les gérer. Des tableaux similaires sont disponibles pour les différentes versions des MOG dans notre publication [75] et pour tous les modèles dans le chapitre de livre [73].

Certaines situations critiques sont donc relatives à la caméra et d'autres à la scène. De plus, elles peuvent être classées de la manière suivante : les variations lumineuses, les variations de mouvement et les variations structurelles de fond.

Toutes ces situations critiques ont des caractéristiques spatiales et temporelles différentes. Par exemple, les variations lumineuses peuvent être locales ou globales spatialement et graduelles ou soudaines temporellement. De plus, les changements lents ont une variation continue alors que les changements rapides et surtout ultra-rapides présentent une variation discrète. En outre, certains challenges présentent une variation unimodale alors que d'autres présentent des variations multimodales. On peut avoir également des variations locales ou globales de mouvement. Toutes ces situations critiques et leurs caractéristiques particulières doivent être prises en compte à chaque étape lors de la conception d'une méthode de suppression de fond pour obtenir une bonne segmentation.

Les premières méthodes ont été conçues pour être robustes à tel ou tel challenge rencontré dans une application spécifique alors que maintenant les méthodes tentent à être développées pour être robustes à la plus grande partie de ces challenges. Il existe d'ailleurs des outils récents de visualisation (TimeSearcher [424], TimeViewer [433]) permettant de voir les variations du pixel au cours du temps lors de ces différents challenges, facilitant ainsi la conception d'un mécanisme robuste à ceux-ci.

En outre, cette conception doit être réalisée dans le contexte d'applications temps réel qui doivent fonctionner sur un PC classique ou sur un système embarqué. Ainsi deux contraintes

Challenges	Solutions
Image bruitée	Modélisation [84][280], Caractéristiques [175][63]
Ajustements automatiques de la caméra	Background Maintenance [615], Caractéristiques [260]
Vibration de la caméra	*Modélisation :* Statistiques [490][462][537][513] Statistiques Avancés[231][155][338][51][248], Flous [46]
Variations lumineuses	*Modélisation :* Réduction de dimension[387][523][83][315], Filtres [521][366][103] RPCA [92][226][156][643][240], Parcimonieux [255][254][628][575] Réduction de dimension discriminatifs et mixtes [177][359] Caractéristiques [455]
Bootstrapping	Initialisation [323][249], Modélisation [84][280][454][567] Caractéristiques [249][7], Stratégies [427]
Camouflage	Caractéristiques [202][239]
Foreground aperture	Maintenance [529]
Objets du fond déplacés	Maintenance [626][515][535]
Objets du fond insérés	Maintenance [626][515][535]
Fonds dynamiques	*Modèlisation :* Statistiques [490][167][513] Statistiques Avancés [231][155][338][51][248] Flous [46], Transformation de domaine[563][411] Caractéristiques [245][623][539][577][576][645][307][385]
Reprise du mouvement d'un objet (Ghost)	Maintenance [404][514][383]
Arrêt du mouvement d'un objet	Maintenance [404][514][383]
Ombres	Caractéristiques [413][625][12][442][17][57][492], Stratégies [492]

TABLE 2.2 – Les challenges et certaines solutions pour les gérer. La première colonne indique le challenge et la seconde indique certaines solutions permettant de les gérer et à quelle étape.

apparaîssent : le temps de calcul qui doit être le plus rapide possible et l'espace mémoire employé qui doit être le plus faible possible.

Dans la section suivante (Section 3), nous allons d'abord présenter une vue générale des différents modèles existants pour la modélisation, l'initialisation et la mise à jour du fond. Ensuite, nous verrons les différents modèles de classification Fond/Objet mobiles employés jusque maintenant dans la Section 3.4. Puis, dans la Section 3.5 et la Section 3.6, nous verrons comment la taille et le type des caractéristiques employées peuvent influer également sur la robustesse. Dans la Section 4, nous présentons une vue plus détaillée des différents modèles de fonds et nous abordons aussi leur robustesse aux différents challenges. Cette section est aussi l'occasion de situer notre travail dans le domaine comme nos états de l'art ainsi que les trois approches développées sous ma direction scientifique.

Ensuite, les différentes stratégies d'applications de ces modèles sont présentées dans la

Section 5. Nous verrons aussi les différentes approches temps-réel dans la Section 6 ainsi que les différents capteurs employés dans la Section 7. Finalement, nous aborderons la problématique de l'évaluation des performances des algorithmes de suppression de fond. Pour cela, nous présentons d'abord les mesures existantes avec ou sans vérité-terrain dans la Section 8, et ensuite les bases de données existantes dans la Section 9.

a) Traffic b) Aband. Box c) Sofa

FIGURE 2.6 – La première ligne présente respectivement des images originales de scènes contenant un mouvement de camèra, un objet déplacé et un objet inséré dans le fond. a) Traffic, b) Abandoned Box et c) Sofa. Ces trois séquences proviennent de la base de données ChangeDetection.net [204]. La seconde ligne montre les masques de détection obtenus avec les MOG [490].

a) Parking b) Camouflage c) Fore. Apert.

FIGURE 2.7 – La première ligne présente respectivement des images originales de scènes contenant des objets mobiles qui commencent à se déplacer (une voiture), une situation de camouflage et une situation de "foreground aperture". a) Parking [204], b) Camouflage [182] et c) Foreground Aperture [521]. La seconde ligne montre les masques de détection obtenus avec les MOG [490].

a) Backdoor b) Bungalows c) PeopleInShade

FIGURE 2.8 – La première ligne présente respectivement des images originales de scènes contenant des ombres : a) Backdoor, b) Bungalows et c) PeopleInShade. Ces trois séquences proviennent de la base de données ChangeDetection.net [204]. La seconde ligne montre les masques de détection obtenus avec les MOG [490].

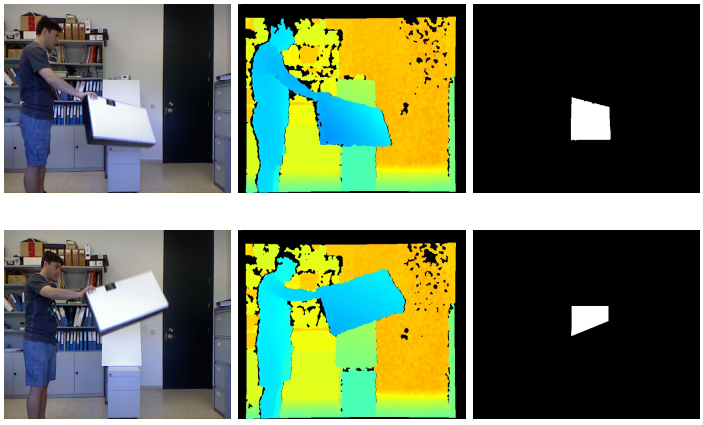


FIGURE 2.9 – La première colonne montre des images originales de scènes contenant des camouflages en couleur. La seconde colonne montre les images de profondeur qui permettent de gérer le camouflage en couleur. La troisème colonne montre les vérités-terrain. Cette séquence appelée "ColCamSeq" vient de la base de données RGB-D Dataset [89].

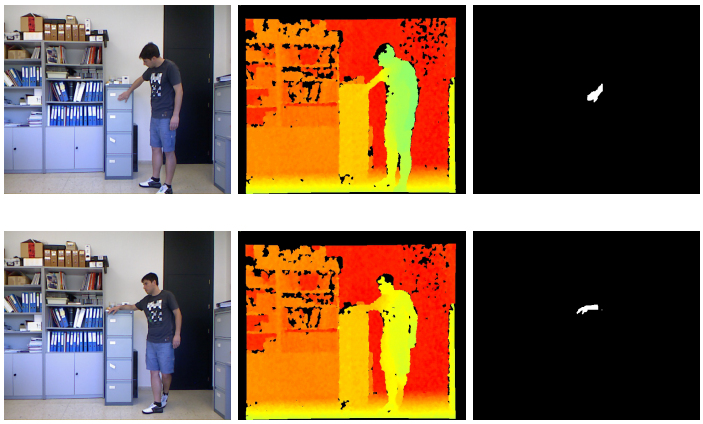


FIGURE 2.10 – La première colonne montre des images originales de scènes contenant des camouflages en profondeur. La seconde colonne présente les images de profondeur correspondantes qui montrent le camouflage en profondeur. La troisème colonne montre les vérités-terrain. Cette séquence appelée "DCamSeq" vient de la base de données RGB-D Dataset [89].

3 Suppression de fond : Les différentes étapes

Les sections ci-dessous donnent un aperçu global des différentes étapes et des points clés des méthodes de suppression de fond. Les tableaux 2.3, 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 2.10 et 2.11 donnent une vue synthétique des méthodes développées pour ces étapes et points clés.

Remarque : Les différents tableaux fournis comportent différents acronymes correspondant à des méthodes. Pour chaque méthode, la référence fournie est celle de l'article original. Le lecteur pourra trouver la signification des acronymes soit dans le texte de ce manuscrit soit sur le site "Background Subtraction Web Site", et de la même façon le lecteur pourra avoir plus de références vers d'autres publications pour la méthode concernée.

3.1 Modélisation de fond

La modélisation de fond décrit le modèle utilisé pour représenter le fond. Ce modèle de fond conditionne essentiellement la capacité du modèle à gérer les fonds uni-modaux ou multi-modaux. Dans la Section 4, les différents modèles de la littérature sont classés en modèles actuels et prospectifs :

— **Les modèles actuels** comprennent les modèles actuellement utilisables dans les systèmes de vidéo-surveillance et par conséquent les premiers modèles développés dans le domaine. Ils ont été souvent élaborés pour prendre en compte des challenges spécifiques. De plus, un grand nombre d'améliorations a été apporté à leur version originale et il semble que leurs limites soient quasiment atteintes. Ces modèles sont regroupés dans la Section 4.1 et le tableau 2.3.

— **Les modèles prospectifs** sont des modèles plus sophistiqués conçus pour prendre en compte un grand nombre de challenges. Par contre, ils nécessitent encore des recherches pour être incrémentaux et atteindre des performances temps-réel. Ces modèles sont regroupés dans la Section 4.2 et les tableaux 2.4 et 2.5.

Les tableaux 2.3, 2.4, 2.5 et 2.6 montrent une vue synthétique de ces différents modèles de fond. Les approches développées sous ma direction scientifique sont 1) les approches floues (T2-FMOG-UM [46], T2-FMOG-UV [46]), 2) les approches discriminatives et mixtes de réduction de dimension (IMMC [177], PCA-ILDA [359]) et 3) des variantes de l'ACP robuste (IRLS [226], Polynomial RPCA [228]). Les travaux 1) et 2) sont des travaux précurseurs dans le domaine et sont développés respectivement dans le Chapitre 3 et le Chapitre 4. Les travaux concernant les variantes des récentes avancées en ACP robustes sont développés dans le Chapitre 5.

3.2 Initialisation du fond

L'initialisation du fond consiste à déterminer la première image de fond. Contrairement à la modélisation de fond et la mise à jour du fond, relativement peu d'articles portent sur cette étape. La principale raison réside dans l'hypothèse que la première image de fond peut être obtenue soit en prenant la première image courante soit en utilisant des images sans objets mobiles au début du processus. Cependant, le challenge principal est d'obtenir une image de fond correcte alors que des objets mobiles apparaissent plus de la moitié du temps dans la séquence d'apprentissage. L'initialisation du fond est un processus d'apprentissage et peut être effectué hors ligne. Ainsi, l'algorithme peut être "batch". Pratiquement, les algorithmes peuvent être classés en trois catégories : 1) "batch" en utilisant N images d'apprentissage qui sont consécutives ou non [371], 2) incrémental avec N images d'apprentissage avec N fixés et 3) progressif avec N pas fixé jusqu'à l'obtention d'une image de fond complète [127][132]. De plus, les algorithmes différent suivant la modalité du fond [414]. Les principales recherches concernent la médiane [371], l'histogramme [127][132], les intervalles stables [340][536][220][107] et les SVM [323]. Le tableau 2.7 montre une vue synthétique de ces différents modèles d'initialisation.

3.3 Maintenance du fond

La maintenance du fond concerne le mécanisme de mise à jour du modèle de fond pour s'adapter aux changements qui apparaissent au cours du temps. La maintenance doit être effectuée en direct (on-line) et donc l'algorithme doit être incrémental. Les principales caractéristiques clés de cette étape sont les suivantes :

— **Modèle de maintenance :** Dans la littérature, il existe trois modèles de maintenance : le non-sélectif, le sélectif et l'adaptatif flou [43]. Le modèle non-sélectif consiste à mettre à jour tous les pixels avec la même règle qui est couramment un filtre à réponse impulsionnelle fini comme suit :

$$B_{t+1}(x,y) = (1-\alpha)\, B_t(x,y) + \alpha I_t(x,y) \tag{2.1}$$

où α est un coefficient d'apprentissage qui est généralement constant dans l'intervalle $[0,1]$. L'inconvénient de ce modèle est que la valeur des pixels détectés comme "objet mobile" est utilisé dans le calcul de la nouvelle image du fond. Pour résoudre ce problème, certains auteurs ont proposé un modèle sélectif qui consiste à calculer l'image de fond avec un coefficient d'apprentissage différent suivant que le pixel ait été détecté comme "objet mobile" ou "fond" comme suit :

$$\begin{aligned} B_{t+1}(x,y) &= (1-\alpha)\, B_t(x,y) + \alpha I_t(x,y) \\ &\quad if\ (x,y)\ is\ background \end{aligned} \tag{2.2}$$

$$\begin{aligned} B_{t+1}(x,y) &= (1-\beta)\, B_t(x,y) + \beta I_t(x,y) \\ &\quad if\ (x,y)\ is\ foreground \end{aligned} \tag{2.3}$$

L'idée est de mettre à jour très rapidement les pixels classés comme "fond" et très lentement ceux qui sont classés comme "objet mobile". Pour cela, $\beta << \alpha$ et couramment $\beta = 0$. Ainsi l'équation précédente devient :

$$B_{t+1}(x, y) = B_t(x, y) \tag{2.4}$$

Cependant, les fausses classifications peuvent conduire à une image de fond incorrecte. Ce problème peut être résolu par un modèle adaptatif flou qui prend en compte l'incertitude de la classification du pixel. Cela peut se faire en graduant la règle de mise à jour en utilisant le masque de détection comme dans [43].

— **Coefficient d'apprentissage :** Le coefficient d'apprentissage détermine la vitesse d'adaptation aux changements qui apparaissent au cours du temps. Il peut être constant ou déterminé de façon statistique ou floue. Dans le premier cas, il est généralement identique sur toute l'image et toute la séquence. Ce qui n'est pas optimal car chaque zone de l'image présente des activités différentes et donc nécessite un seuil adapté. Ce coefficient est généralement déterminé de façon expérimentale [613] ou de façon automatique par un algorithme d'optimisation [559]. On peut aussi faire varier le coefficient suivant la phase d'apprentissage et la phase de mise à jour [267]. On peut également utiliser la phase de suivi pour l'adapter [283]. Pour le cas statistique, Lee [305] emploie des coefficients différents pour chaque Gaussiennes améliorant ainsi la convergence. Pour le cas flou, Sigari el al. [477][476] ont proposé un coefficient adaptatif calculé en fonction du degré d'appartenance à la classe "fond" ou "objet mobile". Récemment, Maddalena et Petrosino [353][352] ont amélioré cette adaptativité en utilisant l'information de cohérence spatiale.

— **Mécanisme de maintenance :** Le coefficient d'apprentissage détermine la vitesse d'adaptation aux variations lumineuses mais aussi la vitesse d'incorporation des objets déplacés ou insérés dans la fond et la vitesse d'incorporation des objets mobiles qui s'arrêtent. Ainsi, un coefficient gère différents processus qui ont des caractéristiques temporelles différentes. Pour séparer le mécanisme d'adaptation du mécanisme d'incorporation, certains auteurs [626][534] ont utilisés un ensemble de compteurs qui compte le nombre de fois où un pixel a été classé comme "objet mobile". Quand ce nombre est plus grand qu'un seuil, l'objet est incorporé dans le fond et donc le pixel est classé comme "fond".

— **Fréquence de mise à jour :** L'objectif est de mettre à jour le fond seulement quand cela est nécessaire. La mise à jour doit être idéalement effectuée à chaque instant t mais dans le cas de changement faible cela peut être différé. Par exemple, Porikli et al. [407] détermine la fréquence de mise à jour en fonction d'un indicateur de changement lumineux. L'idée est que la mise à jour est nécessaire que si une variation lumineuse est détectée. Dans la même idée, Magee [356] a utilisé un indicateur d'ac-

tivité du pixel permettant de mettre à jour lentement les pixels stables et rapidement les pixels moins stables.

Le tableau 2.8 montre une vue synthétique de ces différents modèles de mise à jour. L'approche développée sous ma direction scientifique est celle concernant la règle de mise à jour adaptive [43]. Ce modèle est précurseur dans le domaine et est développé dans le Chapitre 3.

3.4 Classification fond/objets mobiles

Cette classification est obtenue à l'aide de la comparaison entre l'image de fond et l'image courante. Pour cela, on peut utiliser des outils de classification classiques, statistiques ou flous. Les points clés de cette étape sont les suivants :

— **Pré-traitement :** Le pré-traitement permet d'éviter de fausses détections dues à des changements comme ceux engendrés par le mouvement de la caméra et les variations lumineuses. Cela peut se faire en faisant des corrections géométriques [422] et des corrections lumineuses [520][519][390] entre les deux images.

— **Test :** Le test qui permet de classer un pixel comme appartenant au fond ou à un objet mobile est couramment la différence absolue entre l'image de fond et l'image courante. Cette différence est ensuite seuillée pour obtenir le masque de détection. D'autres possibilités existent comme les tests statistiques [568][480][5][4][3][367][2][6].

— **Seuil :** Il y a plusieurs façons de seuiller. La première consiste à prendre un seuil fixe et constant pour toute l'image et toute la séquence. Ce schéma est simple mais bien sûr pas optimal dans la mesure où tous les pixels n'ont pas la même activité suivant la zone dans laquelle ils se trouvent. Pour résoudre cela, un seuil adaptatif spatio-temporellement doit être utilisé et mis à jour au cours du temps. On peut ainsi le faire en utilisant la variance d'un pixel [562]. On peut le faire également en utilisant une méthode floue [99].

— **Post-traitement :** L'idée est d'améliorer le masque de détection en éliminant les pixels isolés avec un opérateur morphologique classique ou statistique [491]. Un autre moyen est d'utiliser des règles d'inférence floue [481].

De plus, cette phase de détection d'objets mobiles (foreground detection) est un cas particulier de la détection de changement entre deux images (change detection) dans le cas où 1) une image est le fond qui est modélisé et mis à jour, et l'autre est l'image courante, 2) le changement est du aux objets mobiles. Ainsi, toutes les techniques développées dans le cadre de la détection de changement entre deux images peuvent être utilisées dans cette phase de détection d'objets mobiles. Des états de l'art se trouvent dans [422][437].

Le tableau 2.8 montre une vue synthétique de ces différents méthodes de classification fond/objets mobiles. L'approche développée sous ma direction scientifique est celle concernant l'utilisation de l'intégrale de Choquet [42] pour agréger les résultats provenant de différentes caractéristiques. Ce modèle basé sur l'approche utilisant l'intégrale de Sugeno précédemment développée par Zhang et Xu [620] est présenté dans le Chapitre 3.

3.5 Taille de la sous-entité de l'image

Cette sous-entité de l'image qui va servir à la comparaison peut être de la taille d'un pixel [490], d'un bloc [175] ou d'un cluster :

— **Pixel :** Le modèle de fond est appliqué comme un modèle indépendant des pixels voisins. C'est le cas le plus courant mais cela ne prend pas en compte les contraintes spatiales et temporelles.

— **Bloc :** Les blocs peuvent être disjoints ou pas [217]. Le bloc est couramment obtenu à partir d'un vecteur voisinage 3×3 du pixel courant. L'intérêt est de prendre en compte l'aspect spatial pour améliorer la robustesse et cela permet aussi de réduire le temps de calcul. De plus, des blocs spatio-temporels peuvent être utilisés [405]. Dans ce cas, une technique de réduction de dimension (ACP) est employée pour obtenir une représentation compacte de chaque bloc. Ces blocs fournissent à la fois une information sur la texture et le mouvement, donnant ainsi une meilleure robustesse aux bruits et aux fonds multi-modaux. L'inconvénient est une détection moins précise car seulement des blocs sont détectés, rendant ainsi la méthode inappropriée dans le cas d'application nécessitant l'information de forme.

— **Cluster :** Dans ce cas, les clusters sont obtenus par une méthode de segmentation couleur [63]. Ensuite, chaque cluster contient des pixels ayant quasiment les mêmes valeurs dans l'espace couleur HSV. Enfin, le modèle de fond est appliqué sur chaque cluster dont les pixels sont ensuite classés en "fond" ou "objet mobile". Contrairement aux approches blocs, la détection obtenue est de l'ordre de la précision du pixel.

Finalement, le taille de l'élément détermine essentiellement la robustesse de la méthode aux bruits et aussi sa précision. Une méthode pixel est plus précise mais moins robuste aux bruits qu'une méthode bloc ou cluster. Le tableau 2.9 montre une vue synthétique des différentes tailles utilisées dans la littérature.

3.6 Choix de la caractéristique

La caractéristique caractérise la sous-entité de l'image choisi (pixel, bloc ou cluster). Dans la littérature, on trouve couramment les caractéristiques suivantes : couleur, contour,

stéréoscopique, mouvement et texture. Ces caractéristiques sont souvent discrètes et certaines d'entre elles peuvent être obtenues à partir de concept statistique ou flou :

— **Caractéristiques de bases :** La couleur dans l'espace RGB est la caractéristique qui est la plus employée car elle est directement fournie par les caméras. Cependant, l'espace RGB comporte plusieurs inconvénients : les trois composantes sont dépendantes et sensibles aux variations lumineuses. Par exemple, si une zone est couverte par une ombre portée, les trois composantes sont affectées car les informations de luminance et de chrominance ne sont pas séparées. Des espaces alternatifs comme YUV or YCrCb peuvent être employés. Des comparaisons entre ces différents espaces existent dans la littérature [293][431][270] et quasiment dans toutes ces comparaisons, YCrCb apparaît comme l'espace couleur le plus approprié. Bien que la couleur soit souvent discriminative pour détecter les objets, elle présente plusieurs limitations en présence de certains challenges comme les variations lumineuses, le camouflage et les ombres. Pour résoudre cela, certains auteurs ont employé d'autres caractéristiques comme le contour ou la texture en plus de la couleur. Dans le cas des contours, ceux-ci sont calculés à partir d'une approche gradient comme le détecteur de Canny. L'avantage des contours apparaît lors des variations lumineuses et dans le cas de détection "ghost" quand un objet à l'arrêt commence à se déplacer. Dans le cas des caractéristiques stéréoscopiques, deux caméras doivent être utilisées pour obtenir l'information de disparité ou de profondeur. Celles-ci sont efficaces quand il y a camouflage en couleur. L'information de mouvement qui est obtenue par flot optique permet quant-à-elle de gérer les fonds dynamiques. Par contre, son temps de calcul en limite son usage. Enfin, la texture est appropriée quand il y a des variations lumineuses et des ombres. Les caractéristiques "texture" les plus utilisés est le Local Binary Pattern (LBP) [245] et le Local Ternary Pattern (LTP) [321]. Un grand nombre de variantes de LBP et LTP ont été développées comme le montre le tableau 2.9 dans la ligne "Type de la Caractéristique".

— **Caractéristiques statistiques :** Les textures peuvent nécessiter des propriétés statistiques dans leur calcul. La première développée par Satoh et al. [447] est appelée Radial Reach Correlation (RRC) qui comporte plusieurs variantes : Bi-polar Radial Reach Correlation (BP-RCC) [446], Fast Radial Reach Correlation (F-RRC) [258], Probabilistic Bi-polar Radial Reach Correlation (PrBP-RCC) [598]. De façon similaire, Yokoi [597] a proposé le Peripheral Ternary Sign Correlation (PTESC). Récemmment, Yoshinaga et al. [600] ont développé le Statistical Local Difference Pattern (SLDP). L'objectif de calculer ces caractéristiques est d'être plus robuste dans le cas de variations lumineuses.

— **Caractéristiques floues :** Les textures peuvent nécessiter des propriétés floues dans

leurs calculs. Par exemple, Chiranjeevi et Sengupta ont introduit l'utilisation de corrélogrammes flous [119], de textures statistiques et floues [120] et de 3D Histons flous [121]. Dans ce cas, l'idée est d'être à la fois robuste aux variations lumineuses et aux fonds multi-modaux.

Le tableau 2.9 montre une vue synthétique de ces différentes caractéristiques utilisées dans la littérature. Toutes ces caractéristiques ont des propriétés spatiales ou temporelles qui leur permettent d'être robustes à tel ou tel challenge. De plus, dans le cas de plusieurs caractéristiques, un opérateur d'agrégation doit être employé pour combiner les résultats. Les opérateurs logiques ET et OU sont souvent employés. Cependant, les opérateurs flous comme l'intégrale de Sugeno [620] ou l'intégrale de Choquet [42] sont les plus appropriés. Les stratégies multi-caractéristiques sont abordées dans la Section 5. Un autre point clé encore peu développé est la sélection de caractéristiques [168] permettant d'employer les caractéristiques les plus adaptées pour une zone de l'image où apparait tel ou tel challenge.

3.7 Autres points clés

Les autres points clés du domaine de la détection d'objets mobiles sont les suivants :

— **Domaine des données :** La détection d'objets mobiles s'effectue usuellement dans le domaine pixel mais on peut également faire cette détection dans le domaine de la mesure [241] ou dans le domaine compressé (H.264 [151], HEVC [630]).
— **Stratégies :** Plusieurs stratégies peuvent être mises en oeuvre pour appliquer et améliorer la robustesse des différents modèles de fond vis-à-vis des situations critiques. Il est important de noter que la plupart du temps ces stratégies sont indépendantes du modèles et peuvent donc être appliquées quelque soit le modèle de fond (voir Section 5).
— **Implémentation Temps-Réel :** La détection des objets mobiles doit être réalisée la plupart du temps dans des applications dites temps-réel. Ainsi, le temps de calcul et l'espace mémoire requis doivent être réduits le plus possible. Cela peut être résolu en partie par des implémentations GPU (voir Section 6).
— **Evaluation des performances :** L'idée est de comparer de façon quantitative les différents modèles de fond. L'évaluation des performances peut s'exprimer en temps de calcul ou en espace mémoire requis mais aussi en termes de fausses détections. (voir Section 8).
— **Bases de données :** Les bases de données de séquence vidéos acquises par des caméras fixes sont importantes pour faire une comparaison complète, fiable et juste des différents algorithmes (voir Section 9).

Les tableaux 2.10 et 2.11 montrent une vue synthétique des différents développements concer-

nant ces points clés.

Modélisation de Fond	**Sous-categories**	**Méthodes**
Modèles Actuels	Basiques	Mean [66], Median [362], Histogram [635]
Survey (Bouwmans (2014) [73])	Statistiques	*1) Gaussien*
	Survey (Bouwmans (2010) [76])	SG [562], SGG [279]
	Survey (Bouwmans (2011) [70])	MOG [490][93] (**Survey (Bouwmans et al. (2008) [75])**), MOGG [16]
		KDE [167]
		2) Vecteurs de Support
		SVM [323], SVR [537], SVDD [513]
		*3) Réduction de dimension (**Survey (Bouwmans (2011) [69])**)*
		3.1) Matrices
		PCA [387], ICA [579], INMF [83], LoPP [291]
		3.2) Tenseurs
		IRT [315][21], TenLoPP [292], DB [165]
	Clusters	K-means [84], Codebook [280], Basic Sequential Clustering [567]
	Réseaux de Neurones	General Regression [141][142], Multivalued [346], Competitive [343]
		Dipolar Competitive [345]
		Self Organizing (SOBS [350], SOBS-SC [354], 3dSOBS+ [355])
		Growing Hierarchical Self Organizing [391]
		Adaptive Resonant Theory [344]
		Artificial Neural Network [473]
		Multilayer Neural Network [423], Neural Network Mapped [28]
		Nonlinear Autoregressive (NARX) [606]
		Discrete Time Cellular Neural Network [101]
		Weightless Neural Networks [206]
	Estimation	Wiener filter [521], Kalman filter [271], Correntropy filter [172]
		Chebychev filter [103]

TABLE 2.3 – Vue d'ensemble des méthodes actuelles de modélisation de fond. Mes états de l'art sont indiqués en gras.

Modélisation de Fond	**Sous-categories**	**Méthodes**
Modèles Prospectifs **Survey (Bouwmans (2014) [72])**	Statistiques	*1) Mixture Models* Student's t [374] DMM [231], VarDMM [173] APMM [178] *2) Hybrid Models* KGMM [155], KGHM [337] *3) Non Parametric Models* ViBe [51], PBAS [248], VuMeter [203]
	Fuzzy Models **Survey (Bouwmans (2012) [71])**	1) Fuzzy Color Histogram FCM [286], FCH [269], MHC [590] 2) Fuzzy Mixture of Gaussians **T2-FMOG-UM** [46], **T2-FMOG-UV** [46] 3) Fuzzy Neural Network Fuzzy SOM [98], Adaptive Fuzzy SOM [100] Fuzzy ART [105], Spatial Fuzzy ART[106] 4) Fuzzy Estimation TSK (Takagi-Sugeno-Kang) [607]
	Discriminative Subspace Models	**IMMC** [177]
	Mixed Subspace Models	**PCA-ILDA** [359]
	Kernel Subspace Models	Euler PCA (e-PCA) [339]
	Robust Subspace Models (via Outliers Suppression)	RSL [517][518] R1-PCA [154], PCA-l_1 [299], Half-Quadratic-PCA (HQ-PCA) [243]

TABLE 2.4 – Vue d'ensemble des méthodes prospectives de modélisation de fond (Partie 1). En gras, les approches développées sous ma direction scientifique ainsi que mes états de l'art.

Modélisation de Fond	**Sous-categories**	**Méthodes**
Modèles Prospectifs **Survey (Bouwmans (2014) [72])**	Robust Subspace Models (via Decomposition in Low Rank and Sparse Matrices)	*1) Robust Principal Components Analysis (RPCA)* **(Survey (Guyon et al. (2012) [227])** *1.1) Principal Component Pursuit* **(Survey (Bouwmans and Zahzah (2014) [79])** PCP [92], SPCP [647] QPCP [55], BPCP [509], LPCP [560] *1.2) Outlier Pursuit* OP [570], SOP [570] *1.3) Sparsity Control* SpaCtrl [361], SpaCorr [252] *1.4) Log-sum heuristic Recovery* LHR [150] *1.5) Iteratively Reweighted Least Square* **IRLS** [226], **Spatial IRLS** [224], **Spatial and temporal IRLS** [223] *1.6) Bayesian RPCA* Bayesian RPCA [156], Variational Bayesian RPCA [36] Factorized Variational Bayesian RPCA [10] Variational Bayesian based Sparse Estimator (VBSE) [113][112] *1.7) Approximated RPCA* GoDec [643], Semi-Soft GoDec [643] *1.8) Sparse Additive Matrix Factorization* SAMF [35] **1.9) Polynomial RPCA** [228] *2) Robust Non-negative Matrix Factorization (RNMF)* MahNMF [212], NS-NMF [295] *3) Others Subspace Learning* *3.1) Robust Subspace Discovery* RSD [548] *3.2) Robust Subspace Recovery via Bi-Sparsity* RoSuRe [64] *3.3) Robust Orthonomal Subspace Learning* ROSL [474], ROSL+ [474]
	Subspace Tracking	GRASTA [240], t-GRASTA [242], GOSUS [571], FARST [9] pROST [230][451]
	Low Rank Minimization	DECOLOR [646] DRMF [569], DRMF-R [569] PRMF [542], BRMF [543], MBRMF [543] PLRMF [636], LRMF-MOG [365], UNN-BF [85] RRF [465]
	Matrix Completion	FRMC [357], RMC[592], MAMC[583]
	Sparse Models	Compressive Sensing [97] Structured Sparsity [255], Dynamic Group Sparsity [254] Dictionary Learning [147], Sparse Error Estimation [153]
	Robust Tensor Models (via Decomposition in Low Rank and Sparse Tensors)	*1) RPCA Tensor* LTR-MALM [501][502], MALM-TR [503] T-RPCA [522], SO-IR [312], RSTD [318] *2) NMF Tensor* RTF [128]
	Transform Domain	FFT [563], DCT [411] Walsh [516], Wavelet [193], Hadamard [49]

TABLE 2.5 – Vue d'ensemble des méthodes prospectives de modélisation de fond (Partie 2). En gras, les approches développées sous ma direction scientifique ainsi que mes états de l'art.

Modélisation de Fond	Sous-categories	Méthodes
Modèles "Autres"	Basiques	Pixel Intensity Classification (PIC) [250][629][601]
		Pixel Change Classification (PCC) [188]
	Statistiques	*1) Gaussien*
		Co-occurrence [319]
		Multi-Kernel (WGKVS [370])
		2) Réduction de dimension
		Pursuing Dynamic Spatio-Temporal Models (STDM [324])
		3) Non Parametric Models
		Projection Pursuit Density Estimation (PPDE [400])
	Estimation	Mean Shift [403][320][236]
		Particle Filter [618]
		Radial Reach Filter [448][445][443]
	Outlier Detection	Local Outlier Factor (LOF) [397]
		Connectivity-based Outlier Factor (COF)[397]
		incLOF [397], incCOF[397][428]

TABLE 2.6 – Vue d'ensemble des méthodes "Autres" de modélisation de fond.

Initialisation du Fond	Sous-categories	Méthodes
Batch Models	Modèles Basiques	Mean, Median [371]
Progressive Models	Basiques	Histogram [127][132]
	Intervalles Stables	Smoothness Detector (SD) [340]
		Adaptive Smoothness Detector (ASD)[340]
		Local Image Flow (LIF) [220], Optical Flow [107]
		Adaptive-Scale Sample Consensus (ASSC) [536]
		Dynamic Analysis [122], Spatio-Temporal Similarity [393]
	Statistiques	SVM [323]
		ECD based Mean Shift [336], Wronskian based SG [493]
		Image Content Sensitivity based MOG [414]
	Transform Domain	Block-based DCT [427]
		Block-based Hadamard Transform [49]
	Consecutive Frames	Block-based Sum of Absolute Differences (SAD) [514]
		Patch-based Sum of Squared Distances (SSD) [131]
		Moving Object Region Detection by Difference Product (DP) [229]
		Evolutionary Algorithm (EA) [284]
		Labeling Cost [129]
Strategical Models	Strategical Models	Markov Random Fields (MRF) [427], Multi-Levels [135]

TABLE 2.7 – Vue d'ensemble des méthodes d'initialisation du fond.

Categories	**Sous-categories**	**Méthodes**
Maintenance du Fond	Règle de Maintenance	IIR filter (MOG [490])
		Kalman filter (MOG [235])
	Modèle de Maintenance	Blind update
		Selective update
		Fuzzy adaptive update [43]
	Coefficient d'Apprentissage	*1) Fixed learning rate*
		One value (Good settings [613], Optimized settings [559])
		2) Variable learning rate
		Two values (Initialization-Maintenance [267], Tracking feedback [283])
		3) Statistical learning rate [305] [239] [329]
		4) Fuzzy learning rate
		Exponential function [477][476][456], Adaptive exponential function [594]
		Saturating linear function [351], Saturating linear function - Spatial coherence [353][352]
	Mécanisme	Set of counters [626][534]
		CUSUM detector [329]
	Fréquence	Illumination changes score [407]
		Pixel activity [356]
Classication Fond/Objets Mobiles	Pré-traitement	1) Corrections géométriques [422]
		2) Corrections de luminosité
		Intensity Normalization [422]
		Homomorphic Filtering [520][519][390]
	Test	*1) Difference*
		2) Test Statistique
		Bhattacharyya Distance Test [568], Fast Significance Test [480]
		Probability Test [5][4][3]
		Total Least Squares Test [367][2][6]
		3) Test de Dépendance Linéaire
		Linear dependence detector (LDD) [163], Wronskian Detector (WD) [164]
		Linear Approximation Model (LAM) [191]
	Seuil	*1) Seuil Fixe*
		2) Seuil Statistique
		Adaptive Threshold [480][4][3][367][2][6], Spatio-Temporal Adaptive Threshold [519]
		Significance Invariance [1]
		Standard Deviation [130], Variance [562]
		3) Seuil Flou [99]
	Post-traitement	Opérateurs Morphologiques
		Opérateurs Morphologiques Statistiques [491]
		Règle d'inférence floue [481]

TABLE 2.8 – Vue d'ensemble des méthodes de mise à jour du fond et des méthodes de classification des pixels. En gras, l'approche développée sous ma direction scientifique.

Categories	Sous-categories	Méthodes
Taille de la Caractéristique	Pixel	Taille la plus courante
	Bloc	Blocks [217]
		Spatiotemporal Blocks [405]
	Cluster	Color Clusters [63]
Type de la Caractéristique	Caractéristique	*1) Couleur*
		RGB, YUV [174], YCrCb [293], HSV [632], HSI [546]
		Normalized RGB [572], Luv [588], Improved HLS Color Space [452]
		Cylindrical RGB [280], Conical RGB [256]
		2) Histogramme
		LH [640], LDH [624]
		3) Contours
		Gradient [261], Histogram of Oriented Gradient (HOG) [400] [262], SIFT Flow [162]
		4) Texture
		4.1) Local Binary Pattern (LBP
		LBP [245], STLBP [623], ELBP [539], SCS-LBP [577], HCS-LBP [576]
		SC-LBP [645], OC-LBP [307], SA-LBP [385]
		U-LBP [602], SLBP-AM [596], MV-LBP [544]
		4.2) Local Ternary Pattern (LTP)
		LTP [321], SILTP [347]
		MC-SILTP [347], SCS-LTP [627], CS-STLTP [324]
		4.3) Local States (LS)
		LS [605]
		4.4) Local Neighborhood Patterns (LNP)
		LNP [18]
		4.5) Textons
		Textons [488]
		5) Stereo
		Disparity [202], Depth [239][479][303][302]
		6) Mouvement
		Optical flow [642]
		7) Caractéristique Robuste
		Speeded Up Robust Features (SURF) [453]
	Caractéristique Statistique	*1) Texture*
		1.1) Peripheral Increment Sign Correlation (PISC)
		PISC [444]
		1.2) Radial Reach Correlation (RRC)
		RRC [447], BP-RCC [446], F-RRC [258], PrBP-RCC[598], PTESC [597]
		1.3) Peripheral Ternary Sign Correlation (PTESC)
		PTESC [597]
		1.4) Local Difference Patterns (LDP)
		LDP [599], SLDP [600]
	Caractéristique Floue	*1) Histogramme*
		Fuzzy 3D Histons [121]
		2) Texture
		Fuzzy Correlograms [119], Fuzzy Statistical Texture Features [120]
Agrégation de Caractéristiques	Opérateur Classique	Moyenne [288]
	Opérateur Logique	OU, ET
	Opérateur Flou	Sugeno integral [620]
		Choquet integral [42]

TABLE 2.9 – Vue d'ensemble des caractéristiques utilisées dans la littérature. En gras, l'approche développée sous ma direction scientifique.

Stratégies	**Sous-catégories**	**Méthodes**
Multi-Caractéristiques	Deux Caractéristiques	Couleur et Contour [261], Couleur et Texture [125][620][42]
		Couleur et Disparité [202], Couleur et Profondeur [239]
	Trois Caractéristiques	Couleur, Contour et Texture [33], Couleur, Gradient et Caractéristiques de Haar [288][233]
Multi-Echelles	Pixel et Bloc	Pixel et Bloc [627][146][145][143][116][214]
Multi-Niveaux	Deux Niveaux	Pixel et Région [134][137][138]
		Pixel et Image [589]
	Trois Niveaux	Pixel, Région et Image [261][614][521][641]
Multi-Resolutions	Multi-Resolutions	Multi-Resolutions [342][638][540]
	Hierarchique	Hierarchical [111][639]
	Coarse-to-Fine	Coarse-to-Fine [61][60][192]
Multi-modèles		Short Term Background - Long Term Background
	Same Background Models	MOG-MOG [253][408][412]
	Different Background Models	Histogram (STLBP) - MOG [469]
		SG (RRC) - KDE [506][505][504]
		Predictive model - MOG [468][470]
Multi-Couches	Deux Couches	Deux Couches [582]
	Multiple Couches	Multiple Couches [410]
Multi-Classifieurs	Trois Classifieurs	Trois Classifieurs [426]
Multi-Objectifs	Two cues	Stratégies Bottom-up et Top-down [527]
		Local et Global [237]
	Three cues	Couleur, Contour et Intensité [256]
		Couleur, Texture et Region [385]
Multi-Critères	Trois Critères	Trois Critères [268]
Carte de Sélection	Deux Zones	Deux Zones [266]
Markov Random Field	Markov Random Field	Markov Random Field [296][449]
		Dynamic Hidden Markov Random Field [552][553][549]
		Hidden Conditional Random Fields [124]

TABLE 2.10 – Vue d'ensemble des différentes stratégies : Multi-caractéristiques, multi-résolutions, multi-modèles, multi-couches, multi-classifieurs, Multi-objectifs, Multi-critères, Champs de Markov.

Catégories	Sous-catégories	Méthodes
Implémentations Temps-Réel	Implémentation GPU	*Statistical :* SG [179], MOG [401][297],SVM [201][114][115]
		Clusters :, Codebook [179]
		Neural Network : SOBS [179]
	Implémentation Embarquée	Multimodal Mean [23]
		Light-Weight Foreground Detection [94][95]
		CS-MOG [463]
	Architecture Dédiée	Digital Signal Processor (DSP) [27][68]
		Very-large-Scale Integration (VLSI) [524][399]
		Field-Programmable Gate Array (FPGA) [247][561][550][24][264][316][196][436]
Capteurs	Cameras	Cameras Mono-chromatique [478], Color CCD Cameras
		Web Cameras [485][289][290]
		Cameras Haute Définition[195]
	Infra-Red Cameras	Infrared Cameras [375][133]
	RGB-D Cameras	Systémes Stéreoscopiques[202][239]
		Time-of-Flight (ToF) Cameras [479][303][302]
		Microsoft's Kinect [90][89]
	Light Field Cameras	Light Field Cameras [467][467]
Domaine des Données	Pixel	Domaine le plus courant
	Mesure	Compressive Sensing (CS) [556][97][310][547]
	Compression	DCT [545], H.264 Streaming Video [151], HEVC [630]
Evaluation	Vérité Terrain	*Courbes ROC :* True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN)[149]
		Courbes ROC : Detection Rate (DR), Precision, Recall [149]
		Courbes ROC : False Positive Rate (FPR), True Positive Rate (TPR) [149]
		F-Measure [353]
		D-Score [300]
		Similarity Measure [311]
	Sans Vérité Terrain	Perturbation Detection Rate [282]
		Boundary, Model and Assisted based Evaluation [368]
Bases de Données	Traditionnelles	*Articles :* Wallflower [521], I2R [311], Carnegie Mellon [461]
		Laboratoires : LIMU [599], USCD [102], SZTAKI [57], ETISEO [382]
		Conférences : VSSN 2006, OTCBVS 2006, PETS
		Projets : ATON [413], CAVIAR, i-LIDS
	Récentes	*Articles :* SABS [82]
		Workshops : ChangeDetection.net [204], Background Models Challenge (BMC) [530]
		Projets : MAR [67], RGB-D [89], CITIC RGB-D Dataset [182], Fish4knowledge [274], Aqu@theque [37]
Publications	Journaux	IEEE (PAMI, IP, etc..), ELSEVIER (PR, PRL, IVC, CVIU, etc..), SPRINGER (MVA, IJCV, etc..)
		Special Issue on "Background Modeling and Foreground Detection for Video Surveillance" dans MVA
		Special Issue on "Background Models Comparison" dans CVIU
	Conferences	ICIP, ICPR, CVPR, ACCV, ECCV
		AVSS
	Workshops	*Video Surveillance d'activités humaines :* ChangeDetection.net [204], Background Models Challenge (BMC) [530]
		Vidéo surveillance intelligente d'activités animales : VAIB, CVAUI

TABLE 2.11 – Vue d'ensemble des points clés : Implémentations temps-réel, capteurs, domaine des données, mesures pour l'évaluation des performances et les bases de données.

4 Les modèles de fond

Dans cette section, nous présentons les différents modèles classés en modèles conventionnels et récents comme expliqué dans la Section 3.1.

4.1 Les modèles actuels

Les modèles actuels sont les plus examinés dans la littérature et sont classés dans les différentes catégories suivantes (Survey Bouwmans (2014) [73]) :

— **Les modèles basiques :** Dans ce cas, le fond est modélisé soit par une moyenne [66], soit par une médiane [362], soit par un histogramme temporel [635]. Une fois l'image de fond calculée, les pixels de l'image courante sont classées en objets mobiles en seuillant la différence entre l'image de fond et l'image courante comme suit :

Si $d\left(I_t(x,y), B_{t-1}(x,y)\right) > T$ alors le pixel (x,y) est classé comme "objets mobiles" sinon le pixel (x,y) est classé comme "fond".

où T est un seuil constant et $I_t(x,y)$, $B_{t-1}(x,y)$ sont respectivement l'image courante au temps t et l'image de fond au temps $t-1$. $d(.,.)$ est une mesure de distance qui est usuellement la différence absolue.

— **Les modèles statistiques :** Les modèles statistiques sont plus robustes aux variations lumineuses et aux fonds dynamiques tout en offrant un bon compromis performance et temps de calcul [76][70]. Ces modèles comportent différentes catégories [76][70] :

1. Les modèles Gaussiens : Single Gaussian (SG) [562], Single General Gaussian (SGG) [279], Mixture of Gaussians (MOG) [490], Mixture of General Gaussians (MOGG) [16], Kernel Density Estimation (KDE) [167].
2. Les modèles à vecteurs de support : Support Vector Machine (SVM) [323], Support Vector Regression (SVR) [537], Support Vector Data Description (SVDD) [513].
3. Les modèles de réduction de dimension soit sous forme de matrices (PCA [387], ICA [579], INMF [83], LoPP [291]) soit sous forme de tenseurs (IRT [315][21], TenLoPP [292], DB [165]).

Les modèles Gaussiens sont plus adaptés aux fonds dynamiques alors que les modèles de réduction de dimension sont robustes en présence de variations lumineuses.

— **Les modèles à base de clusters :** Ces modèles supposent que chaque pixel dans l'image peut être représenté temporellement par un cluster. Ces modèles sont composés de modèles à base de K-means [84], de modèles à base de codebooks [280] et de

modèles à base de clustering séquentiel [567]. Ces modèles sont relativement robustes lors de l'apprentissage en cas de bootstrapping et pour les fonds dynamiques.

— **Les modèles à base de réseaux de neurones :** Dans ce cas, le fond est modélisé à partir des poids d'un réseau de neurones qui sont déterminés avec une séquence d'apprentissage sans objets mobiles présents. Le réseau de neurones apprend comment classer chaque pixel en fond ou objet mobile. Différents types de réseaux de neurones ont été employés comme les suivants : General Regression Neural Network [141][142], Multivalued Neural Network [346], Competitive Neural Network [343], Dipolar Competitive Neural Network [345], Self Organizing Neural Network (SOBS [350], SOBS-SC [354], 3dSOBS+ [355]), Growing Hierarchical Self Organizing Neural Network [391]. Les approches les plus représentatives sont SOBS [350] et sa variante SOBS-SC [354].

— **Les modèles à base de filtres :** Le fond est estimé à partir d'un filtre. Chaque pixel de l'image courante qui dévie de façon significative de la valeur estimée est classé comme "objet mobile". Les filtres utilisés dans la littérature sont le filtre de Wiener [521], le filtre de Kalman [271] et le filtre de Chebytchev [103]. Ces méthodes sont robustes lors de variations lumineuses lentes et nécessitent par contre un mécanisme au niveau image pour les variations lumineuses qui sont à la fois globales et rapides. Ce mécanisme est celui qui consiste à ré-initialiser l'image de fond dès que plus de 50% des pixels sont détectés comme "objets mobiles" dans l'image [521].

Le tableau 2.3 montre une vue globale de ces différentes méthodes actuelles. Les états de l'art que j'ai effectués sur certaines catégories et sous-catégories sont indiquées en gras.

4.2 Les modèles prospectifs

Les modèles prospectifs tentent d'être plus performants que les modèles actuels en addressant plusieurs challenges et sont quelques fois des améliorations de modèles actuels. On peut les classer dans les différentes catégories suivantes (Survey Bouwmans (2014) [72]) :

— **Les modèles statistiques avancés :** Ces modèles comportent les différentes catégories suivantes :

1. Les modèles à base de mélanges basés sur des distributions autres que des distributions Gaussiennes : Student's t [374], Dirichlet Mixture Model (DMM) [231], Variational Dirichlet Mixture Model (VarDMM) [173], Adaptive Poisson Mixture Model (APMM) [178].
2. Les modèles hybrides : KGMM [155] KGHM [337].
3. Les modèles non-paramétriques : ViBe [51], PBAS [248], VuMeter [203].

Les modèles non-paramétriques présentent l'avantage de nécessiter peu de réglages et sont parmis les méthodes "leaders" actuelles.

— **Les modèles flous :** Ces modèles essaient de prendre en compte l'imprécision et l'incertitude au niveau de la modélisation du fond en présence de fonds dynamiques et de variations lumineuses. Les modèles (T2-FMOG-UM [46], T2-FMOG-UV [46]) font partie des modèles développés lors de mes travaux avec Fida El Baf et sont donc détaillés dans le Chapitre 3.

— **Les modèles de réduction de dimension :** Ces modèles peuvent être classées de la façon suivante :

1. **Discriminatifs ou mixtes :** Ces modèles (IMMC [177], PCA-ILDA [359]) sont les modèles développés lors de mes travaux avec Diana Farcas et Cristina Marghes et sont donc détaillés dans le Chapitre 4.

2. **Robustes basés sur la décomposition en matrices :** Ces modèles se déclinent en trois catégories. La première est basée soit sur des approches classiques d'ACP robuste [517][518] soit sur l'approche d'ACP robustes développée par Candes et al. [92] en décomposition en matrices de rang faible et parcimonieuse. Ces modèles ont été évalués dans le cadre de la thèse de Charles Guyon et sont donc examinés dans le Chapitre 5 ainsi que les variantes proposées lors de mes travaux avec Charles Guyon (IRLS [226], Polynomial RPCA [228]). La deuxième catégorie est basée sur des approches NMF (MahNMF [212], NS-NMF [295]). La troisième est quant à elle basée sur une approche "Robust Orthonomal Subspace Learning" (ROSL)[474].

3. **Robustes basés sur la décomposition en tenseurs :** Ces modéles se déclinent en deux catégories et utilisent des tenseurs pour prendre en compte l'aspect spatio-temporel. La première catégorie est l'extension de la décomposition de la décomposition RPCA de type Candes et al. [92] sous forme de tenseur (LTR-MALM [501][502], MALM-TR [503], T-RPCA [522], SO-IR [312], RSTD [318]). La seconde catégorie est quant à elle l'extension de la décomposition Robust NMF sous forme tenseur (RTF [128]).

— **Les modèles basés sur le suivi des sous-espaces :** Ces modèles ont pour objectif de suivre la structure des sous-espaces au cours du temps (GRASTA [240], t-GRASTA [242], pROST [230][451], GOSUS [571]). Ces méthodes sont souvent incrémentales et donc très adaptées pour la mise à jour du fond.

— **Les modèles basés sur la minimisation en rang faible :** Ces modèles sont basés sur une représentation de rang faible dans laquelle la matrice de rang faible et le support des outliers sont estimés (DECOLOR [646], DRMF [569], DRMF-R [569], PRMF [542], BRMF [543]). Ces approches effectuent la modélisation du fond et la détection des objets mobiles en un seul processus en résolvant la décomposition à l'aide un algorithme alternatif. Contrairement aux approches RPCA de type Candes et al. [92], ces approches n'émettent aucune hypothèse sur la parcimonie de la matrice correspondant aux objets mobiles.

— **Les modèles basés sur les approches parcimonieuses :** Ces modèles se basent sur l'aspect parcimonieux des objets mobiles et peuvent être classés comme suit :

1. Compressive Sensing [97].
2. Structured Sparsity [255].
3. Dynamic Group Sparsity [254].
4. Dictionary Learning [147].
5. Sparse Error Estimation [153].

Ces modèles sont en plein développement notamment les modèles basé sur le Compressive Sensing [97]. Ces modèles nécessitent des solveurs (Basis Pursuit (BP) [97][110][91], Orthogonal Matching Pursuit (OMP) [148], Lattice Matching Pursuit (LaMP) [96], Compressive Sampling Matching Pursuit (CoSaMP) [381], K-CoSaMP [573], CoSaMP Subspace [241], Gradient Projection [547][187]), et une comparaison de leur influence a été effectuée par Davies et al. [148]. De plus, Cevher et al. [555][554] ont développé des approches adaptatives.

— **Les modèles de transformation de domaine :** Ces modèles changent de domaine pour prendre en compte les caractéristiques spatiales et temporelles et sont généralement plus robustes en présence de fonds dynamiques. Ces modèles utilisent différentes transformées :

1. La transformée de Fourier [563].
2. La transformée en cosinus discret [411].
3. La transformée de Walsh [516].
4. La transformée en ondelette [193].
5. La transformée d'Hadamard [49].

La pertinence du choix de telle ou telle transformée n'est pas abordée dans la littérature alors qu'il serait bien d'en mesurer l'influence vis-à-vis des challenges.

Les tableaux 2.4 et 2.5 montrent une vue globale de ces différentes méthodes prospectives. Les approches développées sous ma direction scientifique sont indiquées en gras. Les états de l'art que j'ai effectués sur certaines catégories et sous-catégories sont indiqués en gras.

4.3 Les modèles "Autres"

D'autres modèles font leur apparition ou n'ont pas fait l'objet pour l'instant de recherches plus avancées bien qu'ayant un potentiel intéressant. On peut citer en autres les modèles suivants :

— **Modèle basé sur la fréquence d'apparition :** Hou and Han [250] ont développé un modèle de reconstruction de fond basé sur l'hypothèse que le fond est visible la plupart du temps. Ce modèle appelé "Pixel Intensity Classification" (PIC) calcule la différence entre les images consécutives, et classe ensuite les pixels en fonction de cette différence. Puis, l'intensité ayant la plus grande occurrence est sélectionnée comme valeur représentant le fond. Ce modèle a fait l'objet de différentes améliorations (Seuil adaptatif [629], post-traitement [601]). Fu et al. [188] ont proposé une variante dans le cadre du modèle SG. Cette méthode est appelée "Pixel Change Classification" (PCC). Ces modèles ne semblent être adaptés qu'à des fonds uni-modaux.

— **Modèle basé sur la matrice de co-occurence :** Ce modèle est basé sur l'utilisation de la matrice de co-occurrence. Différentes approches ont été développées [319][637].

— **Modèle basé sur des méthodes multi-kernels :** Molina-Giraldo et al. [370] ont développé une méthode multi-kernels appelée "Weight Gaussian Kernel Video Segmentation" (WGKVS) qui attribue un noyau Gaussien différent à chacune des caractéristiques employées. Le modèle ensuite pondère l'influence de chaque caractéristique permettant ainsi d'être le plus adaptatif possible. Ce modèle donne de très bons résultats en présence de fonds dynamiques et de variations lumineuses.

— **Modèle basé sur la réduction de dimension de blocs locaux spatio-temporels :** Lin et al. [324] ont proposé une méthode de modélisation pour les fonds dynamiques présentant des textures. Cette méthode utilise des caractéristiques de LBP sous forme de blocs locaux spatio-temporels. Une méthode de réduction de type PCA est ensuite employée. Ce modèle donne de très bons résultats en présence de fonds dynamiques.

— **Modèle basé sur la poursuite de densité :** Pham et al. [400] ont proposé une méthode non paramétrique basée sur la "Projection Pursuit Density Estimation" (PPDE). Ce modèle est similaire dans la conception au modèle des KDE sauf que la fonction de densité n'est pas Gaussienne mais un produit de densité. Ce modèle semble robuste en présence de variations lumineuses dans des scènes d'intérieur et d'extérieur.

— **Modèle basé sur l'estimation de densité :** Ces modèles utilisent soit des approches de type "Mean Shift" [403][320][236], soit des approches de filtrage particulaire (Particle Filter) [618] soit des approches de type "Radial Reach Filter" [448][445][443].

Ces méthodes semblent robustes dans le cas de variations lumineuses progressives.

— **Modèle basé sur la détection d'outliers :** Pejcic et al. [397] ont proposé l'utilisation de l'algorithme LOF (Local Outlier Factor) qui attribue un degré d'appartenance à la classe "outlier" pour chaque donnée et l'algorithme COF (Connectivity-based Outlier Factor) qui prend en plus en compte les informations de voisinage des pixels. Pour chacun de ces algorithmes, une version incrémentale a été développée et appelée respectivement incLOF [397] et inCOF [397][428]. Cette méthode donne de bons résultats dans le cas de petits objets dans des séquences filmées dans le domaine infrarouge.

Le tableau 2.6 montre une vue globale de ces différentes méthodes "Autres".

5 Les stratégies

Plusieurs stratégies peuvent être mises en oeuvre pour améliorer la robustesse d'un modèle aux différents challenges. Il est important de noter que ces différentes stratégies peuvent être appliquées la plupart du temps à n'importe quel modèle car elles sont indépendants généralement du modèle de fond choisi. On peut classer ces différentes stratégies comme suit :

— **Multi-caractéristiques :** L'objectif est d'autres caractéristiques en plus de la couleur pour pallier les limites de celle-ci. Par exemple, comme les gradients d'une image sont moins sensibles aux changements lumineux, Javed et al. [261] les utilise en plus de la couleur pour obtenir une méthode relativement robuste aux variations lumineuses. En outre, la plupart des approches multi-caractéristiques utilisent la texture pour améliorer la performance en présence de variations lumineuses et d'ombres comme dans [125][126][620][42]. Pour résoudre la situation de camouflage, plusieurs auteurs [202]
[239] ont employés les caractéristiques stéréoscopiques comme la disparité et la profondeur. D'une autre façon, Azab et al. [33] ont agrégé 3 caractéristiques (couleur, contour et texture) en utilisant l'intégrale de Choquet pour être robuste aux variations lumineuses et aux fond multi-modaux **poursuivant ainsi l'approche initiée par Fida El Baf sous ma direction**. Klare and Sakar [288] ont quant-à-eux proposé un algorithme qui incorpore plusieurs instantiations de l'algorithme MOG avec les caractéristiques couleur, contour et texture (Haar features). L'agrégation est effectuée en utilisant l'opérateur de moyennage. Plus récemment, Han and Davis [233] ont appliqué les SVM sur les vecteurs de vraisemblance pour un ensemble de caractéristiques constitué de la couleur, du gradient et de la texture. Le problème de combinaison et de sélection de caractéristiques reste encore un problème ouvert en détection d'objet mobiles.

— **Multi-échelles :** Cette stratégie consiste à travailler au niveau pixel et au niveau bloc pour combiner leur avantages respectifs (voir Section 3.6). Les caractéristiques utilisées au niveau pixel et au niveau bloc peuvent être de nature différente, comme par exemple la couleur au niveau pixel et la texture au niveau bloc [627]. La stratégie multi-échelle a été employée avec succès pour l'étape de l'initialisation du fond [146][145], la détection d'objets stationnaires [116] et les fonds dynamiques [214]. De plus amples détails sur la stratégie multi-échelle sont disponibles dans [143].

— **Multi-niveaux :** L'idée est de résoudre les différentes situations critiques à différents niveaux qui sont soit au niveau pixel, soit au niveau de la région soit au niveau de l'image [261][614]. Le niveau région permet de prendre en compte l'aspect connexion spatial des pixels constituant les objets mobiles. Par ailleurs, le niveau image permet quant-à-lui de gérer les variations lumineuses globales. En général, l'hypothèse émise est que l'image de fond n'est plus adaptée si plus de la moitié des pixels sont détectés. Dans ce contexte, des auteurs ont utilisés soit deux niveaux (pixel et région, pixel et image) soit trois niveaux (pixel, région et image).

— **Multi-résolution :** L'objectif est de détecter les zones où sont présents les objets mobiles au niveau le plus haut et d'affiner la détection au niveau pixel [342][638][540]. Les zones classées comme fond à un niveau ne sont pas examinées au niveau inférieure. Cette stratégie permet de diminuer le temps de calcul. Elle est aussi appelée hiérarchique [111][639] ou "coarse-to-fine" [61][60][192].

— **Multi-fonds :** Cette stratégie utilise en général deux images de fond s'adaptant respectivement aux changements progressifs et rapides. Le premier modèle est un modèle "court-terme" dont le but est d'intégrer les changements sur un temps court. Les pixels appartenant au fond sont utilisés pour la mise à jour de l'image de fond "court-terme". Ceci n'est pas souhaitable quand les objets mobiles ralentissent et peuvent être intégrés dans le fond. Le seconde modèle est un modèle "long-terme" dont l'objectif est quant-à-lui d'intégrer les changements sur un temps long. Chaque pixel est employé pour la mise à jour du modèle "long-terme". Ceci n'est pas souhaitable quand un objet mobile se déplace lentement ou quand certains objets sont insérés dans le fond, mais cela compense les inconvénients du modèle "court-terme". Ensuite, les deux modèles sont combinés pour obtenir une meilleure image de fond. Un bon exemple dans [408][412] permet de détecter les objets temporairement immobiles en vidéo-surveillance. Tous ces travaux emploient le même modèle de représentation du fond pour le modèle "court-terme" et le modèle "long-terme" mais cela n'est pas forcément optimal. Par exemple, Shimada et Taniguchi [469] ont pris respectivement un modèle histogramme de texture (STLBP) pour le modèle "court-terme", et le modèle MOG pour le modèle "long-terme". Tanaka et al. [506][505][504] ont quant-à-eux

employé un modèle SG sur les textures RRC pour le modèle "court-terme" et le modèle KDE pour le modèle "long-terme". Dans la même idée, Shimada et Taniguchi [468][470] ont préféré un modèle prédictif pour le modèle "court-terme" et le modèle MOG pour le modèle "long-terme". Cette prise de modèles différents permet de prendre en compte les spécificités des changements à chaque échelle de temps.

— **Multi-couches :** Cette stratégie permet de gérer les fonds multi-modaux. Par exemple, Yang et al. [582] ont utilisé un mélange de deux Gaussiennes en considérant deux couches. La première couche est dédiée aux changements progressifs alors que la seconde est consacrée aux changements irréguliers. Cette méthode est effectivement plus robuste et précise dans le cas de scènes présentant des fonds multi-modaux.

— **Multi-classifieurs :** Cette stratégie consiste à utiliser plusieurs classifieurs à l'étape de classification fond/objets mobiles et ensuite de combiner ceux-ci ou de les faire suivre en cascade pour profiter de leurs avantages complémentaires par rapport aux différents challenges comme dans la méthode développée par Reddy et al. [426].

— **Multi-objectifs :** L'idée est de gérer les différents challenges avec des solutions dédiées qui sont ensuite agencer pour que l'ensemble soit robuste à la plupart des challenges comme dans la méthode développée par Huerta et al. [256]. D'autres exemples sont disponibles dans [385][527][237].

— **Carte de sélection :** Le principe est d'appliquer un modèle différents suivant l'activité d'une zone de l'image [266]. Par exemple, Liang et al. [320] ont définies deux zones : 1) les zones uni-modales dans laquelle le modèle SG [562] est appliquée et 2) les zones multi-modales dans lesquelles le modèle KDE [167] est employée. Ainsi, le temps de calcul est optimisé car il évite d'employer le modèle KDE sur toutes les zones de l'image.

— **Champs de Markov :** Les champs de Markov permettent de prendre en compte les contraintes spatiales et temporelles. Par exemple, Kumar et Sengupta [296] ont ainsi amélioré les résultats du modèle MOG réduisant les fausses détections de façon significative.

Le tableau 2.10 montre une vue synthétique de ces différentes stratégies. Une idée pas encore exploitée dans le domaine serait de combiner au mieux ces différentes stratégies pour profiter de chacun de leurs avantages.

6 Les implémentations temps réel

La détection d'objets mobiles s'effectue la plupart du temps dans des applications qui nécessitent de fonctionner en temps réel et dans des systèmes embarqués. Ainsi, le temps de calcul et l'espace mémoire utilisés par les algorithmes doivent être réduits le plus possible. Les différentes approches temps réel peuvent être classées comme suit.

6.1 Implémentations GPU

Une façon d'atteindre des performances temps réel sur PC classique est d'utiliser la carte GPU. Par exemple, Fauske et al. [179] ont proposé une implémentation temps réel utilisant le langage CUDA pour les algorithmes suivants : SG [562], codebook [281] et SOBS [350]. La plupart des autres recherches sur les implémentations temps réel concernent les MOG [401][297]. D'autres concernent le modèle SVM [201][114][115].

6.2 Implémentations embarquées

Les systèmes embarqués avec des contraintes de temps réel requierent d'optimiser les algorithmes de suppression de fond. Par exemple, Apewokin et al. [23] ont développé dans ce cadre une méthode appelée multi-modal mean (MM) qui développe les mêmes performances en terme de précision de détection que les MOG [490] avec un facteur de gain en exécution de 6 et d'occupation mémoire de 17. Pour les caméras embarquées dites intelligentes, Casares et al. [94][95] ont développé une méthode de détection d'objets mobiles robustes aux variations lumineuses et aux fonds dynamiques.

6.3 Implémentations sur architecture dédiée

Une manière efficace d'atteindre des performances temps-réel est de développer une architecture dédiée qui peut être basée soit sur la technologie DSP [27][68], soit sur la technologie VLSI [524][399] soit sur la technologie FPGA [247][561][550][24][264][316][196][436].

Le tableau 2.11 montre une vue synthétique de ces différentes implémentations temps réel. Un des challenges actuel est de concevoir des implémentations temps réel pour les méthodes récentes comme celles basées sur la décomposition en matrices de rang faible et parcimonieuse.

7 Les capteurs

Les capteurs sont essentiels dans l'acquisition de la séquence vidéo car ils déterminent la disponibilité de tel ou tel caractéristique comme celle de profondeur ou celle relative à l'infra-rouge. Les différents capteurs utilisables sont les suivants :

— **Les caméras standards :** Les caméras dans le domaine visible sont naturellement les plus utilisées et peuvent faire l'acquisition de séquence en niveau de gris [478] ou couleur. Les webcams peuvent être aussi utilisées à moindre coût avec une qualité d'image plus faible [485][289][290]. En outre, les caméras HD commencent à être employées dans la mesure où leur prix diminue [195]. Dans chaque cas, des algorithmes spécifiques ont été développés pour tenir en compte la particularité du capteur.

— **Les caméras infra-rouge :** Elles sont soit les seules disponibles dans certaines applications, soit elles sont utilisées en complément des caméras du domaine visible. Dans ce cas, l'idée est de s'affranchir des variations lumineuses et aussi de la situation de camouflage comme dans [375][133].

— **Les caméras RGB-D :** Les systèmes stéréoscopiques fournissent tant l'information de couleur que de profondeur [202][239]. Cependant, ils nécessitent deux caméras. Par contre, les caméras ToF (Time of Flight) apporte à elle-seules les deux informations avec une fréquence d'acquisition relativement importante. Cependant, la qualité des images est moindre que dans le cas des systèmes stéréoscopiques. Par contre, ces systèmes permettent d'acquérir la profondeur qui permet de pallier la situation de camouflage en couleur [479][303][302]. Récemment, les caméras RGB-D low-cost comme la Microsoft Kinect[1] ou la Asus Xtion Pro[2] permettent d'obtenir facilement l'information de profondeur. Dans ce contexte, Camplani et al. [90][89] ont employé une Microsoft Kinect pour la détection d'objets mobiles avec un mélange de Gaussiennes utilisant les caractéristiques RGB et la caractéristique de profondeur. Une méthode similaire a été développée par Fernandez-Sanchez et al. [182][183].

Le tableau 2.11 montre une vue synthétique des différents capteurs utilisés dans la suppression de fond lorsque la caméra est fixe. Dans le cadre d'applications où les caméras sont mobiles, on peut utiliser des caméras PTZ (Pan-Tilt-Zoom) [184][215][34] ou des caméras des systèmes mobiles (smartphones) [460] ou des caméras embarquées (voitures, robots) [460]. Le challenge actuel est de concevoir un système multi-capteurs pour pouvoir sélectionner le capteur le mieux adapté suivant la ou les conditions détectées ou de fusionner les informations venant de ceux-ci.

1. http://www.microsoft.com/en-us/kinectforwindows/
2. http://www.asus.com/Multimedia/

8 Evaluation des performances

L'évaluation des performances est importante pour comparer les différentes méthodes soit en terme de temps de calcul et d'utilisation de mémoire soit en terme de fausses détections et de détections ratées. Pour évaluer les performances dans le premier sens, le temps et la mémoire requise par l'algorithme peuvent être facilement mesurés par l'utilisation d'instructions dans le code. Pour mesurer les performances dans le second sens, différentes méthodes sont possibles avec ou sans vérité terrain.

8.1 Evaluation avec vérité terrain

Ces méthodes d'évaluation nécessitent une vérité-terrain pour calculer les bonnes et fausses détections notées TP (True Positive), TN (True Negative), FP (False Positive) et FN (False Negative). FP et FN concernent les pixels classés incorrectement en objet mobile (FP) ou en fond (FN) alors que TN et TP concernent les pixels correctement classés en objet mobile ou en fond. Ensuite différentes métriques sont calculées et exploitées pour avoir une mesure de la robustesse des méthodes de suppression de fond. Les méthodes les plus courantes dans la littérature sont les suivantes : L'analyse des courbes ROC [149], la F-Measure [353] et la mesure de similarité de Li [311] :

— **Analyse des courbes ROC :** Les courbes ROC (Receiver Operation Characteristics) représentent l'évolution d'une métrique en fonction d'une autre. Celles-ci sont le *Detection Rate* (DR) et la précision calculée comme suit :

$$DR = \frac{TP}{TP + FN} \tag{2.5}$$

$$Précision = \frac{TP}{TP + FP} \tag{2.6}$$

Les courbes ROC utilisent aussi le *False Positive Rate* (FPR) et le *True Positive Rate* (TPR) qui sont calculés comme suit :

$$FPR = \frac{FP}{FP + TN} \quad ; \quad TPR = \frac{TP}{TP + FN}$$

Cette méthode est certainement la plus utilisée mais présente plusieurs limitations. En effet, elle nécessite une vérité terrain qui requiert une segmentation manuelle. De plus, elle mesure les taux d'erreurs de détection d'un objet particulier sur un fond particulier. Ainsi, il y a autant de courbes ROC qu'il y a d'objets d'intérêt différents. De plus, pour une scène donnée, il n'y a aucune connaissance sur ce qui peut correspondre à un objet d'intérêt dans la scène.

— **F-Measure :** La F-Measure caractérise la performance de la classification dans l'espace Précision et Detection Rate et est calculée comme suit :

$$F = \frac{2 * DR * Precision}{DR + Precision} \tag{2.7}$$

L'objectif est de maximiser F à 1 [353].

— **Mesure de similarité de Li :** Li et al. [311] ont introduit cette mesure pour évaluer leur méthode de suppression de fond. Soit A une région détectée et B la vérité terrain correspondante, la mesure de similarité de Li entre A et B est définie comme suit :

$$S(A, B) = \frac{A \cap B}{A \cup B} \tag{2.8}$$

Si A et B sont similaires, $S(A, B)$ est proche de 1, sinon elle se rapproche de 0. Comme pour les courbe ROC, il est nécessaire d'avoir une vérité terrain et donc une segmentation manuelle. Cette mesure s'appelle également indice de Jacquard ou coefficient de Jacquard qui représente donc le rapport entre le cardinal (la taille) de l'intersection des ensembles considérés et le cardinal de l'union des ensembles. Cette mesure permet ainsi d'évaluer la similarité entre les ensembles et par conséquent d'évaluer les méthodes en terme de qualité de segmentation.

D'autres mesures plus rarement utilisées ont été proposées dans [380][379][378][304][330]. Toutes ces méthodes évaluent les modèles de suppression de fond au niveau pixel. Cependant, il peut être intéressant d'introduire une mesure qui dépend du contexte de l'application.

Par exemple, Lallier et al. [300] ont proposé un cadre d'évaluation multi-niveaux (pixel, image et séquence) pour l'application de la détection d'intrusion. Le niveau "pixel" fournit une mesure de précision quant à la classification du pixel en "fond" ou "objet mobile". Le niveau "image" donne une mesure sur la décision d'intrusion ou non. Le niveau "séquence" mesure la précision sur la période de temps où l'objet apparaît. De plus, Lallier et al. [300] ont proposé une nouvelle mesure appelée D-Score qui est plus adaptée à la détection d'intrusion.

Cependant, toutes ces méthodes nécessitent la vérité-terrain et bien que quelques méthodes ont été développées pour la générer de façon interactive [209][208], cette tâche est souvent fastidieuse.

8.2 Evaluation sans vérité terrain

L'idée de ces méthodes d'évaluation est de s'affranchir de la vérité terrain souvent longue à générer. Les plus connues sont les suivantes :

— **Analyse par perturbation ("Perturbation Detection Rate") :** Kim et al. [282] ont développé l'analyse de perturbation PDR pour évaluer les méthodes de suppression de fond. Cette technique définit la sensibilité d'un algorithme à détecter un objet d'intérêt présentant peu de contraste avec une partie du fond comme une fonction de contraste sans connaissance à priori sur la distribution de l'objet d'intérêt. L'idée est de mesurer la distance nécessaire entre deux distributions dans le but d'obtenir un certain coefficient de détection. En d'autres termes, l'approche PDR consiste à déterminer, pour un coefficient de fausses détections, un coefficient de détection comme une fonction de différence entre l'objet d'intérêt et le fond. Dans l'analyse ROC, le principe est basé sur le fait que les distributions du fond et de l'objet d'intérêt sont connues et que les paramètres de l'algorithme doivent varier pour obtenir un certain coefficient de fausses détections et de détections ratées. L'hypothèse avancée par Kim et al. [282] considère que la forme de la distribution de l'objet d'intérêt est localement similaire à celle de la distribution du fond et que la distribution de l'objet d'intérêt de peu de contraste peut être considérée comme une version perturbée de la distribution du fond. Comme le remarquent Kim et al. [282], cette hypothèse est raisonnable parce que tout objet peut être fond ou objet d'intérêt. Par exemple, pour une application en vidéo-surveillanvce une voiture peut être considérée comme une partie du fond dès qu'elle s'arrête dans un parking et comme un objet d'intérêt quand elle se déplace. En variant les paramètres des algorithmes, l'analyse PDR détermine une relation entre les coefficients de fausses alarmes et celui de détections, et la distance entre les distributions, contrairement à l'analyse de ROC, qui détermine une paire de coefficients d'erreurs. Une limitation de l'analyse PDR est qu'elle ne prend pas en compte le flou dû au mouvement d'un objet d'intérêt. Dans le cas d'un mouvement dans le fond, la distribution des objets d'intérêt sera mélangée. Les objets du fond ont souvent des effets de réflexion de lumière sur le fond. Ils ont un impact important dans la détermination du coefficient adéquat de fausses détections.

— **Analyse par multi-variations :** San Miguel et Martínez [368] ont proposé une méthode qui évalue les modèles de suppression de fond sur des mesures extraites de variations de contours, de l'impact des variations de la détection sur l'étape suivante et sur une segmentation obtenue par d'autres méthodes de suppression de fond.

Toutes ces mesures d'évaluation permettent de comparer de façon objective la grande diversité des algorithmes de suppression de fond sur des bases de données communes examinées dans la section suivante.

9 Les bases de données

Les bases de données de séquence vidéos acquises par des caméras fixes sont importantes pour faire une comparaison complète, fiable et juste des différents algorithmes. Le tableau 2.11 montre une vue synthétique des différentes bases de données utilisées dans la suppression de fond. De plus, celles-ci sont directement accessibles dans la section "Available Datasets" de notre site "Background Subtraction Web Site". Nous avons classés les différentes bases de données existantes en traditionnelles et récentes. Les premières permettent des évaluations partielles car elles ne couvrent pas tous les challenges et ne fournissent pas forcément le nombre suffisant d'images vérité terrain alors que les secondes plus récentes sont plus complétes. Le tableau 2.12 montre une vue d'ensemble des différentes bases de données. La première colonne indique l'étape concernée. La seconde colonne indique le concept flou employé avec l'acronyme de la méthode et le nombre de publications entre parenthèses. La troisième colonne indique le nom des auteurs et la date de la première publication. En gras, les approches développées par Fida El Baf sous ma directions scientifique.

Bases de données	Noms	Domaine	Auteurs - Dates
Traditionnelles	Wallflower	Visible (RGB)	Toyama et al. (1999) [521]
	I2R	Visible (RGB)	Li et al. (2004) [311]
	Canergie Mellon	Visible (RGB)	Sheikh and Shah (2005) [461]
	LIMU	Visible (RGB)	Yoshinaga et al. (2010) [599]
	USCD	Visible (RGB)	Chan et al. (2011) [102]
	VSSN 2006	Visible (RGB)	(2006)
	OTCBVS Benchmark	Infra-rouge	(2009)
	PETS	Visible (RGB)	2000, 2001, 2002, 2003, 2006, 2014
	ATON	Visible (RGB)	Prati et al. (2003) [413]
	CAVIAR	Visible (RGB)	(2003)
	ETISEO	Visible (RGB)	Nghiem et al. (2007) [382]
Récentes	ChangeDetection.net (CDW)	Visible (RGB), Infra-rouge	Goyette et al. (2012)
	Background Models Challenge (BMC)	Visible (RGB)	Vacavant et al. (2012) [477]
	Stuttgart Artificial Background Subtraction (SABS)	Visible (RGB)	Brutzer et al. (2011) [82]
	MAR	Visible (RGB)	Bloisi et al. (2014) [67]
	RGB-D Object Detection	Visible (RGB-D)	Camplani et al.(2013) [89]
	CITIC RGB-D	Visible (RGB-D)	Fernandez et al. (2013) [182]
	Fish4knowledge	Visible (RGB)	Kavadis et al. (2013) [274]

TABLE 2.12 – Vue d'ensemble des différentes bases de données.

Figure 9.1 – Wallflower dataset [521] : La première ligne montre l'image test de chaque séquence et la seconde montre la vérité-terrain correspondante suivant l'ordre suivant : Moved Object, Time of Day, Light Switch, Waving Trees, Camouflage, Bootstrapping and Foreground Aperture.

9.1 Les bases de données traditionnelles

Les bases de données traditionnelles sont les suivantes :

— **Wallflower dataset :** La base de données Wallflower[1] est fournie par Toyama et al. [521] et consiste en sept vidéos qui comportent chacune un challenge qu'on peut rencontrer dans des scènes réelles de vidéo-surveillance telles que les variations lumineuses et les fonds dynamiques. Pour chaque séquence, une vérité-terrain est fournie pour une image clef de la séquence où le changement intervient. La Figure 2.11 montre une vue de cette base de données. La première ligne montre l'image test de chaque séquence et la seconde montre la vérité-terrain correspondante. Cette base de données est la plus utilisée à ce jour car elle a été la première dans le domaine mais comme elle ne fournit qu'une vérité-terrain par séquence, son utilisation tend à disparaître au profit des récentes bases de données.

— **I2R dataset :** La base de données I2R[2] est fournie par Li et al. [311] et comporte 9 séquences comportant des fond dynamiques, des variations lumineuses et des fonds peu visibles. Les séquences "Curtain", "Campus", "Water Surface", et "Fountain" présente des fonds dynamiques alors que la séquence "Lobby" comporte des variations lumineuses et les séquences "Shopping Mall", "Airport" et "Restaurant" présentent des fonds peu visibles. Pour chaque séquence, la vérité-terrain est fournie pour vingt images clef où les algorithmes doivent montrer leur robustesse aux changements. La Figure 2.12 montre une vue de cette base de données. La première ligne montre l'image test de chaque séquence et la seconde montre la vérité-terrain correspondante. Cette base de données est fréquemment employée lors de validation des méthodes sur des fonds dynamiques. En particulier, elle est couramment utilisée dans l'évaluation qualitative des méthodes RPCA basées sur la décomposition en matrices de rang faible et parcimonieuse comme par exemple dans la publication de Candès et al. [92].

1. http://research.microsoft.com/en-us/um/people/jckrumm/wallflower/testimages.htm
2. http://perception.i2r.a-star.edu.sg/bk-model/bk-index.html

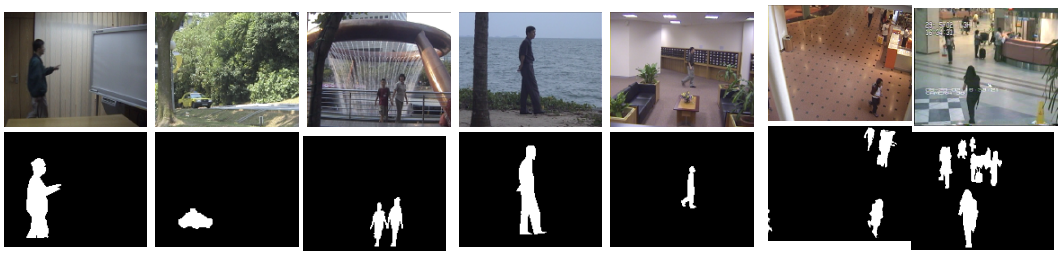


FIGURE 2.12 – I2R dataset [311] : La première ligne montre une image originale de chaque séquence et la seconde montre la vérité-terrain correspondante suivant l'ordre suivant : Curtain, Campus, Fountain, Water Surface, Lobby, Shopping Mall, and Airport

— **Carnegie Mellon Dataset (CMD) :** La séquence CMD[3] fournie par [461] emploie une caméra basée sur un poteau qui vacille à cause de rafale de vent. De plus, des variations lumineuses apparaissent au cours de la séquence. La vérité-terrain est fournie pour toute la séquence, permettant ainsi une évaluation fiable.

— **LIMU Dataset :** Le laboratoire LIMU[4] [599][600] met à disposition cinq séquences avec les vérité-terrain toutes les quinze images. De plus, trois séquences de PETS 2001 sont fournies avec leur vérité-terrain.

— **UCSD Dataset :** Cette base de données[5] [102] comporte 18 séquences vidéos fournies par le Statistical Visual Computing Lab (SVCL). La vérité-terrain est fournie pour quasiment toutes les séquences.

— **VSSN 2006 Dataset :** Cette base de données comporte différentes catégories de vidéos classées comme suit : 1) fonds dynamiques, variations lumineuse progressives et ombres, 2) variations lumineuses rapides, 3) fonds peu visibles et 4) deux ou trois vues de la même scène avec plusieurs personnes à détecter. Dans chaque vidéo, les objets mobiles sont des objets obtenus par synthèse des images et placés sur des fonds réels. Ainsi, les images vérité-terrain sont très précises. De plus, un environnement de développement en C++ était fourni pour permettre une comparaison équitable. VSSN 2006 est le premier workshop dédié à l'évaluation et la comparaison des méthodes de suppression de fond.

— **OTCBVS Benchmark Datasets :** Cette base de données est celle du workshop "Object Tracking and Classification in and Beyond the Visible Spectrum" (OTCBVS[6]) est contient des vidéos infra-rouge. Trois séquences sont intéressantes pour les méthodes

3. http://www.cs.cmu.edu/ yaser/new-backgroundsubtraction.htm
4. http://limu.ait.kyushu-u.ac.jp/dataset/en/
5. http://www.svcl.ucsd.edu/projects/background-subtraction/
6. http://www.vcipl.okstate.edu/otcbvs/bench/

de suppression de fond : (1) Dataset 01 (OSU Thermal Pedestrian) qui concerne la détéction de personnes, (2) Dataset 03 (OSU Color-Thermal Database) et (3) Dataset 05 (Terravic Motion IR Database).

— **PETS Datasets :** Cette base de données provient de la conférence "Performance Evaluation of Tracking and Surveillance" (PETS) et comporte différente base de données comme celles de PETS 2001, PETS 2003 et PETS 2006. Cette base de donnée est plus adaptée à lévalaution des méthodes de suivi qu'à celle des méthodes de suppression de fond car la vérité-terrain est fournie sous forme de boite englobante dans laquelle doit se trouver l'objet mobile.

— **ATON Dataset :** La base de données ATON[7] [413] comporte cinq séquences (Highway I, Highway II, Campus, Laboratory, Intelligent Room) qui présentent des ombres dans diverses scènes d'intérieur et d'extérieur. La vérité-terrain est fournie pour chaque séquence.

— **CAVIAR Dataset :** La base de données CAVIAR[8] contient des vidéos d'intérieur comportant différents types de comportements humains. Comme dans la base de données PETS, la vérité-terrain est fournie sous forme de boite englobante dans laquelle doit se trouver l'objet mobile. Cette base de données est plus pour la validation de détection de scénario que des méthodes de suppression de fond.

— **i-LIDS Dataset :** La base de données i-LIDS[9] (Information about the Imagery Library for Intelligent Detection Systems) comporte 4 vidéos de situations de vidéo-surveillance comme des voitures dans un parking et des objets abandonnés. Son application concerne plus la détection d'événements que la détection d'objets mobiles.

— **ETISEO Dataset :** La base de données ETISEO[10] [382] comporte plus de 70 vidéos de scènes d'intérieur et d'extérieur sur de longues périodes de temps. La vérité terrain fournie pour toutes les séquences contient également des informations pour des étapes de plus haut niveau et permet ainsi de voir l'influence de l'étape de détection sur les étapes suivantes. L'information fournie pour la détection des objets est le nombre d'objets détectés. Cependant, il n'est pas possible de faire des mesures de précision au niveau pixel pour différencier plus précisément les algorithmes [204].

7. http://cvrr.ucsd.edu/aton/shadow/index.html
8. http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/
9. https://www.gov.uk/imagery-library-for-intelligent-detection-systems
10. http://www-sop.inria.fr/orion/ETISEO/

FIGURE 2.13 – ChangeDetection.net dataset [204] : La première ligne montre une image représentative d'une séquence de chaque catégorie et la seconde montre la vérité-terrain correspondante dans l'ordre suivant : Baseline, Dynamic Background, Camera Jitter, Shadows and Intermittent Object Motion.

En résumé, les bases de données Wallflower et I2R comportent diverses scènes représentatives des challenges rencontrées en vidéo-surveillance. La base de données CMU est dédiée à un challenge qui est le mouvement de caméra. La base de données OTCBVS concerne le domaine de l'infra-rouge. La base de données ATON est dédiée aux ombres. Enfin, les bases de données CAVIAR, i-LIDS et ETISEO sont plus adaptées à la validation des étapes de plus haut niveau. Finalement, aucune de ces bases de données offrent à la fois un spectre complet des challenges et des images de verité-terrain suffisantes pour une évaluation fiable.

9.2 Les bases de données récentes

Les bases de données récentes mettent à disposition un spectre large de vidéos avec des images vérité-terrain pour chaque image de chaque séquence. Ces bases de données sont libres d'accès et fournissent de plus un logiciel permettant de calculer les métriques nécessaires à l'évaluation et la comparaison des différents algorithmes. Des bases de données récentes semblent émerger et sont les suivantes :

— **ChangeDetection.net Dataset :** La base de données ChangeDetection.net[11] [204] comporte environ 90000 images sur 31 séquences classées en 6 catégories couvrant ainsi tous les challenges et aussi le domaine de l'infra-rouge. Ces catégories sont les suivantes : Baseline, Dynamic Background, Camera Jitter, Shadows, Intermittent Object Motion and Thermal. La Figure 2.13 montre des images représentatives de chaque catégorie ainsi que les images vérité-terrain correspondante. De plus, les résultats du workshop CDW 2012 sont disponibles sur le même site que la base de données[12].

Pour le workshop CDW 2014, cette base de données a été enrichie de 5 catégories pour couvrir plus de challenges.

11. Authors : **F. Porikli**, P. Jodoin, J. Konrad, P. Ishwar (en gras co-guest editor du handbook).
12. http://www.changedetection.net/

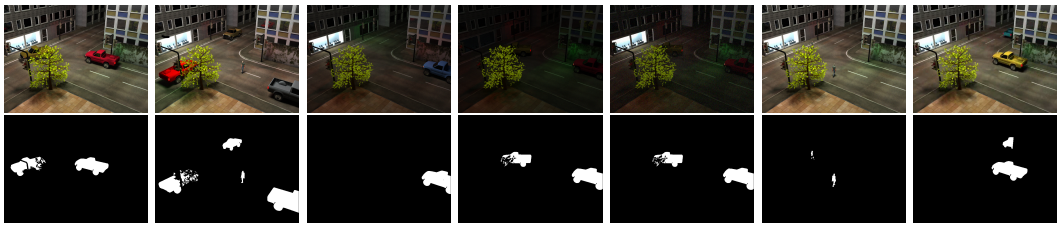


FIGURE 2.14 – SABS dataset [82] : La première ligne montre une image représentative d'une séquence de chaque catégorie et la seconde montre la vérité-terrain correspondante dans l'ordre suivant : Basic, Bootstrapping, Darkening, Light Switch, Noisy Night, Camouflage and Video Compression.

— **SABS Dataset :** La base de données SABS[13] (Stuttgart Artificial Background Subtraction) [82] comporte neuf séquences réalisées par la synthèse d'images et comportant des challenges qui apparaissent dans des scènes réelles. Chaque image vérité-terrain est ainsi fournie de façon très précise. Les catégories suivantes sont ainsi présentées : Basic, Dynamic Background, Bootstrapping, Darkening, Light Switch, Noisy Night, Shadow, Camouflage et Video Compression. La Figure 2.14 montre des images représentatives de chaque catégorie ainsi que les images vérité-terrain correspondantes. De plus, les résultats de la comparaison [82] sont disponibles sur le même site[14].

— **BMC Dataset :** La base de données BMC[15] (Background Models Challenge) [530] comporte à la fois des séquences de synthèse et des séquences réelles pour permettre une mise au point et une comparaison rigoureuse des méthodes de suppression de fond. Cette base de données est celle du workshop correspondant organisé lors de la conférence ACCV (Asian Conference in Computer Vision). Cette base de données est donc constituée comme suit : (1) Un ensemble de 20 séquences de synthèse représentant des scènes d'extérieurs de circulation routière sur une route et un rond point. Plusieurs challenges ont été simulés comme les variations lumineuses dues au brouillard, au soleil et le bruit d'acquisition. Chaque séquence est découpée en deux parties : Une partie d'apprentissage et une partie d'évaluation. (2) Un ensemble de 9 séquences réelles de vidéo-surveillance est aussi fourni pour la phase d'évaluation. Elles ont la caractéristique d'être de longue durée, permettant d'évaluer ainsi l'algorithme au cours d'un nombre important de changements. De plus, certains fonds ont des challenges spécifiques comme les fonds dynamiques, les ombres portées et les variations lumineuses rapides. La Figure 2.15 montre des images représentatives de chaque catégorie ainsi que les images vérité-terrain correspondante. De plus, les

13. Authors : S. Brutzer, **B. Höferlin**, G. Heidemann (en gras co-guest editor du handbook).
14. http://www.vis.uni-stuttgart.de/index.php?id=sabs
15. Authors : **A. Vacavant**, T. Chateau, L. Tougne (en gras co-guest editor du handbook).
16. http://bmc.univ-bpclermont.fr

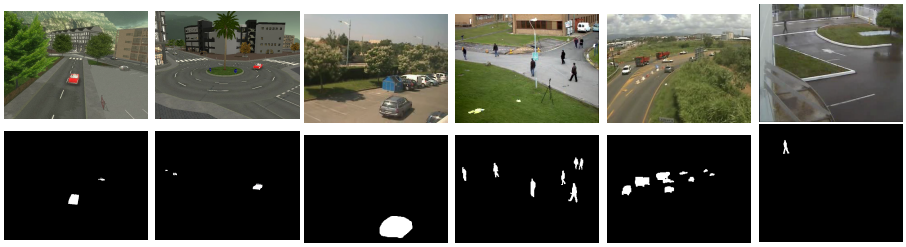


FIGURE 2.15 – BMC dataset [530] : La première ligne montre une image représentative d'une séquence de chaque catégorie et la seconde montre la vérité-terrain correspondante dans l'ordre suivant : Sequence Street, Sequence Rotary, Sequence Video 001, Sequence Video 003, Sequence Video 008 and Sequence Video 009.

FIGURE 2.16 – MAR [67] : La première ligne montre une image représentative d'une séquence de chaque catégorie et la seconde montre la vérité-terrain correspondante

résultats du workshop BMC 2012 sont disponibles sur le même site que la base de données[16].

Pour le futur workshop BMC, cette base de données a été enrichie de plusieurs vidéos provenant des bases de données suivantes : MAR Dataset, RGB-D Object Detection Dataset, Fish4knowledge Dataset[17] et de Aqu@theque Dataset[18]. Cette extension permet d'évaluer les algorithmes sur un plus grand nombre de challenges mais aussi sur des objets d'intérêts de formes et de couleurs beaucoup plus variées.

— **MAR Dataset :** La base de données MAR[19](Maritime Activity Recognition) [67] contient des vidéos de surveillance maritime. Il y a 10 vidéos issues de caméras fixes avec les images vérité-terrain et 15 vidéos issues de caméras PTZ. La Figure 2.16 montre des images représentatives de chaque catégorie ainsi que les images vérité-terrain correspondantes.

— **RGB-D Object Detection Dataset :** La base de données RGB-D[20] [89] est la pre-

17. http://www.fish4knowledge.eu
18. http://sites.google.com/site/thierrybouwmans/recherche—aqu-theque-dataset
19. http://labrococo.dis.uniroma1.it/MAR/
20. http://www.gti.ssr.upm.es/ mac/

FIGURE 2.17 – RGB-D dataset [89] : La première ligne montre une image représentative d'une séquence de chaque catégorie et la seconde montre la vérité-terrain correspondante dans l'ordre suivant : General, Shadows, Depth Camouflage et Color Camouflage

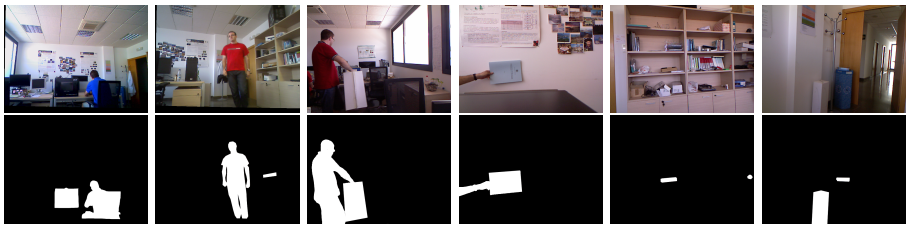


FIGURE 2.18 – CITIC dataset [182] : La première ligne montre une image représentative d'une séquence de chaque catégorie et la seconde montre la vérité-terrain correspondante dans l'ordre suivant : séquences provenant d'un systéme stéréoscopique (LCDScreen, Labdoor), séquences provenant de la Kinect (Chairbox, Wall, Shelves et Hallway)

mière à avoir comporté des séquences acquises avec les caméras de la Microsoft Kinect RGB-D, permettant ainsi d'avoir la caractéristique de profondeur. Chaque séquence comporte un challenge comme les ombres, le camouflage en couleur et le camouflage en profondeur. Chaque séquence est fournie avec les images vérité-terrain. La Figure 2.17 montre des images représentatives de chaque catégorie ainsi que les images vérité-terrain correspondantes.

— **CITIC RGB-D Dataset :** Cette base de données [182] disponible à partir du site du laboratoire CITIC[21] contient des séquences acquises avec un système stéréoscopique et des séquence acquises avec une caméra Microsoft Kinect, permettant ainsi d'avoir la caractéristique de profondeur. Chaque séquence comporte un challenge comme les ombres, le camouflage en couleur et le camouflage en profondeur. Chaque séquence est fournie avec les images vérité-terrain. La Figure 2.18 montre des images représentatives de chaque catégorie ainsi que les images vérité-terrain correspondantes.

En résumé, il existe deux bases de données à large spectre de challenges : 1) ChangeDetection.net qui contient de séquences dans le domaine visible et infra-rouge et 2) BMC qui comporte à la fois des séquences de synthèse et des séquences réelles. La base de données

21. http://atcproyectos.ugr.es/mvision/

SABS ne contient quant à elle que des séquences de synthèse rendant ainsi son domaine de validité plus limité. D'autre part, la base de données MAR concerne la surveillance maritime. Finalement, deux bases de données permettent l'évaluation avec l'information de profondeur. Pratiquement, il serait idéal d'avoir une base de données à large spectre de challenges mais aussi à large spectre d'applications. En effet, pour l'instant, ces bases de données se limitent à des séquences issues de la vidéo surveillance d'activités humaines. Or les applications de la vidéo surveillance d'activités animales et des environnemenst naturels ainsi que des applications mobiles sont en pleine extension. Cette base de données à large spectre de challenges mais aussi à large spectre d'applications permettrait donc d'évaluer les algorithmes sur un plus grand nombre de challenges mais aussi sur des objets d'intérêts de formes et de couleurs beaucoup plus variées.

10 Conclusion

Dans ce chapitre, nous avons défini un cadre commun pour les méthodes de suppression de fond, en identifiant les étapes (modélisation, initialisation, maintenance et classification) et les choix clés (taille et type des caractéristiques) qui déterminent la robustesse de la méthode face aux différents challenges rencontrés. Ceux-ci ont été ensuite classés suivant leurs caractéristiques spatiales et temporelles. Puis, nous avons présenté un état de l'art complet des méthodes traditionnelles et récentes pour donner une vue globale des modèles mathématiques utilisés et pour situer les approches développées dans les travaux de recherche effectués par Fida El Baf, Diana Farcas, Cristina Marghes, et Charles Guyon sous ma direction scientifique. En outre, nous avons identifié et classé les différentes stratégies possibles pour appliquer de façon optimale ces méthodes de suppression de fond. Les aspects capteurs ont aussi été abordés montrant ainsi leur importance quand à la disponibilité de telle ou telle caractéristique. De plus, nous avons abordé les avantages et inconvénients de la plupart des méthodes d'évaluation. Enfin, nous avons effectué une analyse des différentes bases de données pour en définir leurs avantages et leurs limites.

Les perspectives de travail dans ce domaine restent nombreuses comme l'amélioration des méthodes que nous avons proposées et qui sont développées dans les chapitres suivants. En outre, l'étape d'initialisation a reçu relativement peu d'attention et peu d'articles la concernent alors que la qualité de la première image de fond est relativement primordiale pour le reste du processus. S'attacher à cette étape présente donc un intérêt majeur. Pour l'aspect mise à jour du fond, il existe dans la littérature différents mécanismes mis en place pour gérer des challenges spécifiques et il serait intéressant de les associer pour obtenir un mécanisme robuste simultanément à un grand nombre de challenges. En outre, très peu de travaux se sont intéressés vraiment à une étude approfondie des différentes caractéristiques et cela reste donc un problème ouvert dans le cadre d'une méthode de sélection de caractéristiques permettant d'employer les caractéristiques les plus adaptées pour une zone de l'image où apparait tel ou

tel challenge.

Chapitre 3

Apport des concepts flous dans la détection d'objets mobiles

1 Introduction

Dans ce chapitre [1], nous nous focalisons sur l'apport des concepts flous dans la détection d'objets mobiles dans des séquences video. **Ces travaux ont été menés sous ma direction scientifique dans le cadre de la thèse de Fida El Baf.**

Nous développons ainsi comment en utilisant la théorie des sous-ensembles flous peuvent être résolus les problèmes apparaissant lors de la détection d'objets mobiles dans des fonds dynamiques. L'objectif est de prendre en compte l'imprécision et l'incertitude au niveau de la modélisation du fond, de la classification fond/objet et de la mise à jour du fond en présence de fonds dynamiques et de variations lumineuses comme suit :

— **Modélisation floue du fond (Section 3)** : Dans cette section, nous intéressons à la modélisation de fond par le mélange de Gaussiennes floues de Type-2 pour prendre en compte l'incertitude apparaissant en présence de fonds dynamiques pour l'estimation des paramètres (moyenne et variance) de mélanges de Gaussiennes. Cette méthode a été appliquée à des scènes dynamiques du domaine visible et infrarouge.
Publications relatives : 1 journal (MASAUM [74]), 2 conférences (ISVC [46], CVPR workshop OTCBVS [47])

— **Classification floue des pixels en fond/objet (Section 4)** : Dans cette seconde section, nous développons une méthode de détection d'objets mobiles qui utilisent la caractéristique couleur et une caractéristique texture (LBP [245]). Les résultats obtenus par ces deux caractéristiques sont ensuite fusionnés par l'intégrale floue de Choquet permettant ainsi à la méthode développée d'être plus robuste aux variations lumi-

1. Ce chapitre est basé sur le manuscrit de thèse de Fida El Baf.

neuses.
Publications relatives : 3 conférences (WIAMIS [42], FUZZ'IEEE [45], CVPR workshop OTCBVS [44])

— **Mise à jour de l'image du fond (Section 5)** : Finalement, nous présentons une méthode floue de mise à jour adaptative de l'image du fond à partir du degré d'appartenance du pixel à la classe "fond" et à la classe "objet mobile".
Publication relative : 1 conférence (ICIP [43])

Pour chaque partie, un site web a été mis en ligne : T2FMOG[1], Fuzzy Foreground Detection[2] et Fuzzy Background Maintenance[3]. Ces différents sites permettent de visualiser les résultats obtenus et de télécharger les codes correspondants développés en C et OpenCV sous Microsoft Visual C. De plus, ces algorithmes font partie de la librairie relative au Background Subtraction développée en 2012 par Andrews Sobral (Federal University of Bahia, Brazil). Cette librairie développée avec OpenCV et C++ s'appelle BGS Library[4].

2 Approche par les concepts flous : Une vue d'ensemble

Les différents challenges rencontrés dans les vidéos engendrent des imprécisions et des incertitudes tout au long du processus de la détection des objets mobiles. Pratiquement, il existe trois théories mathématiques permettant de gérer les problèmes d'imprécisions et d'incertitudes qui sont les suivantes : la théorie statistique, la théorie des sous-ensembles flous et la théorie de Dempster-Schafer. Dans ce cadre, nous avons choisi en 2004 d'employer la théorie des sous-ensembles flous et nous sommes ainsi dans les premiers à avoir proposé l'utilisation des concepts flous dans ce domaine dans le cadre de la thèse de Fida El Baf. Ces travaux ont donc été précurseurs dans le domaine. Depuis, d'autres travaux ont été développés en parallèle et d'autres sont basés sur nos travaux. Nous présentons donc dans cette section un état de l'art de ces méthodes.

Les concepts flous ont donc été employés dans les différentes étapes des méthodes de suppression de fond pour prendre en compte ces problémes d'imprécisions et d'incertitudes. Les différentes méthodes peuvent être classées suivant les étapes concernées comme développé dans notre état de l'art [71].

1. http://sites.google.com/site/t2fmog/
2. http://sites.google.com/site/fuzzyforegrounddetection/
3. http://sites.google.com/site/fuzzybackgroundmaintenance/
4. http://code.google.com/p/bgslibrary/

2.1 Principe des systèmes basés sur les concepts flous

Nous rappelons tout d'abord brièvement le cadre général d'application des concepts flous. Tous les systèmes employant les concepts flous sont basés sur le même principe qui comporte pour la plupart les étapes suivantes : la fuzzification, une base de règles floues, un moteur d'inférence et la défuzzification. La figure 3.1 présente ces différentes étapes sous forme graphique. Ces étapes peuvent être définies de façon concise comme suit :

— **Fuzzification :** La fuzzification se définit comme l'étape de passage d'une valeur discrète à un valeur floue. Ce passage s'effectue à partir d'une fonction d'appartenance (ou fonction d'adhésion) qui détermine le degré d'appartenance (valeur floue) d'un élément considéré à un ensemble flou correspondant à une valeur linguistique (comme petit, moyen et grand). Les fonctions d'appartenances les plus courantes sont les fonctions triangulaires et sigmoïdes.

— **Base de règles floues :** Les variables floues et leurs relations correspondantes sont modélisées par une systèmes de règles floues de type "si-alors" (if-then) de la façon suivante : *si proposition "antécédent" (ou prémisse) alors proposition "conséquence" (ou conclusion)*. Sur ce principe, ils existent principalement deux grands types de modèles de base à règles floues qui sont : 1) les modèles linguistiques où les règles sont exprimées en utilisant les variables linguistiques. Les propositions "antécédent" et "'conséquence" sont toutes les deux floues ; 2) les modèles de type Takagi-Sugeno (TS) où la proposition "antécédent" est floue et la proposition "conséquence" est discrète.

— **Moteur d'inférence :** Le moteur d'inférence est constitué de l'ensemble des règles floues et définit comment ces règles agissent. Pratiquement, il s'agit d'un algorithme de simulation de raisonnements déductifs, permettant ainsi aux systèmes de conduire des raisonnements logiques et de dériver des conclusions à partir des règles floues. Il existe différent type de moteur d'inférence et le plus connu est celui de Mamdani (*Mamdani-type Fuzzy Inference System*).

— **Defuzzification :** La défuzzification consiste à obtenir à la fin du processus une valeur discrète qui correspond donc à une décision binaire. Il existe plusieurs méthodes pour cette étape comme celle qui consiste à prendre la moyenne, le minimum ou le maximum des valeurs floues.

Dans le cas de plusieurs variables, les variables floues sont agrégées par des opérateurs flous comme les intégrales floues de Sugeno ou de Choquet comme développé dans la section 4.1.

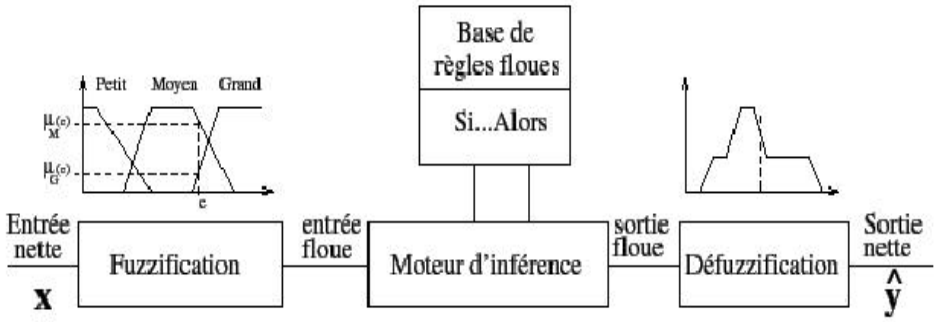


FIGURE 3.1 – Vue d'ensemble d'un système d'inférence basés sur les concepts flous [80].

2.2 Modélisation floue de fond

Le principal challenge visé est la modélisation de fonds dynamiques et les différentes approches peuvent être classées comme suit :

— **Histogramme flou :** Kim et Kim [286] ont adopté un modèle de clustering basé sur les fuzzy c-means qui est appliqué sur des histogrammes couleur flous. Ce modèle permet d'atténuer l'influence des variations couleur générées par les mouvements du fond et de mettre en valeur les objets mobiles. Les expérimentations montrent une robustesse plus grande que les MOG [490] en présence de fonds dynamiques. Kanna et Murthy [269] ont utilisé une approche identique en ajoutant un processus de mise à jour adaptatif. Yang et al. [590] ont quant à eux utilisé en plus un champ de Markov pour prendre en compte la cohérence spatiale.

— **Mélange de Gaussiennes floues :** L'algorithme le plus utilisé pour modéliser les fonds dynamiques est sans nul doute le mélange de Gaussiennes [490] mais les paramètres de ce modèle (moyenne, variance) sont déterminés à partir d'une séquence d'apprentissage qui peut être bruitée ou présenter des données manquantes à cause de la présence des objets mobiles. Dans ce contexte, nous avons proposé l'utilisation des mélanges de Gaussiennes floues de Type-2 [46][47][74]. Ainsi, deux algorithmes modélisant respectivement l'incertitude autour de la moyenne et l'incertitude autour de la variance, ont été développés et sont nommés respectivement T2-FMOG-UM et T2-FMOG-UV. Ces deux algorithmes sont plus robustes que la version originale de l'algorithme MOG [490]. Cette approche est développée dans la Section 3.

— **Réseau de neurones flou :** Chacon-Muguia et Gonzalez-Duarte [98] ont proposé un réseau de neurones basé sur des concepts flous qui permet d'être plus robuste en présence de variations lumineuses qu'elles soient lentes ou rapides ainsi qu'en présence de fonds dynamiques. Chacon-Murguia et al. [100] ont ensuite amélioré cette ap-

proche par un processus adaptatif flou de mise à jour. Par ailleurs, Chen et al. [105] ont proposé une version floue de réseau de neurones de type ART *Adaptive Resonant Theory*. Chen et al. [106] ont ensuite amélioré ce modèle en ajoutant un facteur de cohérence spatial.

— **Estimation floue :** Li et al. [607] ont employé pour estimer le fond un ensemble de règles floues de type TSK (Takagi-Sugeno-Kang) [608]. L'apprentissage du système de règles floues est effectué à partir d'une approche hybride combinant une approche de type *particle swarm optimization* et une approche de *kernel least mean square*. Pour estimer le fond, les objets mobiles sont d'abord vus comme des outliers et séparés du fond par un opérateur. Ensuite, les résultats de l'opérateur sont employés pour entrainer l'estimateur flou et estimer le fond à chaque pixel. Cette approche peut être améliorée par un ensemble de règles floues de type TSK de Type-2 [429].

Pratiquement, la plupart des méthodes de modélisation classiques peuvent être améliorés par l'utilisation des concepts flous.

2.3 Classification floue fond/objets mobiles

A cette étape, une fonction linéaire de type *saturating linear function* peut être employée pour éviter de prendre une décision binaire lors de la classification en fond/objets mobiles. Certains auteurs ont utilisé ce principe dans le cas de fond unimodal en employant la moyenne courante [477][457], et d'autres dans le cas de fonds multimodaux en employant une mesure de confiance [439]. Yao et Ahmad [593] ont quant à eux employer un modèle de fond basé sur un histogramme dans l'espace couleur HSV et ont ensuite appliqué un fuzzy k-means pour la classification.

Une autre approche consiste à employer des concepts flous pour agréger des résultats provenant de différentes caractéristiques. En 2006, Zhang et Xu [620] ont été les premiers à publier une approche utilisant l'intégrale de Sugeno pour agréger les résultats venant de la couleur et de la texture, améliorant ainsi la combinaison de ces résultats par rapport à l'emploi d'un "OU" ou d'un "ET" logique. Cependant, l'hypothèse induite par l'utilisation de l'intégrale de Sugeno est que les valeurs des pixels ont une relation de type ordinal, ce qui n'est pas le cas car c'est une relation de type cardinal. Dans ce contexte, nous avons proposé en 2008 une approche qui emploie l'intégrale de Choquet [42][45][44] (voir Section 4). Cette approche a ensuite été reprise par Ding et al. [160] en changeant la mesure de similarité employée pour comparer le fond et l'image courante. Azab et al. [33] ont quant à eux repris cette approche en employant trois caractéristiques (couleur, contour et texture) au lieu des deux employées dans notre approche. Ces approches multi-caractéristiques floues sont plus robustes aux variations lumineuses que leur version binaire basée sur l'emploi d'un "OU" ou d'un "ET" logique.

D'une autre façon, Bowen [80] a proposé d'effectuer la classification en se basant sur un système de règles d'inférence. Ce système de type Mamdani (*Mamdani-type Fuzzy Inference System*) est appliqué sur les paramètres du modèles (moyenne et variance) calculés pour chaque composante de l'espace couleur YCrCb. Les expérimentations menées sur la base de données Wallflower montrent que cette approche est plus robuste que les méthodes courantes comme les MOG.

Wu et Wu [565] ont proposé de déterminer le seuil optimal en chaque pixel à l'aide d'une fonction floue non symétrique (*Fuzzy Non-Symmetric Membership Function (NSMF)*. La classification est effectuée dans chaque composante de l'espace RGB et les résultats sont ensuite fusionnés à l'aide de l'opérateur floue "max".

Une autre approche floue développée par Revathi et Hemalatha [430] consiste à comparer l'image de fond et l'image courante à l'aide d'une fonction floue appelée *Intuitionistic Fuzzy Function*. Cette approche semble adaptée à l'application de détection d'objets abandonnés.

2.4 Mise à jour adaptative floue du fond

L'idée est de mettre à jour le fond suivant le degré d'appartenance d'un pixel à la classe "fond" ou "objets mobiles". Le degré d'appartenance vient de l'étape précédente de classification des pixels en "fond" ou "objets mobiles" et peut être employé dans la mise en jour de deux manières.

La première manière consiste à adapter le coefficient d'apprentissage en fonction du degré d'appartenance du pixel. Dans ce cas, le coefficient d'apprentissage est une fonction du degré d'appartenance du pixel et plusieurs ont été employés dans la littérature :

1. **Fonction exponentielle :** Sigari et al. [477] ont employé une fonction exponentielle *Exponential function*. Yeo et al. [595] ont quant à eux amélioré cette approche par l'emploi d'une fonction exponentielle adaptative *Adaptive Exponential function*.

2. **Fonction linéaire :** Maddalena et al. [351] ont proposé une fonction linéaire appellée *Saturating linear function*. Maddalena et al. [351] ont ensuite ajouté un terme de cohérence spatiale dans cette fonction. Ce mécanisme fait partie des méthodes SOBS [350], SOBS-SC [354] et 3dSOBS+ [355] qui sont des méthodes "leader" du domaine.

La seconde manière consiste à agir au niveau de la règle de mise à jour. Par exemple, nous avons proposé une règle de mise à jour floue qui est une combinaison de deux règles de mise à jour binaires [43] (voir Section 5). Kashani et al. [272][273] ont repris cette règle de mise à jour. Cette règle de mise à jour adaptative floue permet d'être plus robuste dans le cas de variations lumineuses et d'ombres. Lijun [322] a proposé une règle de mise à jour basée sur des

règles d'inférences floues. Calvo-Gallego et al. [87] ont quant à eux utilisé un système à deux fonctions linéaire floues basées sur des variables linguistiques petites et grandes associées à la différence entre l'image de fond et l'image courante. Calvo-Gallego et al. [87] obtiennent ainsi une règle de mise à jour plus adaptative que celle que nous avons développée. Cette approche peut être améliorée par l'utilisation d'une fonction sigmoide à la place d'une fonction linéaire.

2.5 Caractéristiques floues

Chiranjeevi et Sengupta [119] ont été les premiers dans ce domaine à proposer une caractéristique floue appelée corrélogramme floue *fuzzy correlogram*, obtenu en appliquant un algorithme fuzzy c-means sur un corrélogramme binaire. Chiranjeevi et Sengupta [120] ont ensuite proposé une caractéristique texture floue générée à partir des vecteurs de co-occurence. Les résultats obtenus par cette caractéristique texture appelée *fuzzy statistical texture features (FST)* sont ensuite combinés avec les résultats provenant des caractéristiques couleur en utilisant l'intégrale de Choquet. Les expérimentations [120] montrent une amélioration notable des résultats par rapport à l'utlisation de la version binaire de la caractéristique texture. D'une autre façon, Chiranjeevi et Sengupta [121] ont développé une version floue de la caractéristique 3D histon. La classification emploie la distance de Bhattacharyya pour comparer le 3D fuzzy histon de l'image courante avec l'image du fond. Cette approche semble robuste aux fonds dynamiques.

2.6 Post-traitements flous

Des post-traitement flous peuvent être appliqués aux masques de détection fournis par la classification fond/objets mobiles. De plus, certains auteurs comme Sivabalakrishnan et Manjula [481] ont appliqué une règle d'inférence floue entre les masques consécutifs pour rendre la détection cohérente temporellement.

Les modèles flous se montrent robustes en présence de variations lumineuses et de fonds dynamiques. Le tableau 3.1 montre une vue d'ensemble des différentes méthodes de suppression de fond basées sur des concepts flous. Dans les section suivantes, nous présentons les trois approches que nous avons développées concernant dans l'ordre l'étape de modélisation [46], l'étape de classification [42] et l'étape de mise à jour [43].

Méthodes floues	Algorithmes	Auteurs - Dates
Modélisation floue	1) Fuzzy Color Histogram	
	Fuzzy C-means Clustering (FCM) (1)	Kim and Kim (2012) [286]
	Fuzzy Color Histogram (FCH) (2)	Kanna and Murthy (2013) [269]
	Multi-histogram Clustering (MHC) (1)	Yang et al. (2010) [590]
	2) Fuzzy Mixture of Gaussians	
	Type-2 Fuzzy MOG (T2-FMOG) (3)	**El Baf et al. (2008) [46]**
	3) Fuzzy Neural Network	
	Fuzzy SOMM (2)	Chacon-Muguia and Gonzalez-Duarte (2011) [98]
	Fuzzy ART (2)	Chen et al. (2010) [105]
	4) Fuzzy Estimation	
	Takagi-Sugeno-Kang (TSK) (1)	Li et al. (2013) [607]
Classification fond/objets mobiles floue	1) Cas d'une caractéristique	
	Saturating Linear Function (SLF) (5)	Sigari et al. (2008) [477]
	Fuzzy k-means (1)	Yao and Ahmad (2009) [593]
	2) Cas de plusieurs caractéristiques	
	Sugeno Integral (SI) (1)	Zhang and Xu (2006) [620]
	Choquet Integral (CI) (5)	**El Baf et al. (2008) [45]**
	Mamdani-type Fuzzy Inference System (FIS) (1)	Bowen (2013) [80]
	3) Seuil	
	Fuzzy Non-Symmetric Membership Function (2)	Wu and Wu (2011) [565]
	4) Fonction de comparaison	
	Intuitionistic Fuzzy Logic (1)	Revathi and Hemalatha (2013) [430]
Mise à jour adaptative floue	1) Fuzzy Learning Rate (FLR)	
	Exponential function (4)	Sigari et al. (2008) [477]
	Adaptive Exponential function (2)	Yeo et al. (2011) [595]
	Saturating linear function (1)	Maddalena et al. (2009) [351]
	Saturating linear function - Spatial Coherence (5)	Maddalena et al. (2010) [353]
	2) Fuzzy Maintenance Rule (FMR)	
	Fuzzy Maintenance Rule (FMR) (3)	**El Baf et al. (2008) [43]**
Caractéristiques floues	Fuzzy Correlogram (FC) (1)	Chiranjeevi and Sengupta (2011) [119]
	Fuzzy statistical texture features (FST) (1)	Chiranjeevi and Sengupta (2012) [120]
	Fuzzy 3D Histon (F3DH) (1)	Chiranjeevi and Sengupta (2012) [121]
Post-Processing	Fuzzy Inference (FI) (5)	Sivabalakrishnan and Manjula (2010) [481]

TABLE 3.1 – Vue d'ensemble des différentes méthodes de suppression de fond basées sur des concepts flous. La première colonne indique l'étape concernée. La seconde colonne indique le concept flou employé avec l'acronyme de la méthode et le nombre de publications entre parenthèses. La troisième colonne indique le nom des auteurs et la date de la première publication. En gras, les approches développées par Fida El Baf sous ma directions scientifique.

3 Modélisation du fond par un mélange de Gaussiennes floues de Type-2

Avant d'introduire le modèle de fond par mélange de Gaussiennes floues de Type-2 (T2-FMOG), nous discutons tout d'abord dans le préambule les limites du modèle de mélange de Gaussiennes classique (MOG), le principe des ensembles flous de Type-2 et comment l'incertitude liée aux paramètres des MOG peut être gérée par les ensembles flous de Type-2. Ceci nous permettra enfin de définir le mélange de Gaussiennes floues de Type-2 et ainsi de présenter la méthode proposée de modélisation du fond par T2-FMOG.

3.1 Préambule

Limitations du modèle basique MOG : La modélisation d'un pixel par un mélange de Gaussiennes (MOG) permet de représenter le fond en tolérant certains changements dans la scène. Avec ce modèle, les paramètres sont estimés en maximisant la vraissemblance. Il faut savoir qu'une erreur d'approximation peut entraîner une mauvaise classification du pixel en fond ou objet d'intérêt. D'une manière générale, le MOG s'adapte bien à un fond multimodal et aux changements lumineux mais quand il s'agit de données manquantes ou du bruit associé à ces données, il peut ne pas refléter la forme de la distribution réelle des données. On peut traduire ce fait par une incertitude sur la vraissemblance qui n'est pas prise en compte dans la version classique du MOG. D'où l'idée d'utiliser les ensembles flous de Type-2 qui permet justement de prendre en compte ce genre de problèmes de façon adéquate.

Ensembles flous de Type-2 (T2-EF) : Les ensembles flous de Type-2 (T2-EF) ont été initialement introduit par Zadeh [609]. Ils ont été présentés comme une extension des ensembles flous de Type-1 (T1-EF) et ont l'avantage de fournir plus de liberté aux sytèmes à logique floue (SLF). Plusieurs auteurs ont cherché à étudier les propriétés des T2-EF comme Mizumoto and Tanaka [369], John et Mendel [363][364], Kulr et Folger [294]. Dans [294], Kulr et Folger ont expliqué que les fonctions d'adhésion (FA) du Type-1 peuvent être une problématique, car la représention du flou est effectuée en attribuant aux fonctions d'adhésion des valeurs qui sont eux mêmes des nombres réels précis. Or ces valeurs peuvent être également floues. En d'autre teme, elles peuvent être définies comme une fonction floue et pas comme des nombres précis. D'où l'idée d'étendre le concept des T1-EF aux T2-EF. Cette caractéristique augmente la performance des T2-SLF par rapport aux T1-SLF surtout quand ces systèmes présentent beaucoup d'incertitude. Pour réduire la complexité du calcul que demandent les T2-EF, [364] ont utilisé récemment une opération appelée réduction de type qui permet de réduire l'ensemble de Type-2 à un ensemble de Type-1. Cet ensemble de type réduit joue le rôle d'un intervalle de confiance pour les incertitudes. Le calcul associé aux intervalles T2-EF utilisé dans les T2-SLF ne demande que des opérations arithmétiques simples. On donne aux intervalles des systémes à logique floue le nom IT2-SLF.

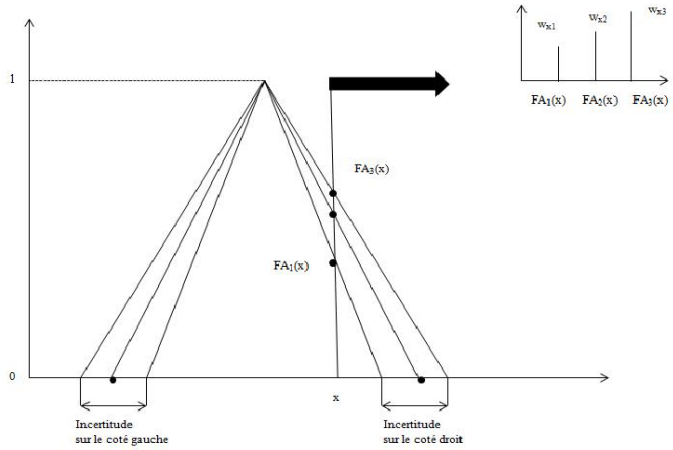
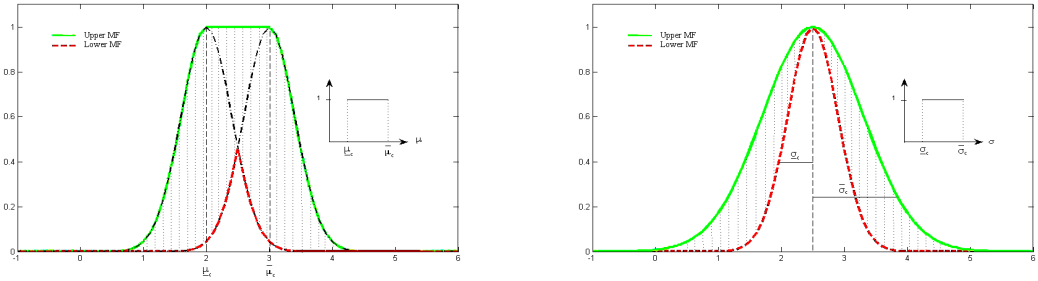


FIGURE 3.2 – Exemple de fonction d'appartenance de Type-2 dans le cas d'une fonction triangle [363].

Dans les ensembles flous de Type-1, la projection d'une valeur donnée x croise la fonction d'adhésion en un seul point. Ce qui fait que sa valeur est clairement définie. A la différence avec les T2-EF, on peut prendre en considération une incertitude liée au T1-EF. Cela peut être traduit par une marge avec laquelle, on tolère à la fonction d'adhésion de varier dans un intervalle et d'intégrer, en quelque sorte, l'incertitude. En d'autre terme, pour une valeur donnée x, on peut avoir autant de valeurs qu'on a de Fonctions d'Adhésion Primaires (FAP). Ce qui crée la troisième dimension des T2-EF (appelée Fonction d'Adhésion Secondaire FAS) ainsi que la zone d'encombrement (appelée zone "FOU"). Dans ce cas, la valeur de la fonction d'adhésion que prend x est considérée elle-même comme une fonction floue. En générale, cette Fonction d'Adhésion Secondaire FAS peut être une Gaussienne, un intervalle à possibilités uniformes, un triangle ou n'importe quelle autre forme des fonctions d'adhésion usuelles. La valeur mimimale ($\underline{h}$) et la valeur maximale des FAP associées à x définissent les bornes inférieure et supérieure de la FAS. Pour simplifier le calcul, la plupart des auteurs utilisent l'intervalle à possibilités uniformes (IT2-SLF) qui permettent d'effectuer des opérations mathématiques simples, car les valeurs secondaires sont toutes égales à l'unité. Notons que la zone d'encombrement reflète le degré d'incertitude liée au système. Elle est d'autant plus grande que l'incertitude est importante et inversement.

Gestion de l'incertitude des paramètres du MOG par les T2-EF : Grâce à la troisième dimension que fournissent les fonctions d'adhésion (FA) des ensembles flous de Type-2, le mélange de Gaussiennes floues de Type-2 (T2-FMOG) peut gagner en gestion de l'incertitude liée aux paramètres du MOG. Cela est réalisé en utilisant la zone d'encombrement illustrée dans la Figure 3.3 (appelé en anglais « footprint of uncertainty » ou « FOU ») aussi bien que l'intervalle des fonctions d'adhésion secondaire. Ainsi, en faisant varier les paramètres dans un intervalle donné avec des possibilités unifomes, on peut considérer que la distribution arrive à approcher la distribution réelle à un moment donné. En d'autre terme, au lieu d'avoir

une valeur exacte de la vraissemblance tel est le cas avec le modèle classique MOG, avec le T2-FMOG cette vraissemblance varie dans un intervalle. Cela permet de gérer l'effet des valeurs aberrantes qui peuvent être présentes dans le jeu de données. En effet, la présence de telles valeurs conduit à une mauvaise approximation des paramètres du mélange, et par conséquent on risque de s'éloigner de la forme exacte de la distribution. Par contre, avec la marge de variation qu'on donne à la vraissemblance, on considère qu'on approche le mieux possible la distribution exacte à un moment donné, ce qui permet d'atténuer l'effet des valeurs aberrantes.



Incertitude sur la moyenne Incertitude sur la variance

FIGURE 3.3 – De la gauche vers la droite : la fonction d'adhésion de la Gaussienne primaire avec une incertitude sur la moyenne et la fonction d'adhésion de la Gaussienne primaire avec une incertitude sur la variance. Les intervalles sont à possibilités uniformes. En rouge et en vert sont les fonctions d'adhésions inférieure et supérieure respectivement. La région hachurée correspond à la zone du FOU

Selon la théorie des ensembles flous de Type-2 (T2-EF), les fonctions d'adhésion de Type-2 utilisent une première dimension pour les fonctions d'adhésion primaires pour décrire l'incertitude de l'observation, et une deuxième dimension pour les fonctions d'adhésion secondaires. Ces dernières évaluent l'incertitude associée aux fonctions d'adhésion primaires. Pour illustrer ce schéma, on montre une fonction d'adhésion Gaussienne typique avec une incertitude sur la moyenne variant dans un intervalle à possibilités uniformes $\left[\underline{\mu}, \overline{\mu}\right]$. La fonction d'adhésion primaire d'une observation x est limitée par les bornes inférieure et supérieure des degrés d'appartenance $\underline{\mu}$ et $\overline{\mu}$. L'incertitude sur cet intervalle est décrite par une fonction d'adhésion secondaire qui, elle même, admet un intervalle à possibilités uniformes $\left[\underline{h\left(x\right)}, \overline{h\left(x\right)}\right]$. L'union de toutes les fonctions d'adhésion primaires constitue la zone d'encombrement ou la zone "FOU". La partie grisée dans la Figure 3.3 correspond à cette zone.

3.2 T2-FMOG

Le mélange de Gaussiennes (MOG) a été largement utilisé dans les problèmes de modélisation et de classification pour sa capacité à approcher la forme de la distribution des données grâce aux différentes Gaussiennes qui la composent. Pour une observation o, ce mélange de

Gaussiennes est exprimée en fonction de K composantes Gaussiennes comme suit :

$$P(o) = \sum_{i=1}^{K} \omega_i \eta(o, \mu_i, \sum_i) \tag{3.1}$$

où $\sum_{i=1}^{K} \omega_i = 1$ et $\omega_i > 0$ est un poids associé à la Gaussienne i. La densité de probabilité du mélange de Gaussiennes est donnée par :

$$\eta(o, \mu, \sum) = \frac{1}{(2\pi)^{\frac{d}{2}} |\sum|^{\frac{1}{2}}} exp\left(-\frac{1}{2}(o-\mu)^T \sum^{-1}(o-\mu)\right) \tag{3.2}$$

Etant définie l'observation o dans un espace de dimension d, le vecteur de la moyenne μ, et la matrice diagonale de la covariance $\sum = diag(\sigma_1^2, \ldots, \sigma_d^2)$, la distribution du mélange de Gaussiennes avec une incertitude autour de la moyenne est introduite de la façon suivante :

$$\begin{aligned} \eta(o, \tilde{\mu}, \sum) &= \frac{1}{(2\pi)^{\frac{d}{2}} |\sum|^{\frac{1}{2}}} exp\left[-\frac{1}{2}\left(\frac{o_1 - \mu_1}{\sigma_1}\right)^2\right] \ldots exp\left[-\frac{1}{2}\left(\frac{o_d - \mu_d}{\sigma_d}\right)^2\right] \\ & \mu_1 \in \left[\underline{\mu}_1, \overline{\mu}_1\right], \ldots, \mu_d \in \left[\underline{\mu}_d, \overline{\mu}_d\right] \end{aligned} \tag{3.3}$$

Dans le cas d'une incertitude autour de la variance, cette densité est définie comme suit :

$$\begin{aligned} \eta\left(o, \mu, \tilde{\sum}\right) &= \frac{1}{(2\pi)^{\frac{d}{2}} |\sum|^{\frac{1}{2}}} exp\left[-\frac{1}{2}\left(\frac{o_1 - \mu_1}{\sigma_1}\right)^2\right] \ldots exp\left[-\frac{1}{2}\left(\frac{o_d - \mu_d}{\sigma_d}\right)^2\right] \\ & \sigma_1 \in \left[\underline{\sigma}_1, \overline{\sigma}_1\right], \ldots, \mu_d \in \left[\underline{\sigma}_d, \overline{\sigma}_d\right] \end{aligned} \tag{3.4}$$

ou $\tilde{\mu}$ et $\tilde{\sigma}$ désignent les vecteurs de la moyenne et de la covariance incertaines respectivement. En supposant qu'aucune connaissance *a priori* n'est fournie sur l'incertitude des paramètres du modèle, on peut varier la moyenne et l'écart type dans un intervalle à possibilites uniformes i.e., $\mu \in \left[\underline{\mu}, \overline{\mu}\right]$ or $\sigma \in [\underline{\sigma}, \overline{\sigma}]$.
Dans le cas de la Gaussienne avec une incertitude sur la moyenne, la FA supérieure est donnée par :

$$\overline{h}(o) = \begin{cases} f\left(o; \underline{\mu}, \sigma\right), & \text{if } o < \underline{\mu} \\ 1, & \text{if } \underline{\mu} \leq o < \overline{\mu} \\ f(o; \overline{\mu}, \sigma), & \text{if } o > \overline{\mu} \end{cases} \tag{3.5}$$

oú $f\left(o; \underline{\mu}, \sigma\right) = exp\left[-\frac{1}{2}\left(\frac{o-\underline{\mu}}{\sigma}\right)^2\right]$ et $f(o; \overline{\mu}, \sigma) = exp\left[-\frac{1}{2}\left(\frac{o-\overline{\mu}}{\sigma}\right)^2\right]$.
La FA inférieure est donnée par :

$$\underline{h}(o) = \begin{cases} f(o; \overline{\mu}, \sigma), & \text{if } o \leq \frac{\underline{\mu}+\overline{\mu}}{2} \\ f\left(o; \underline{\mu}, \sigma\right), & \text{if } o > \frac{\underline{\mu}+\overline{\mu}}{2} \end{cases} \tag{3.6}$$

Dans le cas de la Gaussienne avec une incertitude autour de la variance, la FA supérieure est donnée par $\underline{h}(o) = f(o; \mu, \underline{\sigma})$ et la FA inférieure est décrite par l'expression : $\overline{h}(o) = f(o; \mu, \overline{\sigma})$.
Les facteurs k_m et k_ν déterminent les intervalles dans lesquels les paramètres du mélange de Gaussiennes varient :

$$\underline{\mu} = \mu - k_m\sigma, \quad \overline{\mu} = \mu + k_m\sigma, \quad k_m \in [0, 3], \tag{3.7}$$

$$\underline{\sigma} = k_\nu\sigma, \quad \overline{\sigma} = \frac{1}{k_\nu}\sigma, \quad k_\nu \in [0.3, 1]. \tag{3.8}$$

Parce que les données sont représentées par 99% dans l'intervalle $[\mu - 3\sigma, \mu + 3\sigma]$, Zeng et al. [616] ont contraint $k_m \in [0, 3]$ et $k_\nu \in [0.3, 1]$. Ces facteurs contrôlent la largeur de la zone du FOU. Plus k_m est grand (ou plus k_ν est petit), plus la zone du FOU est grande représentant ainsi une plus grande incertitude.

3.3 Application à la modélisation du fond

Chaque pixel est caracterisé par son intensité dans l'espace couleur RGB. Une observation o est donc un vecteur X_t de longueur 3 ($d = 3$) correspondant aux trois composantes couleurs de l'espace RGB. Le modèle de mélange de Gaussiennes (MOG) est composé d'un mélange de K composantes de Gaussiennes multivariées comme suit :

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t}\eta\left(X_t, \mu_{i,t}, \sum_{i,t}\right) \tag{3.9}$$

où les paramètres sont : K est le nombre des distributions, $\omega_{i,t}$ est le poids associé à la i^{ème} Gaussienne caractérisée par sa moyenne $\mu_{i,t}$ et son écart type $\sum_{i,t}$ mesurés à l'instant t. η est la densité de probabilité Gaussienne définie par :

$$\eta(X_t, \mu, \sum) = \frac{1}{(2\pi)^{\frac{3}{2}}|\sum|^{\frac{1}{2}}} exp\left(-\frac{1}{2}(X_t - \mu)^T \sum^{-1}(X_t - \mu)\right) \tag{3.10}$$

Nous pouvons définir ainsi le modèle de mélange de Gaussiennes floues de Type-2 avec une incertitude autour de la moyenne (T2-FMOG-UM) et le modèle de mélange de Gaussiennes floues de Type-2 avec une incertitude autour de la variance (T2-FMOG-UV) comme suit :

— **T2-FMOG-UM avec une incertitude autour de la moyenne** : La densité de probabilité est définie par :

$$\eta(X_t, \tilde{\mu}, \sum) = \frac{1}{(2\pi)^{\frac{3}{2}}|\sum|^{\frac{1}{2}}} \prod exp\left[-\frac{1}{2}\left(\frac{X_{t,c} - \mu_c}{\sigma_c}\right)^2\right] \tag{3.11}$$

avec $\mu_c \in \left[\underline{\mu}_c, \overline{\mu}_c\right]$ et $c \in \{R, V, B\}$.

— **T2-FMOG-UV avec une incertitude autour de la variance** : Dans ce cas, la densité de probabilité est définie par :

$$\eta\left(X_t,\mu,\tilde{\sum}\right) = \frac{1}{(2\pi)^{\frac{3}{2}}|\sum|^{\frac{1}{2}}}\prod exp\left[-\frac{1}{2}\left(\frac{X_{t,c}-\mu_c}{\sigma_c}\right)^2\right] \quad (3.12)$$

où $\sigma_c \in [\underline{\sigma}_c, \overline{\sigma}_c]$ et $c \in \{R, V, B\}$.

Ces deux versions des T2-FMOG peuvent être utilisées pour modéliser le fond. Cependant, T2-FMOG-UM est plus robuste que T2-FMOG-UV. En effet, cela est directement lié au fait que seulement les moyennes sont estimées et mises à jour correctement et que les variances quant à elles sont instables comme mis en evidence par l'étude approfondie effectuée par Greiffenhagen et al. [207]. Le système complet de modélisation du fond et de détection des objets mobiles par T2-FMOG comportent donc les étapes décrites dans les trois sous-sections suivantes : Initialisation de T2-FMOG, classification Fond/Objet et mise à jour du modèle de fond.

1- Initialisation de T2-FMOG

L'initialisation consiste à estimer les paramètres μ et σ et le facteur k_m ou k_ν. Zeng et al. [616] proposent de choisir des valeurs constantes pour k_m et k_ν qui dépendent d'une connaissance *a priori*. Dans l'application à la modélisation de fond, ces valeurs sont fixées en fonction de la vidéo. Pratiquement, $k_m = 2$ et $k_v = 0.9$ sont globalement des valeurs optimales comme le montre l'étude dans la section (3.4). L'estimation des paramètres comportent donc trois étapes :

— **Etape 1 :** Choisir K entre 3 et 5.

— **Etape 2 :** Estimer les paramètres des MOG par un algorithme EM.

— **Etape 3 :** Ajouter le facteur k_m ou k_ν aux MOG pour obtenir le T2-FMOG-UM ou le T2-FMOG-UV.

Une fois que l'initialisation est effectuée, une première détection peut être réalisée.

2- Classification Fond/Objet

En utilisant le rapport $r_i = \omega_i/\sigma_i$, nous ordonnons les K Gaussiennes comme dans Stauffer Grimson [490]. Les premières distributions normales qui excèdent un certain seuil T sont retenues pour représenter un fond multimodal selon :

$$B = argmin_b \left(\sum_{i=1}^{b} \omega_{i,t} > T \right) \tag{3.13}$$

Les autres distributions représentent la distribution des objets d'intérêt.

Quand une nouvelle image arrive au temps $t+1$, un test de correspondance est appliqué en chaque pixel pour déterminer son appartenance. Pour cela, nous maximisons la vraissemblance et ainsi nous nous intéressons uniquement à la longueur définie par les deux bornes :

$$H\left(X_t\right) = \left| ln\left(\underline{h}\left(X_t\right)\right) - ln\left(\overline{h}\left(X_t\right)\right) \right| \tag{3.14}$$

comme le montre la figure 3.4. Cette longueur décrit l'incertitude de l'entrée X_t au T2-EF.

Plus H est longue, plus l'incertitude est élevée. Ce cas est illustré par le gris foncé dans la figure 3.4. Par exemple, X_1 dévie beaucoup de la moyenne, donc cette donnée a non seulement un degré d'appartenance faible (H_l) mais aussi une longue H_1. les trois intervalles H, H_u et H_l mesurent l'incertitude de la donnée X.

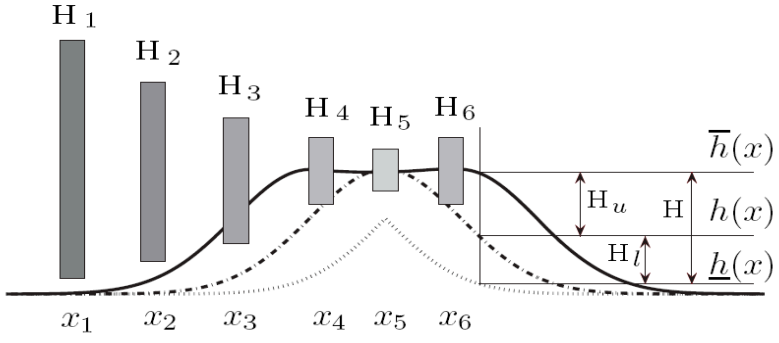


FIGURE 3.4 – La longueur $H\left(X\right) = \left| ln\left(\underline{h}\left(X\right)\right) - ln\left(\overline{h}\left(X\right)\right) \right|$ qui décrit l'incertitude de l'entrée X au T2-EF.

La densité de probabilité avec une incertitude autour de la moyenne est donnée par :

$$H\left(X_t\right) = \begin{cases} \frac{2k_m|X_t-\mu|}{\sigma}, & \text{si } X_t \leq \mu - k_m\sigma \text{ ou } X_t \geq \mu + k_m\sigma \\ \frac{|X_t-\mu|}{2\sigma^2} + \frac{k_m|X_t-\mu|}{\sigma} + \frac{k_m^2}{2}, & \text{si } \mu - k_m\sigma < X_t < \mu + k_m\sigma \end{cases} \tag{3.15}$$

La densité de probabilité avec une incertitude autour de la variance est donnée par :

$$H\left(X_t\right) = \left(\frac{1}{1/k_\nu^2 - k_\nu^2}\right) \frac{\left|X_t - \mu\right|^2}{2\sigma^2} \tag{3.16}$$

μ et σ sont la moyenne et la variance exactes (ou certaines).

Les deux équations 3.15 et 3.16 sont des fonctions croissantes du terme : $|X_t - \mu|$. Par exemple, soit k_m fixé, plus X_t dévie de μ, plus large sera $H\left(X_t\right)$ dans ces équations. cela permet de détecter la présence des valeurs aberrantes. Si une valeur aberrante X_t dévie du centre de la distribution, son $H\left(X_t\right)$ est très grande. Ainsi, un pixel est représenté par une Gaussienne si :

$$H\left(X_t\right) < k\sigma \tag{3.17}$$

où k est un seuil constant déterminé expérimentalement et généralement égal à 2.5.

Finalement, deux cas peuvent se présenter lors de la labellisation des Gaussiennes :

1^{er} Cas : Le pixel est représenté par l'une des Gaussiennes. Dans ce cas, si la Gaussienne a été identifiée en tant qu'une distribution représentant le fond, le pixel est classé comme "fond", sinon il est classé comme "objets mobiles".

$2^{ème}$ Cas : Le pixel n'est représenté par aucune des Gaussiennes. Dans ce cas, il est considéré comme pixel "objets mobiles".

Dans cette étape, les objets mobiles sont marqués à l'aide d'un masque binaire. Pour une nouvelle détection, les paramètres doivent être mis à jour à nouveau.

3- Mise à jour

La mise à jour de T2-FMOG s'effectue comme la version précise de Stauffer et Grimson [490], c'est-à-dire :

1er Cas : Le pixel est représenté par l'une des Gaussiennes. Pour cette Gaussienne, ses paramètres sont mis à jour selon les règles suivantes :

$$\omega_{i,t+1} = (1-\alpha)\,\omega_{i,t} + \alpha \tag{3.18}$$
$$\mu_{i,t+1} = (1-\rho)\,\mu_{i,t} + \rho X_{t+1} \tag{3.19}$$
$$\sigma^2_{i,t+1} = (1-\rho)\,\sigma^2_{i,t} + \rho\,(X_{t+1}-\mu_{i,t+1})\,.\,(X_{t+1}-\mu_{i,t+1})^T \tag{3.20}$$

où α est un coefficient d'apprentissage et $\rho = \alpha.\eta\,(X_{t+1}, \mu_i, \sum_i)$.

Pour les autres Gaussiennes, μ et $\sum$ gardent leur valeur, et c'est uniquement leur poid qui sera mis à jour avec la formule :

$$\omega_{j,t+1} = (1-\alpha)\,\omega_{j,t} \tag{3.21}$$

2ème Cas : Le pixel n'est représenté par aucune des Gaussiennes. Dans ce cas, la Gaussienne ayant le poids le moins important est remplacé par une nouvelle Gaussienne ayant les paramètres suivants :

$$\omega_{i,t+1} = Poids\ le\ plus\ petit\ des\ K\ Gausssiennes \tag{3.22}$$
$$\mu_{i,t+1} = X_{t+1} \tag{3.23}$$
$$\sigma^2_{i,t+1} = Variance\ la\ plus\ grande\ des\ K\ Gausssiennes \tag{3.24}$$
$$\tag{3.25}$$

Une fois les paramètres sont mis à jour, une nouvelle détection peut être réalisée et ainsi de suite.

3.4 Influence des coefficients de flous

Nous présentons dans cette partie l'influence des paramètre k_m et k_ν respectivement sur les algorithmes T2-FMOG-UM et T2-FMOG-UV.

Dans un premier temps, nous avons donc fait varier k_m dans l'intervalle $[0, 3]$. Pour les évaluations, on prend pour k_m les valeurs suivantes $\{0, 0.5, 1, 1.5, 2, 2.5, 3\}$. La figure 3.5 montre les résultats obtenus. Ensuite, nous nous intéressons au paramètre k_ν utilisé dans l'algorithme T2-FMOG-UM. Cette fois-ci, on fait varier k_ν dans l'intervalle $[0.3, 1[$ pour étudier son influence sur T2-FMOG-UV. Les valeurs de k_ν sont choisis dans l'ensemble

$\{0.3, 0.5, 0.7, 0.9\}$. La figure 3.7 montre les résultats obtenus avec T2-FMOG-UV pour les différentes valeurs de k_ν.

Dans la figure 3.5, on remarque que pour des valeurs $k_m \in \{0, 1\}$, aucune vrai positif (TP) n'a été détecté. En effet, quand $k_m = 0$, T2-FMOG-UM ne rejoint pas exactement les MOG car la fonction $H(X_t)$ (équations 3.15 et 3.16) utilisée dans le test de correspondance 3.17 (qui détermine à quelle Gaussienne correspond à un pixel) n'est pas le même test que celui adopté dans les MOG classiques.

D'après l'équation 3.7, nous pouvons constaté que les paramètres k_m et k_ν sont choisis de telle sorte que la zone du FOU soit contrôlée par rapport à la gestion de l'incertitude. Plus k_m est grand (ou plus k_ν est petit), plus la zone du FOU est grande, représentant ainsi une plus grande incertitude sur les données. Ceci est vu également dans la figure 3.5 qui montre les résultats de l'algorithme T2-FMOG-UM pour différentes valeurs de k_m. Plus on augmente k_m, plus la zone de FOU devient grande, intégrant ainsi une plus grande incertitude et offrant aussi une meilleure adaptation du modéle à son environnement. Dans les évaluations quantitatives réalisées et regroupées dans le tableau 3.5 et sous forme graphique dans la figure 3.6, nous avons choisi $k_m = 2$ car cette valeur est celle qui a offert le meilleur compromis entre qualité de détection et adaptation à un fond dynamique avec une erreur totale moyenne égale à 2657 sur les quatres images tests et une erreur totale moyenne de 1375 sur les 500 images de toute la séquence.

De même pour k_ν, dans la figure 3.7 qui montre les résultats de l'algorithme T2-FMOG-UV pour différentes valeurs de k_ν. Plus k_ν est grand, moins est prise en compte l'incertitude. Ce qui a pour conséquence de générer plus de fausses détections comme on peut le voir dans les résultats des images segmentées obtenues pour $k_\nu = 0.3$ et $k_\nu = 0.5$. Dans les évaluations quantitatives réalisées et regroupées dans le tableau 3.5 et sous forme graphique dans la figure 3.8, le meilleur réglage de la valeur k_ν dans l'algorithme T2-FMOG-UV est donné pour $k_\nu = 0.9$ avec une erreur totale moyenne égale à 2667 sur les quatres images tests et une erreur totale moyenne de 2624 sur les 500 images de toute la séquence.

3.5 Discussion

Dans cette Section 3, nous avons présenté une approche originale qui applique la théorie des ensembles flous de Type-2 à la modélisation de fond dynamique. Nous avons ainsi proposé deux versions de mélange de Gaussiennes floues : l'une qui gére l'incertitude associée à la moyenne (T2-FMOG-UM) et l'autre introduit le caractère flou à la variance (T2-FMOG-UV). Cette approche s'est montrée particulièrement adaptée aux fonds dynamiques comme le montre son évaluation dans la Section 6.1.

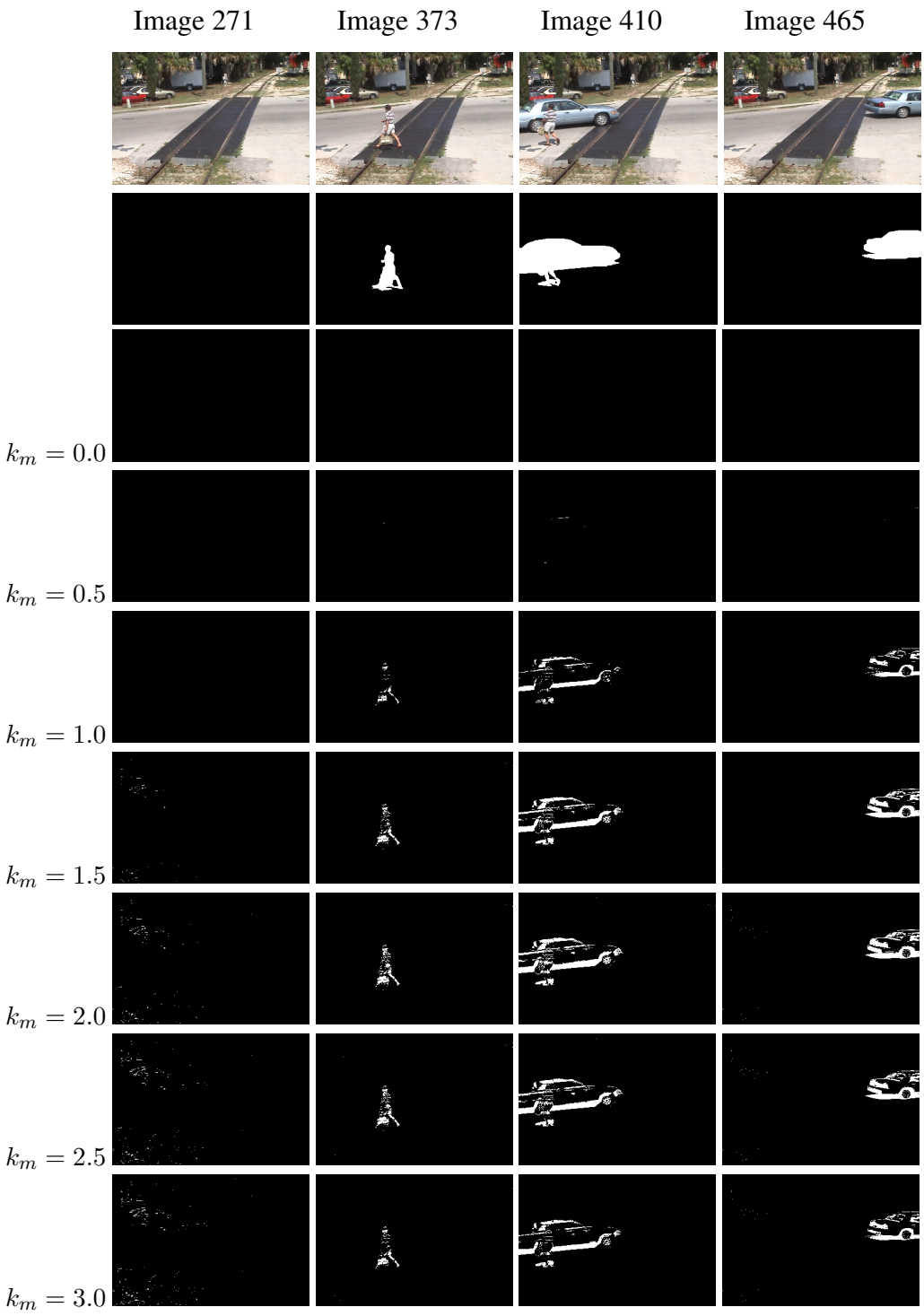


FIGURE 3.5 – Influence de k_m sur l'algorithme T2-FMOG-UM (Cas des MOG de Stauffer et Grimson [490]) : La première ligne montre des images originales de la séquence de Shah et al. [461]. La deuxième ligne présente les vérités terrain. De la troisième jusqu'à la neuvième ligne, sont présentées les images segmentées obtenues en utilisant l'algorithme T2-FMOG-UM pour des valeurs respectives de $k_m \in \{0, 0.5, 1, 1.5, 2, 2.5, 3\}$.

Méthode	Type d'erreurs	Image 271	Image 373	Image 410	Image 465	Total d'erreurs	Moyenne d'erreurs	Moyenne d'erreurs sur toute la séquence
T2-FMOG-UM	FN	0	1902	9237	4623			
$k_m = 0$	FP	0	0	0	0	15762	3940	1861
T2-FMOG-UM	FN	0	1902	9216	4621			
$k_m = 0.5$	FP	0	2	0	2	15743	3935	1860
T2-FMOG-UV	FN	0	1674	7399	3299			
$k_m = 1$	FP	0	60	111	16	12559	3139	1534
T2-FMOG-UV	FN	0	1496	6376	2660			
$k_m = 1.5$	FP	85	117	212	29	10975	2743	1400
T2-FMOG-UV	FN	0	1414	6043	2520			
$k_m = 2$	FP	203	153	252	46	**10631**	**2657**	**1375**
T2-FMOG-UV	FN	0	1437	6292	2698			
$k_m = 2.5$	FP	300	143	222	46	11138	2784	1531
T2-FMOG-UV	FN	0	1550	6912	3145			
$k_m = 3$	FP	362	94	153	34	12250	3062	1781

TABLE 3.2 – Influence de k_m sur T2-FMOG-UM : Calculs des faux positifs (FP) et des faux négatifs (FN). La valeur qui minimise le nombre de total d'erreurs est obtenue pour $k_m = 2$. Les deux dernières colonnes indiquent respectivement le nombre moyen d'erreurs sur les quatres images tests et sur les 500 images de toute la séquence.

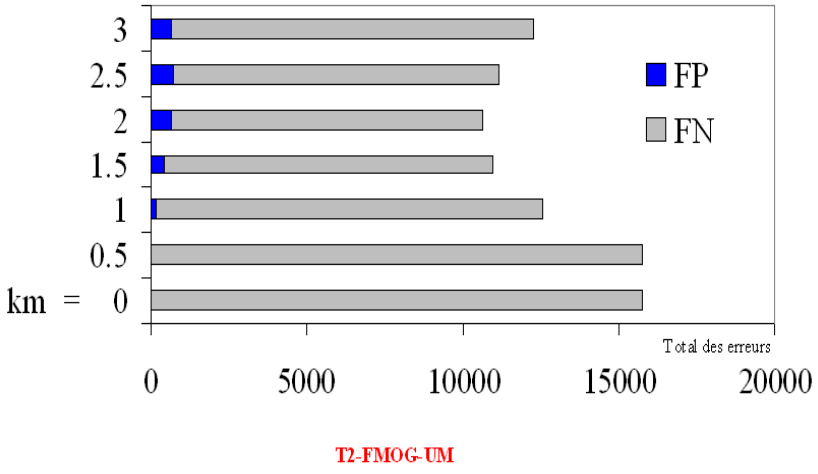


FIGURE 3.6 – Représentation graphique de l'influence de k_m sur l'algorithme T2-FMOG-UM (version des MOG de Stauffer et Grimson [490]) ainsi que de la répartition en termes de faux positifs (FP) et faux négatifs (FN) pour des valeurs de k_m égale à 0, 0.5, 1, 1.5, 2, 2.5, 3 respectivement. La valeur minimale en terme de total d'erreurs est obtenue pour $k_m = 2$.

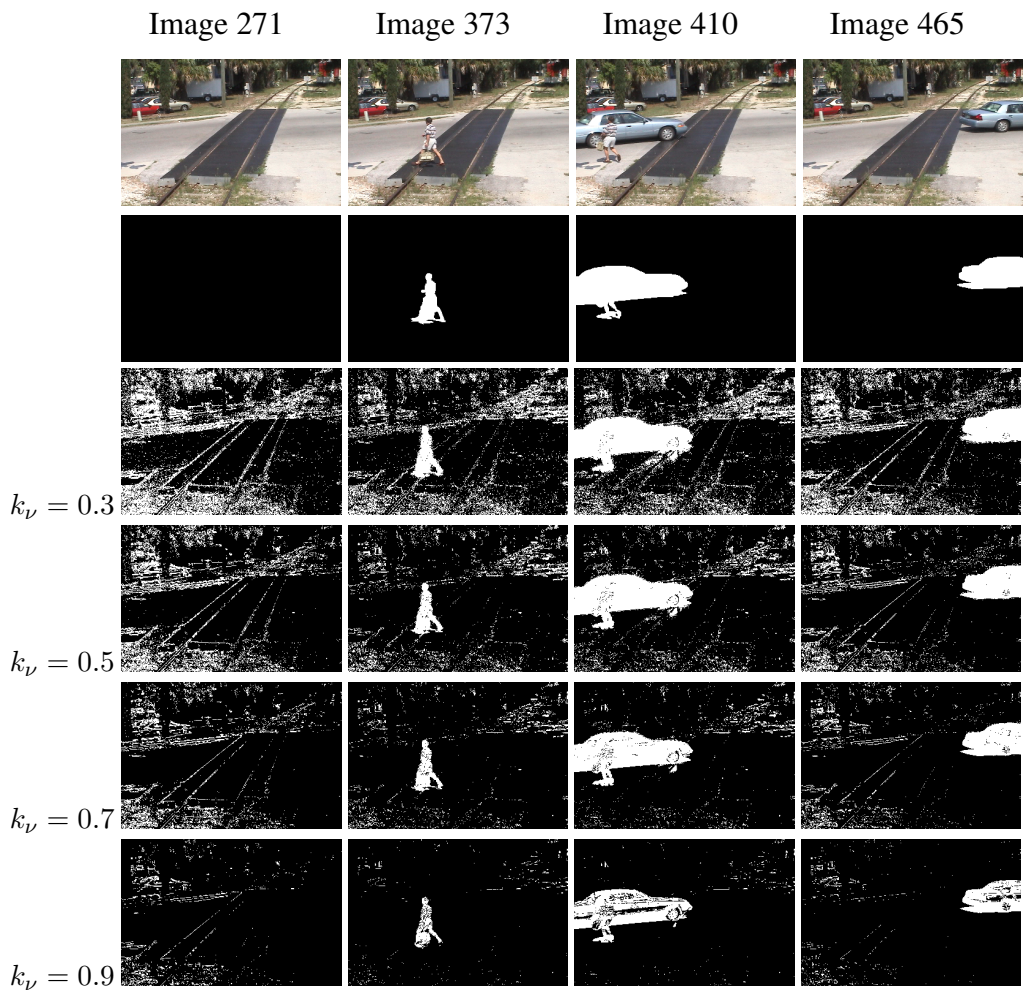


FIGURE 3.7 – Influence de k_ν sur T2-FMOG-UV (Cas des MOG de Stauffer et Grimson [490]) : La première ligne montre des images originales de la séquence de Shah et al. [461]. La deuxième ligne présente les vérités terrain. De la troisième jusqu'à la sixième ligne, sont présentées les images segmentées obtenues en utilisant l'algorithme T2-FMOG-UV pour des valeurs respectives de $k_\nu \in \{0.3, 0.5, 0.7, 0.9\}$.

Méthode	Type d'erreurs	Image 271	Image 373	Image 410	Image 465	Total d'erreurs	Moyenne d'erreurs	Moyenne d'erreurs sur toute la séquence
T2-FMOG-UM	FN	0	430	317	331			
$k_\nu = 0.3$	FP	22676	15552	14586	14951	68843	17290	19392
T2-FMOG-UM	FN	0	559	491	422			
$k_\nu = 0.5$	FP	13330	7303	6898	7487	36490	9122	10444
T2-FMOG-UV	FN	0	668	738	524			
$k_\nu = 0.7$	FP	7957	3623	3367	3817	20694	5173	6022
T2-FMOG-UV	FN	0	957	2217	1069			
$k_\nu = 0.9$	FP	3069	1081	1119	1158	**10670**	**2667**	**2624**

TABLE 3.3 – Influence de k_ν sur l'algorithme T2-FMOG-UV : Calculs des faux positifs (FP) et des faux négatifs (FN). La valeur qui minimise le nombre de total d'erreurs est obtenue pour $k_\nu = 0.9$. Les deux dernières colonnes indiquent respectivement le nombre moyen d'erreurs sur les quatres images tests et sur les 500 images de toute la séquence.

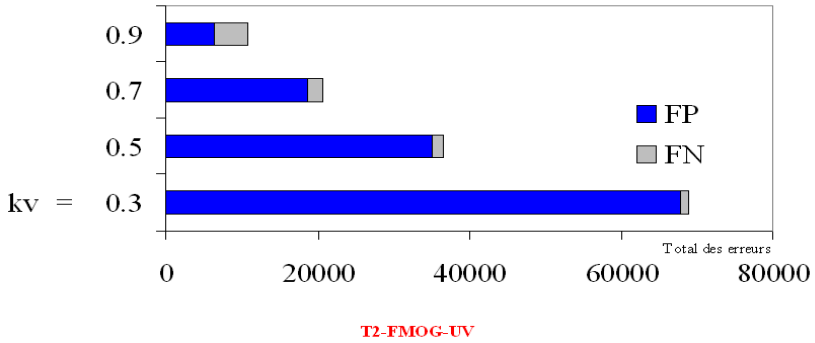


FIGURE 3.8 – Représentation graphique de l'influence de k_ν sur l'algorithme T2-FMOG-UV (version des MOG de Stauffer et Grimson [490]) ainsi que de la répartition en termes de faux positifs (FP) et faux négatifs (FN) pour des valeurs de k_ν égale à 0.3, 0.5, 0.7, 0.9 respectivement. La valeur minimale en terme de total d'erreurs est obtenue pour $k_\nu = 0.9$.

4 Détection d'objets mobiles par l'intégrale de Choquet

L'identification des objets mobiles dans des séquences vidéo peut se faire en se basant sur certaines caractéristiques comme l'intensité (mesuré au niveau de gris ou en espace couleur), la texture, le contour, les vecteurs mouvement, la profondeur ou la disparité. Chacune de ces caractéristiques a des propriétés discriminatives différentes.

Comme nous avons vu précédemment dans la Section 3.6 du Chapitre 2, lors de l'acquisition des séquences vidéo, la présence de certaines situations critiques, qui peuvent apparaître dans la scène, peut perturber la segmentation. Certaines de ces situations sont gérables par le choix de la caractéristique utilisée. La caractéristique la plus utilisée est l'intensité ou la couleur car c'est la première information fournie par les capteurs. Dans certains cas, la couleur ne permet pas une bonne discrimination des objets mobiles par rapport au fond, notamment en présence des variations lumineuses, du camouflage et des ombres. Pour pallier cela, il faut combiner d'autres caractéristiques avec la couleur comme les contours, la texture, le flot optique ou la stéréovision. Cette combinaison permet d'affiner la segmentation en élargissant le champ des informations fournies sur l'évolution du pixel. Chacune des caractéristiques a ses propres avantages mais certaines sont coûteuses en termes de calculs et de complexité. Par exemple, la stéréovision opère bien en présence du camouflage mais deux caméras sont nécessaires. Les contours gèrent les changements de luminosité locaux et les trous causés par la reprise du mouvement d'un objet d'intérêt dans la scène. Quant aux textures, elles sont appropriées au changement de luminosité et aux ombres qui représentent un challenge important dans notre étude. L'idée ici est donc de combiner différentes caractéristiques mais plusieurs problèmes se posent : Quelles caractéristiques utiliser ? Quels sont les outils les mieux adaptés pour les combiner ? Quelle pondération donner à chacune d'elles ? Dans le cas de l'intensité mesurée à partir des composantes d'un espace couleur, il y a eu, dans la littérature, une proposition d'opérateur où les auteurs ont établi la détection, indépendemment, dans chacune des composantes couleurs puis les images segmentées ont été combinées à l'aide de l'opérateur classique « OU ». Le problème avec un tel opérateur est que la qualité de la détection risque de se dégrader très rapidement, du fait qu'un faux positif obtenu dans l'une des images segmentées peut générer un faux positif dans le résultat final.

Afin de réaliser la fusion des caractéristiques, nous avons choisi l'intégrale floue de Choquet. Le choix de cet opérateur de fusion floue a été motivé par quatre raisons principales. Tout d'abord, les opérateurs flous (de Sugeno ou de Choquet) permettent de prendre en compte, à la fois, l'incertitude et l'imprécision associées aux données réelles. Deuxièmement, ils présentent une grande flexibilité, grâce aux opérateurs de combinaison. Troisièmement, leur mise en oeuvre passe par des opérations relativement simples et rapides, ce qui les rend particulièrement adaptés pour les applications pratiques. Et quatrièmement, vu que les données sont de nature cardinale, l'intégrale de Choquet est plus adaptée que l'intégrale de Sugeno pour cette application. Nous proposons également une règle de mise à jour adaptative du système de la détection basée sur le résultat obtenu par l'intégrale de Choquet. L'analyse

des résultats en segmentation obtenus au travers des intégrales de Sugeno et de Choquet permet de conclure que l'opérateur de fusion floue de Choquet est le mieux adapté dans le contexte de la détection d'objets mobiles.

4.1 Opérateurs d'agrégation

L'être humain a souvent inspiré la science par son mécanisme de raisonnement dans l'élaboration d'une décision. Il est capable d'évaluer, comparer, trier, synthétiser et décider même dans des conditions où l'on est amené à prendre en compte différentes informations ou critères conflictuels parfois au sujet d'un même ensemble d'objets. Cette capacité a intéressé certains chercheurs pour développer des systèmes approchant au mieux ce mécanisme sophistiqué d'aide à la décision que dispose l'être humain. On parle donc d'agrégation de données ou de fusion de données. Le mot agrégation réfère à un processus de combinaison des valeurs (numériques ou linguistiques) qui permet de synthétiser n valeurs en une seule, de sorte que toutes les valeurs agrégées sont prises en compte dans le résultat final.
Dans les problèmes d'aide à la décision, les valeurs à agréger sont typiquement des préférences ou des degrés de satisfaction d'un ensemble de solutions potentielles ou « alternatives » par rapport à des critères. Un degré de préference décrit une certaine relation de préference entre deux alternatives données, alors qu'un degré de satisfaction exprime la satisfaction d'une alternative donnée par rapport à un critère donné. D'une façon plus générale, et selon l'application, ces valeurs peuvent être des degrés de confiance en une alternative donnée pour qu'elle soit vraie, des degrés de similarité, des avis des experts, etc...

Les valeurs peuvent être définies sur une échelle numérique ou linguistique. Que ce soit l'une ou l'autre, l'opération de l'agrégation exige une certaine relation de cohérence ou une structure commune à toutes les valeurs. L'échelle linguistique est typiquement utilisée dans la théorie des ensembles flous, par exemple, $\{petit, moyen, grand\}$. Une échelle numérique part sur la base des nombres numériques comme son nom l'indique, et peut être *ordinal* ou *cardinal*. Le caractère *ordinal* intervient pour décrire l'unique relation qui existe entre les valeurs, celle de l'ordre. L'exemple typique d'une échelle ordinale est l'attribution d'une note de 1 à 5 pour évaluer le niveau d'un élève ; 1 correspond à « très faible », 2 correspond à « faible », 3 correspond à « moyen », 4 correspond à « bien » et 5 correspond à « excellent ». Dans ce schéma, on voit bien qu'aucune relation de différence ou de distance ne peut être définit entre les nombres. En d'autre terme, la distance entre 3 et 4 n'est pas égale à la distance entre 4 et 5. Par contre, sur une échelle cardinale, les nombres ont une signification concrète, c-à-d, pour des valeurs de 1 à 5, la différence entre 3 et 4 est égale à la différence entre 4 et 5.

Certains problèmes d'aide à la décision sont de nature multicritères. Les systèmes de nature multicritères doivent être capables d'évaluer la performance des différentes alternatives selon chacun des points de vues jugés pertinents pour le problème. Ces systèmes traitent des problèmes de choix pour sélectionner les meilleures alternatives, des problèmes de rangement pour ordonner des alternatives selon leurs qualités relatives, ou des problèmes de tri

pour affecter des alternatives à des catégories. Pour décider si une alternative est préferable à une autre, il s'agit donc, en général, d'un problème de comparaison d'une performance ou d'une similarité entre ces deux alternatives. Cette comparaison est réalisé en évaluant les performances des deux alternatives selon chacun des points de vues ou critères. Lorsque les quantités à comparer sont exprimées par des nombres réels dans l'intervalle $[0, 1]$, cette relation peut s'interpréter comme la fonction d'appartenance caractéristique des ensembles flous.

Une fois que les valeurs sont définies, on peut les agréger. Cette opération peut être réalisée de différentes façons selon la nature des valeurs à agréger et le type d'échelle utilisée. Dans la littérature, une multitude d'opérateurs d'agrégation a été proposée pour aggréger ces informations. Ces opérateurs ont été appliqués dans différents domaines comme la reconnaissance des formes [26], l'évaluation des candidats, le dépouillement des enquêtes ou des sondages. Ici, nous avons appliqué l'opérateur d'agrégation flou de Choquet pour la détection d'objets mobiles. Ces opérateurs d'agrégation respectent certaines propriétés dont les plus utiles sont explicitées dans la sous section suivante. Nous verrons ensuite certains opérateurs classiques et surtout les opérateurs flous de Sugeno et de Choquet. Une étude plus approfondie sur les opérateurs d'agrégation est presentée dans [205].

4.2 Intégrales floues : Sugeno-Choquet

Les intégrales floues représentent en quelque sorte la valeur moyenne d'une fonction. Dans le cas discret, elles sont considérées comme des opérateurs de type moyenne. Elles ont été introduites par Choquet [123] et par Sugeno [495][496] et leur définition respective est liée à la définition d'une mesure floue. Pour pouvoir les décrire dans le cadre discret, nous commençons tout d'abord par donner la définition d'une mesure floue. On désigne par $\mathbb{P}(S)$ l'ensemble des parties de S.

Definition 1 - ***Mesure floue*** *:*
Une mesure floue notée μ définie sur un ensemble fini de critères S est une fonction $\mu : \mathbb{P}(S) \rightarrow [0, 1]$ qui vérifie les axiomes suivants :

$$\mu(\emptyset) = 0 \quad , \quad \mu(S) = 1 \tag{3.26}$$

$$A \subset B \quad \Rightarrow \quad \mu(A) \leq \mu(B), \quad pour\ A,\ B \in \mathbb{P}(S) \tag{3.27}$$

où $\mu(A)$ et $\mu(A)$ représentent l'importance ou le poids de la coalition des sources A et B dans le processus d'agrégation.

Pour cette mesure, en plus des poids associés à chacun des critères (les singletons) comme nous avons vu avec les ω_i, nous avons besoin aussi de définir des poids associés à chaque combinaison de critères faisant partie de $\mathbb{P}(S)$, ce qui fait $2^n - 2$ coefficients en total.
Pour tout A et B dans $\mathbb{P}(S)$, une mesure floue est dite :

— **Additive** si :

$$\mu(A \cup B) = h(A) + \mu(A), A \cap B = \emptyset \tag{3.28}$$

— **Suradditive** si :

$$\mu(A \cup B) \geq \mu(A) + \mu(A), A \cap B = \emptyset \tag{3.29}$$

— **Sousadditive** si :

$$\mu(A \cup B) \leq \mu(A) + \mu(A), A \cap B = \emptyset \tag{3.30}$$

— **Cardinale ou Symétrique** si elle ne dépend que de la cardinalité des ensembles, i.e. :

$$\mu(A) = \mu(B), |A| = |B| \tag{3.31}$$

Pour une mesure additive, il suffit de définir les n coefficients $\mu(\{1\})$, $\mu(\{2\})$,...,$\mu(\{n\})$ pour pouvoir définir entièrement la mesure pour toutes les coalitions.
Une fois la mesure floue définie, on peut introduire le concept des intégrales floues discrètes.

Definition 2 - ***Intégrale de Sugeno :***
Soit μ une mesure floue définie sur N, l'intégrale de Sugeno [495] de $a_1, \cdots, a_n$ dans $[0,1]$ par rapport à μ est définie par :

$$S_\mu(a_1, \cdots, a_n) = Max\left(Min\left(a_{(i)}, \mu\left(A_{(i)}\right)\right)\right) \tag{3.32}$$

où $A_{(i)} = \{(i), \cdots, (n)\}$ après avoir ordonné les $a_{(i)}$ selon la relation $a_{(1)} \leq \cdots \leq a_{(n)}$

Definition 3 - ***Intégrale de Choquet :***
L'intégrale de Choquet de $a_1, \cdots, a_n$ dans $[0,1]$ par rapport à μ est définie par :

$$C_\mu(a_1, \cdots, a_n) = \sum_{i=1}^{n} a_{(i)}\left(\mu\left(A_{(i)}\right) - \mu\left(A_{(i+1)}\right)\right) \tag{3.33}$$

avec les même notations définies ci-dessus, $a_{(0)} = 0$ et $\mu\left(A_{(n+1)}\right) = 0$.

Discussion autour des intégrales de Sugeno et de Choquet : Les intégrales de Sugeno et de Choquet partagent quelques propriétés communes comme celles de l'idempotence, la continuité, la monotonie, le compromis et bien d'autres. Toutes les deux calculent en quelque sorte la moyenne et contiennent les statistiques d'ordre (le « $\wedge$ », le « $\vee$ » et la « médiane »). L'intégrale de Sugeno permet d'établir des opérations de comparaison (Min et Max) alors que celle de Choquet permet les opérations linéaires usuelles ($\sum$ et $\prod$). Egalement, l'intégrale de Choquet est de plus stable pour le changement d'échelle linéaire ($\sum$ et $\prod$). Dans ce sens, elle est vue comme adaptée pour une agrégation cardinale alors que l'intégrale de Sugeno est adaptée pour une agrégation ordinale [497][377][500]. En effet, sur une échelle cardinale, la distance entre les valeurs à agréger peut être bien définie du fait qu'elles ont un vrai sens. Pour illustrer ce cas, on reprend le même exemple cité dans l'introduction ; la différence entre 3 et 4 est égale à la différence entre 4 et 5 ce qui n'est pas le cas si ces valeurs étaient définies sur une échelle ordinale, car cette dernière permet uniquement d'établir des opérations

de comparaison ($\wedge$ et $\vee$) et la notion de différence n'a plus le même sens du fait que le 3 correspond à une valeur linguistique et non au nombre 3. Ces intégrales ont été appliquées à des problèmes concrets comme la fusion des sources et fournissent des résultats satisfaisants.

4.3 Application à la détection d'objets mobiles

La détection d'objets mobiles est basée sur la comparaison entre l'image courante et l'image du fond pour identifier les zones qui ont changé. En général, une simple soustraction entre ces deux images peut être effectuée pour détecter les régions qui correspondent aux objets d'intérêts. Une autre manière d'établir cette comparaison consiste à définir une mesure de similarité entre chaque pixel dans l'image courante et son homologue dans l'image du fond. Cette similarité est mesurée par le rapport de ces deux valeurs. L'idée ici consiste à dire que les pixels du fond sont similaires dans les deux images alors que ceux qui correspondent aux objets mobiles ne le sont pas. Mais ce qui peut fournir une information sur l'évolution du pixel sont les caractéristiques. Comme on l'a vu précédemment, la caractéristique la plus utilisée est la couleur ou l'intensité. Mais la combinaison d'autres caractéristiques peut fournir des informations supplémentaires grâce à leur propriétes différentes. De cette façon, on peut ainsi calculer plusieurs degrés ou mesures de similarités pour différents types de caractéristiques. Une fois que les valeurs de similarités sont calculées, il est possible de les agréger avec un opérateur d'agrégation.

Dans le contexte de la détection d'objets mobiles, on peut appliquer les integrales floues pour fusionner différents critères ou caractéristiques. La mesure μ peut être vue comme le facteur qui décrit la pertinence des sources d'informations et a désigne les valeurs du pixel par rapport à chacun des critères ou sources. Un pixel est donc évalué selon différents critères fournissant des informations sur l'état du pixel selon s'il correspond à un fond ou a un objet d'intérêt. Les valeurs à agréger que prennent les critères sont les similarités mesurées entre chaque pixel dans l'image courante et son homologue dans l'image du fond. Le poids d'un critère est d'autant plus important dans la décision de l'état du pixel que l'information qu'il représente est importante.

Dans ce qui suit, nous explicitons le choix des caractéristiques, le calcul des quantités obtenues à partir de ces caractéristiques qui sont à agréger et comment les agréger avec une intégrale de Choquet. Après expérimentations, nous montrerons dans ce qui suit que l'intégrale de Choquet est mieux adaptée à ce type de valeurs que l'intégrale de Sugeno, vu qu'elle est définie sur une échelle cardinale.

Choix des caractéristiques : Plusieurs caractéristiques sont exploitées pour étudier l'évolution du pixel comme la couleur, les contours [261][260][329], la disparité [202], la profondeur [239], le flot optique [511] et la texture (LBP) [245]. Nous avons vu dans l'introduction de ce chapitre que les propriétés qui les caractérisent sont différentes. Dans notre application, les problèmes de changement lumineux et d'ombres constituent un challenge important. Les caractéristiques de la couleur sont faciles à manipuler mais dans certains cas, mais seules elles

ne fournissent pas une bonne discrimination entre les objets mobiles et le fond, notamment en présence des variations lumineuses, du camouflage et des ombres. Il existe plusieurs espaces couleur comme l'espace RGB, HSV, Ohta, YCrCb à partir desquels on peut obtenir les valeurs d'intensité pour chaque pixel dans chacune de leurs composantes. Mais la question qui se pose est la suivante : Lequel d'entre eux s'adapte mieux au type de séquences liées à notre application ? Il est connu que certains espaces couleur offrent une meilleure discrimination par rapport aux ombres dues à leurs propriétés respectives. Pour cela, nous avons testé les composantes de ces différentes espaces couleur pour en choisir les plus adaptées. D'autre part, la texture semble être robuste aux changements lumineux locaux et aux ombres. Par ailleurs, les autres caractéristiques bien qu'elles aient beaucoup d'avantages, elles sont toutes liées à une nécessité de les calculer et de les stocker, générant ainsi un probléme en termes de mémoire ou de calcul. Pour cette raison, notre choix s'est posé finalement sur les couleurs et/ou les textures car ces deux caractéristiques peuvent répondre à notre challenge en permettant de réaliser un traitement relativement rapide.

Couleurs : Le choix de l'espace couleur est un facteur important dans l'extraction de l'information portée par la couleur. Comme vu précédemment, les composantes couleurs de l'espace RGB sont les plus utilisées dans la détection d'objets mobiles car ce sont les premieres informations fournies par les capteurs. L'inconvénient de cet espace est que ses trois composantes R, V et B sont corrélées, ce qui augmente la sensibilité aux changements de luminosité. Par exemple, si un pixel du fond est couvert par une ombre, les valeurs des trois composantes sont affectées en ce pixel, car les composantes de luminance et de chrominance ne sont pas séparées. Dans la littérature, plusieurs études [293][431] ont comparé les différentes espaces couleur et ont mis en évidence trois espaces couleur en particulier : Ohta, HSV et YCrCb. Ces espaces permettent de séparer les composantes de luminance et de chrominance.

Caractéristique LBP : Le modèle de texture LBP (Local Binary Pattern) a été dévelopé par Heikkila and Pietikäinen [245] pour la caractérisation de la texture. LBP est invariant aux changements "monotonic" en niveau de gris, ce qui augmente sa robustesse fâce aux variations lumineuses. Cet opérateur procéde par "bloc" autour d'un pixel donné. Il donne un label au pixel par seuillage de sa valeur par rapport aux huit valeurs des pixels qui l'entourent.

$$LBP\,(x,y) = \sum_{i=0}^{N-1} s\,(g_i - g)\,2^i \tag{3.34}$$

Où g correspond à la valeur du pixel central (x, y) en niveau de gris et g_i celles des N

pixels voisins. La fonction s est définie par :

$$s(x) = \begin{cases} 1 & si \quad x \geq 0 \\ 0 & si \quad x < 0 \end{cases} \tag{3.35}$$

L'opérateur original LBP utilise dans un voisinage de 3×3 autour d'un pixel. La figure 3.9 illustre un calcul de LBP. On peut remarquer que la valeur du code LBP varie entre 0 et 255 comme la valeur d'intensité.
Une fois les codes LBP calculés pour chaque bloc dans l'image courante et son homologue dans l'image du fond, on peut mesurer une similarité correspondante $S^T(x, y)$ du pixel central à partir des codes $I^T(x, y)$ et $B^T(x, y)$ des mêmes pixels calculés dans l'image courante et celle du fond.

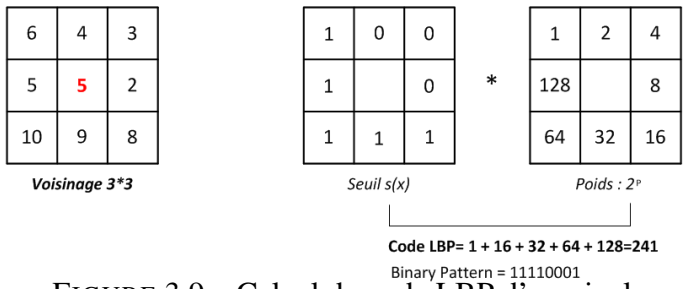


FIGURE 3.9 – Calcul du code LBP d'un pixel central.

Quantités à agréger : Que ce soit les couleurs ou les textures que l'on veuille agréger, la détection d'objet se fait, pour un pixel donné, par comparaison de sa valeur dans l'image courante par rapport à son homologue dans l'image du fond. En principe, les pixels du fond doivent être similaires alors que ceux qui correspondent aux objets mobiles subissent une différence. Mais ces valeurs ne peuvent pas être utilisées directement dans l'agrégation par une intégrale floue car les mesures utilisées dans ces intégrales doivent être comprises dans l'intervalle $[0, 1]$. On peut donc étudier la similarité de deux pixels homologues situés dans l'image courante et celle de référence en faisant le rapport de leurs deux valeurs qui est toujours compris entre 0 et 1. Si ce rapport est proche de 1 cela veut dire que les deux pixels sont similaires (il s'agit donc d'un pixel du fond) et s'il est égal à 0, la similarité est très faible (il est donc fortement possible qu'il s'agisse d'un pixel d'objet d'intérêt.).
On désigne par C, la valeur de la caractéristique qui peut être l'une des composantes couleurs d'un espace donné ou le code LBP (noté T) attribué à un pixel. Donc $C \in \mathbb{C}$ tel que $= \{R, G, B, H, S, V, Y, Cr, Cb, Ohta(1), Ohta(2), T\}$.
Quel que soit $C \in \mathbb{C}$, la mesure de similarité notée S pour une caractéristique donnée C est

calculée de la façon suivante :

$$S^C(x,y) = \begin{cases} \frac{I^C(x,y)}{B^C(x,y)} & si \quad I^C(x,y) < B^C(x,y) \\ 1 & si \quad I^C(x,y) = B^C(x,y) \\ \frac{B^C(x,y)}{I^C(x,y)} & si \quad I^C(x,y) > B^C(x,y) \end{cases} \tag{3.36}$$

où $I^C(x,y)$ et $B^C(x,y)$ sont respectivement les valeurs du pixel (x,y) de la caractéristique C dans l'image courante I et l'image du fond B. Notons que $S^T(x,y)$ est proche de 1 si $I^C(x,y)$ et $B^C(x,y)$ sont très similaires et proche de 0 dans le cas contraire.

4.4 Détection d'objets mobiles par l'utilisation de l'intégrale de Choquet

Ici, nous avons retenu trois critères soit trois caractéristiques, et nous allons déterminer l'état des pixels à partir de l'information portée par chacun d'eux.

Soient $X = \{1,2,3\}$ l'ensemble des critères. Nous associons à chacun une mesure floue telle que : $\mu(1) = \mu(\{1\})$, $\mu(2) = \mu(\{2\})$ et $\mu(3) = \mu(\{3\})$ de sorte que $\mu(i)$ est d'autant plus grande que l'influence du critère i dans la décision finale est grande.

Pour calculer les mesures floues pour toutes les coalitions possibles des critères i.e, $\{1,2\}$, $\{1,3\}$, $\{2,3\}$ et $\{1,2,3\}$, nous avons utilisé une version opérationnelle proposée par Sugeno et qui s'appelle λ-*mesure floue*. Pour simplifier les notations cette mesure est appelée μ_λ-*mesure floue*, où λ est un paramètre de la mesure floue qui décrit une interaction entre les valeurs des critères qui sont combinées. Sa valeur est déterminée grâce aux conditions aux limites, i.e. $\mu(X) = \mu(\{1,2,3\}) = 1$.
Pour un sous ensemble de X tel que $K \subset X$, les valeurs des densités floues sont calculées par l'expression suivante :

$$\mu_\lambda(K) = \frac{1}{\lambda}\left[\prod_{i \in K}(1 + \lambda\mu_\lambda(i)) - 1\right] \tag{3.37}$$

Une fois que les densités floues de tous les sous ensembles de X sont calculées, la valeur de l'intégrale de Choquet est calculée et la classification des pixels est obtenue selon la règle suivante :

$$si \quad C_\mu(x,y) < T \quad alors \quad (x,y) \in Objets\ Mobiles \tag{3.38}$$

où T est un seuil constant déterminé experimentalement. Une amélioration possible serait de faire un seuillage adaptif flou où T serait différent à chaque pixel, c'est-à-dire qu'on aurait $T(x,y)$.

4.5 Discussion

Dans cette Section 4, nous avons présenté une approche originale pour la détection d'objets mobiles en utilisant la caractéristique couleur et une caractéristique texture (LBP). Les résultats obtenus par ces deux caractéristiques sont ensuite fusionnés par l'integrale floue de Choquet permettant ainsi à la méthode développée d'être plus robuste aux variations lumineuses comme le montre son évaluation dans la Section 6.2.

5 Règle de mise à jour adaptative

Une fois que la première détection est établie, le système doit, naturellement, être capable de mettre à jour l'image du fond afin de fournir une nouvelle image integrant les changements récents qui ont apparu dans la scène. Ceci a l'avantage d'augmenter la robustesse du modèle et par conséquent, son adaptation par rapport à son environnement. Certains auteurs ont adopté une règle de mise à jour non sélective. Celle-ci met à jour tous les pixels de l'image selon la même règle :

$$B_{t+1}(x,y) = (1-\alpha)\, B_t(x,y) + \alpha I_t(x,y) \tag{3.39}$$

où α est un coefficient d'apprentissage compris entre 0 et 1 qui contrôle la vitesse d'intégration de la nouvelle information dans la mise à jour de l'image de référence. L'inconvénient d'une telle règle réside dans le fait qu'un pixel détecté comme objet d'intérêt contribue dans la mise à jour de l'image de référence ce qui risque de la détériorer.

Pour pallier cela, d'autres chercheurs ont définis un mode de mise à jour sélectif basé sur deux règles distinctes :

$$\begin{aligned} B_{t+1}(x,y) &= (1-\alpha)\, B_t(x,y) + \alpha I_t(x,y) && (3.40)\\ &\quad if\ (x,y)\ is\ background \\ B_{t+1}(x,y) &= (1-\beta)\, B_t(x,y) + \beta I_t(x,y) && (3.41)\\ &\quad if\ (x,y)\ is\ foreground \end{aligned}$$

(3.42)

La première est chargée de mettre à jour tous les pixels détectés comme pixels du fond selon la même règle et la deuxième traite uniquement les pixels étant détectés comme pixels d'objets d'intérêts. L'idée est de mettre à jour très rapidement les pixels du fond et d'adapter le modèle plus lentement avec les pixels classés comme objets mobiles. Pour cette raison, on prend, $\beta << \alpha$. En général, on prend $\beta = 0$. L'equation 3.41 devient alors :

$$B_{t+1}(x,y) = B_t(x,y) \tag{3.43}$$

Mais ce qui régie ces deux règles selectives est une décision binaire selon si le pixel a été détecté comme fond ou comme objet d'intérêt. Il en resulte qu'un pixel mal détecté comme fond met à jour l'image de référence au même degré qu'un vrai pixel du fond. Ceci peut entraîner une mauvaise mise à jour de l'image de référence et, par conséquent, une mauvaise classification de pixels lors de la prochaine détection.

D'ou notre idée de définir différents degrés d'appartenance des pixels en fond ou objet d'intérêt qui permettent d'intégrer la nouvelle information au fond d'une façon adaptative. La règle de mise à jour que nous proposons se sert de la valeur de l'intégrale de Choquet obtenue lors de l'étape de la détection. Cette valeur est comprise entre 0 et 1 et a l'avantage de donner des poids différents aux pixels classés comme fond et comme objets d'intérêt. Elle est d'autant plus élevée que le pixel est considéré comme fond et inversement. L'idée est de prendre en compte le degré de la contribution de chaque pixel dans la mise à jour de l'image de référence selon sa part d'implication dans la décision de son état (Fond/Objet d'intérêt). On introduit la règle adaptative dans la formule suivante :

$$\begin{aligned} B_{t+1}(x,y) &= \mu_F B_t(x,y) + \\ &\quad \mu_B\left((1-\alpha) B_t(x,y) + \alpha I_t(x,y)\right) \end{aligned} \tag{3.44}$$

où $\mu_F = 1 - \mu_B$.

μ_F et μ_B sont respectivement les degrés d'appartenance du pixel (x,y) aux classes fond/objet d'intérêt. μ_B est une fonction de la valeur de l'intégrale de Choquet $C_\mu(x,y)$ obtenue à l'issue de l'étape de la détection telle que : $\mu_B = 1$ pour $Max\left(C_\mu(x,y)\right)$ et $\mu_B = 0$ pour $Min\left(C_\mu(x,y)\right)$.

En se basant sur cette règle adaptative, nous assurons que tous les pixels de l'image sont mis à jour suivant leur degré d'appartenance. Par conséquent, un pixel mal classé comme fond ne sera pas mis à jour au même degré qu'un vrai pixel du fond comme systématiquement dans le mode selectif. Cette règle est une généralisation du mode sélectif de mise à jour. En effet, si le pixel est classé comme fond avec une valeur de l'intégrale de Choquet égale à 1, on retrouve l'equation 3.41, et si le pixel est classé comme objet d'intérêt avec une valeur de l'intégrale de Choquet égale à 0, l'équation 3.44 est équivalente l'équation 3.43.

Dans cette Section 5, nous avons présenté une méthode floue de mise à jour adaptative de l'image du fond à partir des degrés d'appartenance du pixel à la classe "fond" et à la classe "objet mobiles". Cette approche s'est montrée particulièrement adaptée à des fonds dynamiques comme le montre son évaluation dans la Section 6.3. En perspective, il serait pertinent que le coefficient d'apprentissage soit déterminé par une fonction plus évoluée des degrés d'appartenance pour être plus adaptatif.

6 Expérimentations et évaluations

Pour tester les trois contributions de ce chapitre, nous avons utilisé des bases de données tests disponibles sur le web à l'époque comme Wallflower [521], I2R [311] et PETS. Ces différentes bases de données vidéo ont permis de tester et d'évaluer les algorithmes qualitativement ou quantitativement dans différents contextes selon la disponibilité des vérites terrain des séquences sur une partie ou sur toutes les séquences.

Nous avons donc décidé, dans un premier temps, d'évaluer qualitativement les algorithmes sur des séquences filmées dans des environnements différents (intérieurs et extérieurs, scène visible et infrarouge). Cela a permis d'avoir une première idée sur le comportement des algorithmes proposés dans différentes situations critiques. Ensuite, pour les séquences fournies avec leurs vérités terrain, nous avons évalué la méthode proposée quantitativement, en terme de faux positifs (FP) et de faux négatifs (FN) dans le domaine du visible et de l'infrarouge. Tous ces algorithmes ont été implémentés en utilisant la librairie OpenCV sous Microsoft Visual C et leurs codes sont disponibles dans la librairie BGS Library[5].

De plus, certaines vidéos des résultats sont visibles sur les sites respectifs de chacune de ces trois contributions : T2FMOG[6], Fuzzy Foreground Detection[7] et Fuzzy Background Maintenance[8]. Une liste compléte de celles-ci est disponible sur Vimeo[9].

Pour une évaluation plus complète de ces trois méthodes, le lecteur pourra se reporter à l'évaluation disponible depuis Mai 2014, et effectuée par Sobral et Vacavant [486] sur la base de données BMC 2012.

6.1 Modèlisation du fond par mélange de Gaussiennes floues de Type-2

1- Evaluations qualitatives du modèle T2-FMOG

Nous avons testé l'algorithme T2-FMOG avec ses deux versions T2-FMOG-UM et T2-FMOG-UV en les comparant, à chaque fois, au modèle MOG classique développé par Stauffer et Grimson [490]. Dans cette partie, nous présentons donc une évaluation qualitative de nos algorithmes sur des séquences filmées dans des environnements différents en extérieurs et dans le domaine du visible.

Dans la thèse de Fida El Baf et dans les publications relatives [46][47][74], des résultats sur des scènes d'intérieur présentant des problèmes de variations lumineuses et d'ombres sont présentés ainsi que des résultats sur des scènes filmées par des capteurs infrarouge.

5. http://code.google.com/p/bgslibrary/
6. http://sites.google.com/site/t2fmog/
7. http://sites.google.com/site/fuzzyforegrounddetection/
8. http://sites.google.com/site/fuzzybackgroundmaintenance/
9. http://vimeo.com/thierrybouwmans/videos/all

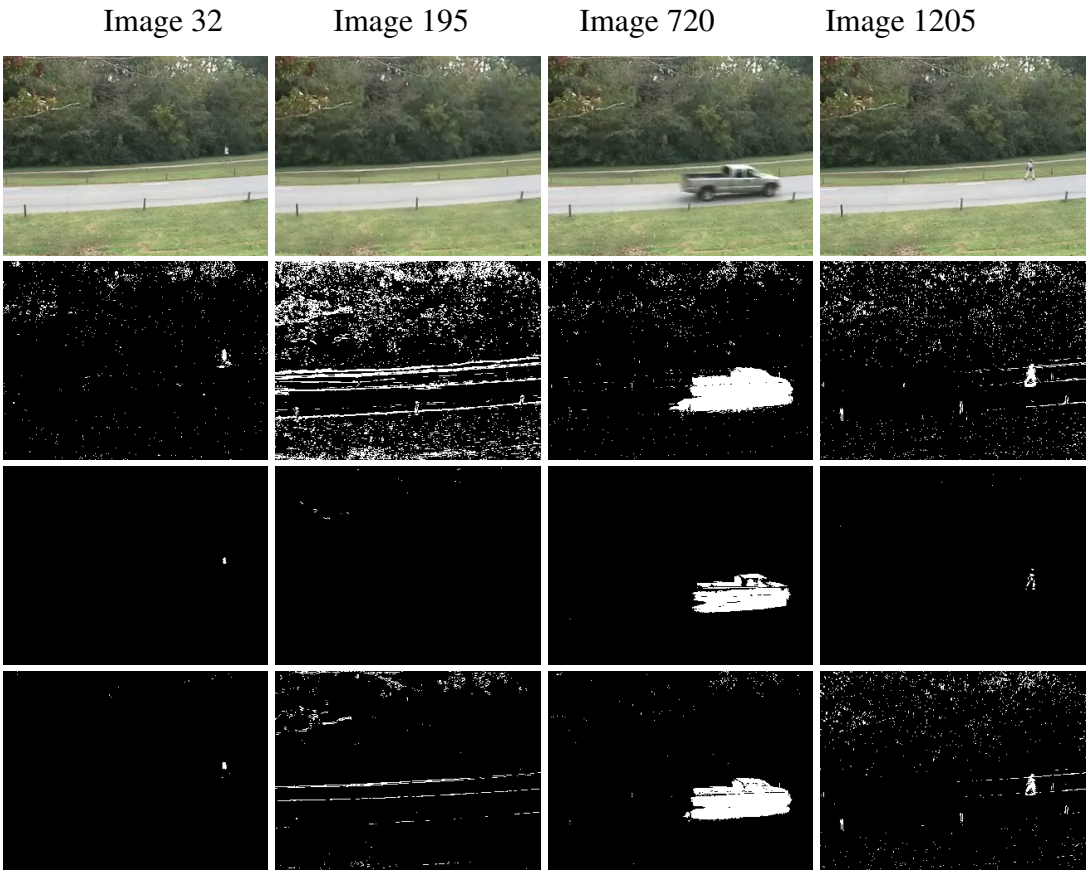


FIGURE 3.10 – Cas de la vibration de caméra. La première ligne montre des images originales provenant des séquences employées par Cheng et al. [116]. La deuxième ligne présente les images segmentées par les MOG. La troisième et la quatrième ligne illustrent les résultats obtenus respectivement par T2-FMOG-UM et T2-FMOG-UV.

Finalement, pour montrer la robustesse de notre algorithme aux fonds dynamiques, nous avons choisi des séquences filmées à l'extérieur présentant les cas suivants : vibration de la caméra, mouvement des arbres, écoulement d'eau (fontaine) et mouvement de la surface de l'eau (vague).

— **Cas de la vibration de caméra :** Cette séquence est issue de la base de données de Bhandarkar et Luo [62]. Les images sont de dimension 320×240 et filmées en RGB. La Figure 3.10 montre les résultats obtenus. Sur l'image 195, les MOG génèrent beaucoup de fausses détections (FP) dûes à la vibration de la caméra qui affecte le champs visuel filmé. Avec T2-FMOG-UM, cette situation est prise en compte rapidement par le modèle et par conséquent ce mouvement est parfaitement intégré dans le fond en détectant toujours bien le véhicule dans l'image 720. Par rapport au T2-FMOG-UV, la marge accordée à la variance permet de mieux gérer le mouvement du fond dû à la vibration de la caméra que les MOG mais moins bien que T2-FMOG-UM.

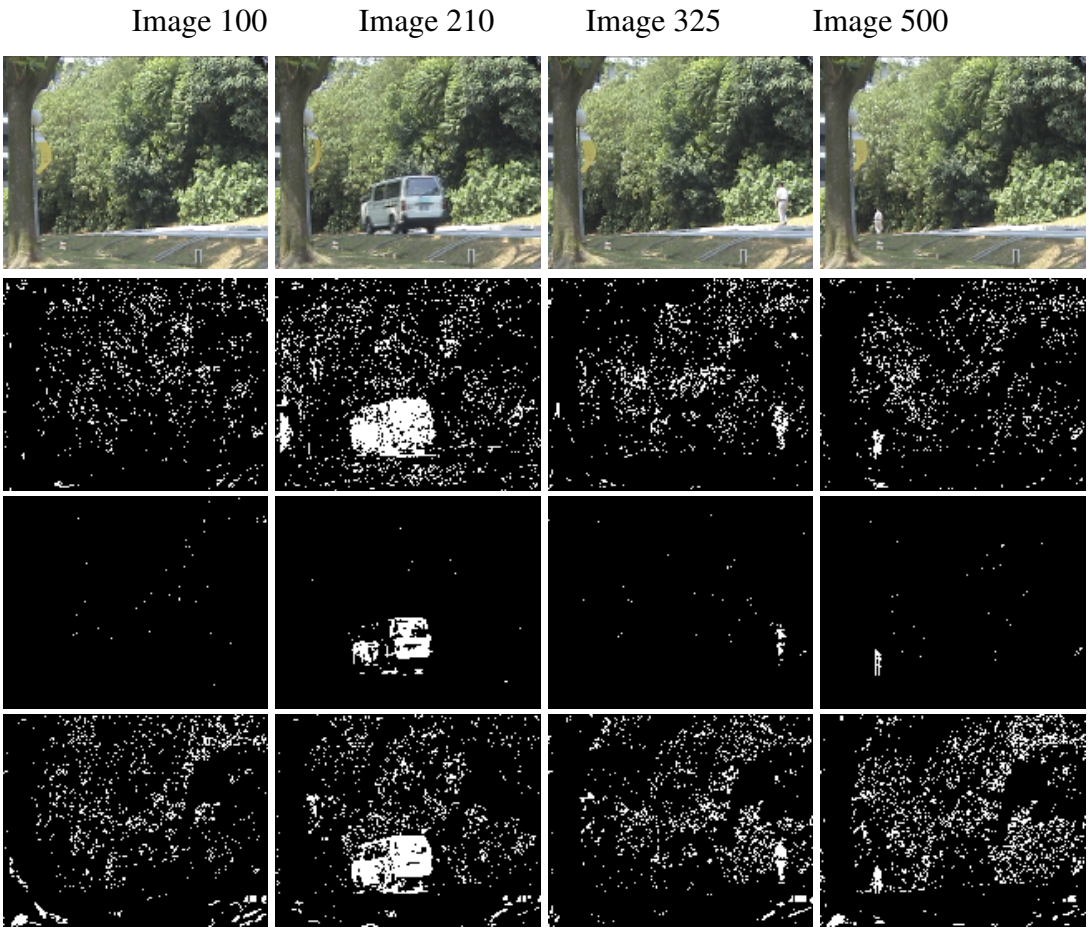


FIGURE 3.11 – Cas d'un mouvement des arbres. La première ligne montre des images originales de la séquence "Campus" de la base de données I2R [311]. La deuxième ligne présente les images segmentées par les MOG. La troisième et la quatrième ligne illustrent les résultats obtenus en utilisant T2-FMOG-UM et T2-FMOG-UV respectivement.

— **Cas du mouvement des arbres :** Cette séquence s'appelle "Campus". Elle est issue de la base de données de Li et al. [311]. Les images sont de dimension 160×128 et filmées en RGB. La Figure 3.11 montre les résultats obtenus. Les mouvements réguliers dans le fond dûs au mouvement des arbres sont une source de fausses détections pour les MOG ainsi que pour T2-FMOG-UV. Néanmoins, notons que les objets d'intérêt sont relativement bien détectés. Par contre, T2-FMOG-UM se montre plus robuste à ce type de mouvement du fond, mais remarquons que les objets d'intérêt sont détectés avec plus de faux négatifs (FN).

— **Cas d'un écoulement d'eau :** Cette séquence s'appelle "Fountain". Elle est issue de la base de données de Li et al. [311]. Les images sont de dimension 160×128 et filmées en RGB. La Figure 3.12 montre les résultats obtenus. Les jets d'eau qui jaîllissent de la fontaine représentent un mouvement assez fort.

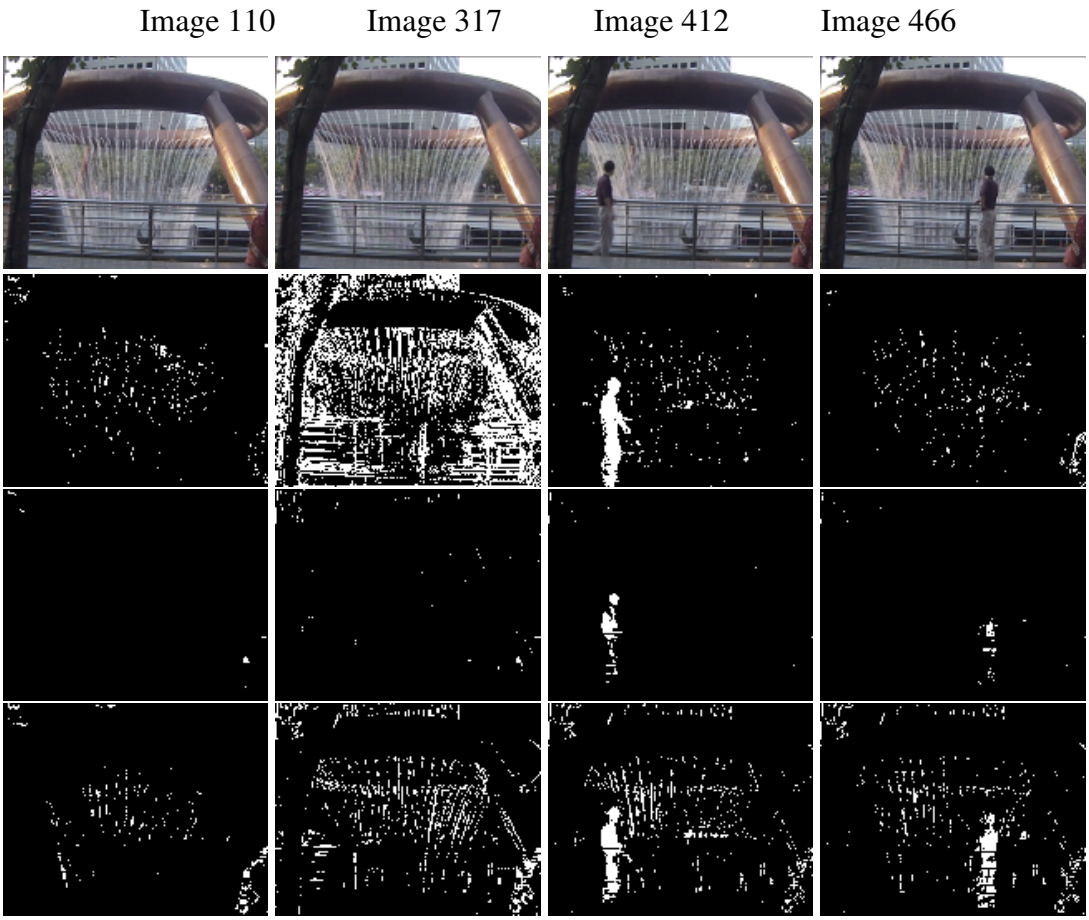


FIGURE 3.12 – Cas d'un écoulement d'eau. La première ligne montre des images originales de la séquence "Fountain" de la base de données I2R [311]. La deuxième ligne présente les images segmentées par les MOG. La troisième et la quatrième ligne illustrent les résultats obtenus en utilisant T2-FMOG-UM et T2-FMOG-UV respectivement.

En plus des variations lumineuses, ces jets provoquent beaucoup de fausses détections avec les MOG comme nous pouvons le voir dans l'image 317. Dans cette situation, T2-FMOG-UV est plus robuste, et est suivi par T2-FMOG-UM. Cependant, en présence d'objets d'intérêt dans la scène, ce dernier détecte la silhouette moins bien que l'algorithme T2-FMOG-UM ou l'algorithme MOG. Mais, quand la personne reste à la même position pour un certain moment (image 466), les MOG la perdent complètement. Par contre, T2-FMOG-UV la détecte toujours bien avec une certaine sensibilité au fond dynamique alors que T2-FMOG-UM signale faiblement sa présence avec une insensibilité très élevée au fond dynamique. Pour résumer, d'un point de vue visuel et par rapport à un mouvement du fond causé par un écoulement d'eau, T2-FMOG-UM se montre plus performant que T2-FMOG-UV ou les MOG. Néanmoins, les objets d'intérêt sont moins bien détectés.

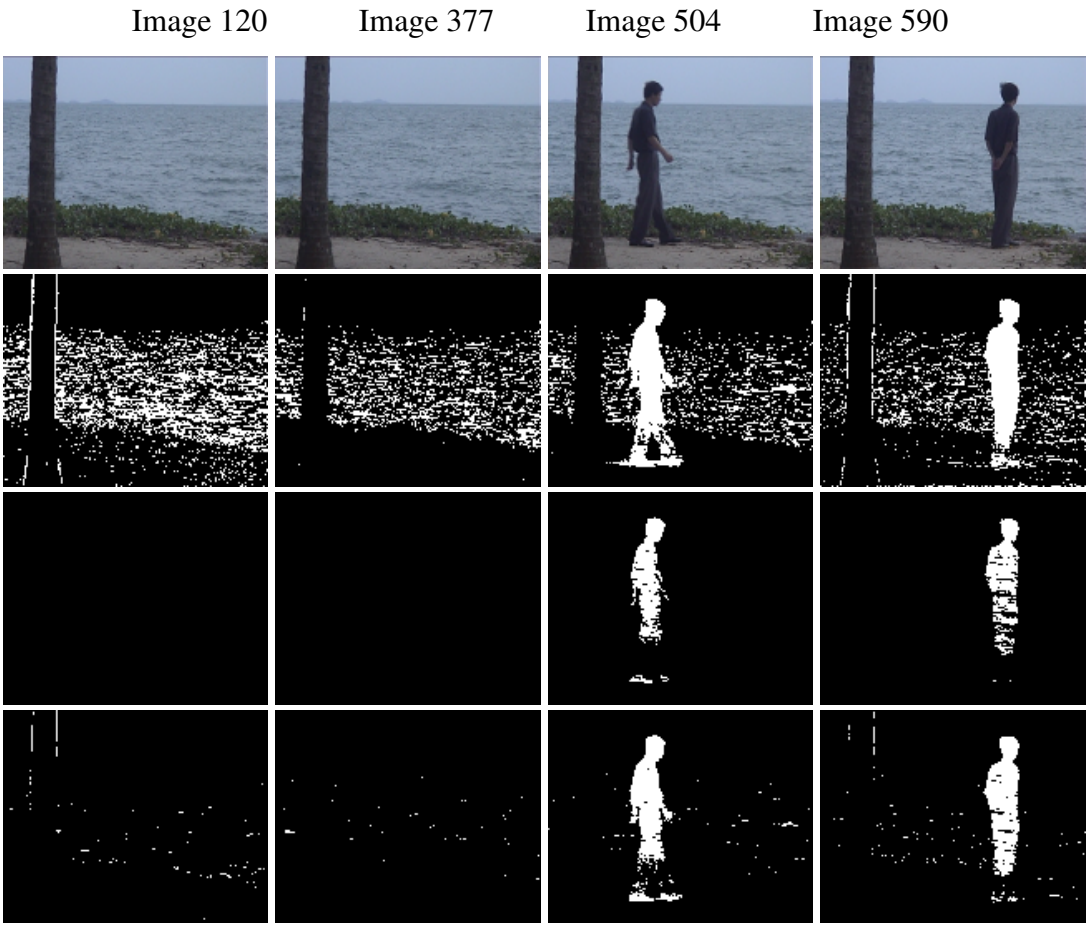


FIGURE 3.13 – Cas d'un mouvement de la surface de l'eau. La première ligne montre les images originales de la séquence "Water Surface" de la base de données I2R [311]. La deuxième ligne présente les images segmentées par les MOG. La troisième et la quatrième ligne illustrent les résultats obtenus en utilisant T2-FMOG-UM et T2-FMOG-UV respectivement.

— **Cas du mouvement de la surface de l'eau :** Cette séquence s'appelle "Water Surface". Elle est issue de la base de données de Li et al. [311]. Les images sont de dimension 160×128 et filmées en RGB. Les vagues génèrent des fausses détections avec les MOG comme le montre la Figure 3.13- image 165. Par contre, T2-FMOG-UV et T2-FMOG-UM donnent des performances similaires. De plus, avec ces deux algorithmes, les silhouettes sont bien détectés. On peut penser que cela est dû au fait que la personne, dans cette séquence, occupe une proportion de l'image relativement plus importante que celle occupée par la personne qui figure dans la séquence précédente (Fountain).

2- Evaluations qualitatives et quantitatives du modèle T2-FMOG

Pour évaluer d'une façon à la fois qualitative et quantitative la performance du modèle T2-FMOG proposé pour la modélisation de fond, nous avons choisi de présenter l'évaluation effectuée sur la séquence de Carnegie Mellon (Sheikh et Shah [461]) qui est fournie avec les vérités terrain pour toute la séquence. Les images sont de taille 360×240. Cette dernière a été filmée à l'extérieur dont l'objectif est de surveiller le trafic routier et les piétons pouvant traverser une route. Les situations critiques qui apparaissent dans cette scène se résument par la vibration de la caméra, les variations lumineuses produites au cours du temps et le fond multimodal représenté par les mouvements des arbres. Dans ces évaluations quantitatives, nous montrons, dans un premier temps, les résultats pour T2-FMOG-UM et T2-FMOG-UV en les comparant premièrement avec les MOG classiques de Stauffer et Grimson [490]. Ensuite, nous présentons des versions floues T2-FMOG adaptées à d'autres versions améliorées des MOG notamment à celle de KaewTraKulPong et Bowden [267] et celle de Zivkovic [649].

Finalement, les fausses détections sont mesurées dans les tableaux 3.4, 3.5 et 3.6 en termes de faux positifs (FP) et de faux négatifs (FN) pour chacun des algorithmes (MOG, T2-FMOG-UM et T2-FMOG-UV) et pour les trois versions des MOG ([490][267][649]). Les valeurs quantitatives montrent que la deuxième et la troisième version des MOG sont meilleures que la version originale car le nombre total des erreurs a été bien réduit de 18576 à 13981 pour [649] et 7830 pour [267]. Ceci est expliqué par l'adaptation des versions améliorées des MOG à mieux interagir avec des fonds multimodaux.

Maintenant, si nous regardons la performance des algorithmes T2-FMOG-UM et T2-FMOG-UV pour toutes les versions des MOG, nous remarquons une baisse nette du nombre total d'erreurs par rapport au MOG comme par exemple, de 18576 pour les MOG (version originale) à 10631 pour T2-FMOG-UM. Néanmoins, T2-FMOG-UM et T2-FMOG-UV donnent des performances presque similaires au niveau des valeurs du nombre total d'erreurs dans le cas des MOG de [490] et [649] bien que, visuellement, T2-FMOG-UM semble être plus performant au niveau de la détection. Ceci est expliqué par le fait que le nombre de faux négatifs (FN) générées avec T2-FMOG-UM est plus que ceux générées avec T2-FMOG-UV avec beaucoup moins de faux positifs (FP) (les objets d'intérêt sont moins bien détectés). Par contre, avec la version de [267], l'erreur totale est supérieure avec T2-FMOG-UV à cause des faux négatifs (FN) générées par la perte des objets d'intérêts (aucun objet n'est quasiment détecté). Les figures 3.17, 3.18 et 3.19 illustrent les performances des trois algorithmes en termes de faux positifs (FP) et de faux négatifs (FN). L'algorithme T2-FMOG-UM enregistre la meilleure performance avec une moyenne d'erreurs la plus faible pour les trois versions des MOG et le moins de faux positifs (FP) sauf pour la version des MOG améliorée par KaewTraKulPong et Bowden.

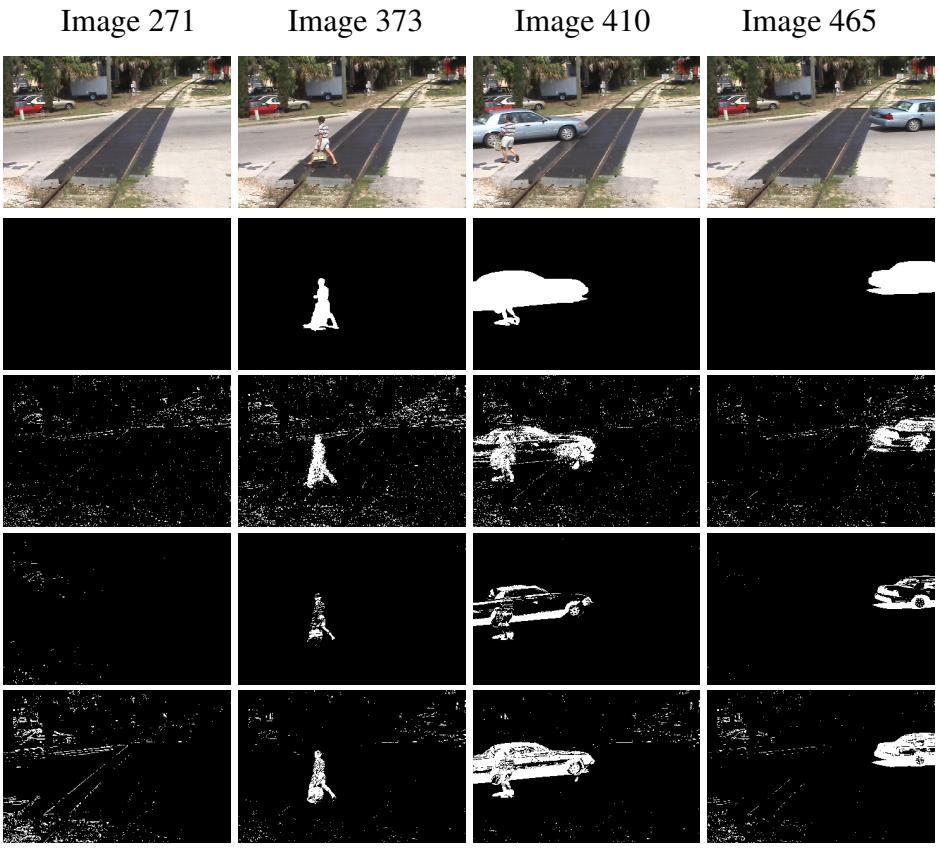


FIGURE 3.14 – Cas des MOG de Stauffer et Grimson [490] : La première ligne montre des images originales de la séquence de Shah et al. [461]. La deuxième ligne présente les vérités terrain. La troisième ligne montre les images segmentées par les MOG [490]. La quatrième et la cinquième lignes illustrent les résultats obtenus par l'algorithme T2-FMOG adapté à l'algorithme des MOG de Stauffer et Grimson [490] sous ses deux versions T2-FMOG-UM et T2-MOG-UV respectivement.

— **Cas des MOG de Stauffer et Grimson [490] :** La Figure 3.14 montre des images segmentées avec les MOG de Stauffer et Grimson [490] ainsi que les résultats obtenus avec l'algorithme T2-FMOG qui est adapté à cette version. Les MOG classiques permettent de détecter les piétons et les voitures avec beaucoup de faux positifs (FP) dus aux changement de luminosité et au mouvement du fond régulier et rapide dans certaines zones de l'image. T2-FMOG avec ces deux versions reste moins sensible à ces problèmes. En effet, T2-FMOG-UM donne le moins de faux positifs (FP) sur toute la séquence. Néanmoins les objets d'intérêt sont moins bien détectés. Par contre, nous pouvons remarquer qu'entre les trois algorithmes, T2-FMOG-UV est celui qui fournit la meilleure détection des objets d'intérêt avec une sensibilité aux problèmes du fond qui reste moindre que MOG et T2-FMOG-UM.

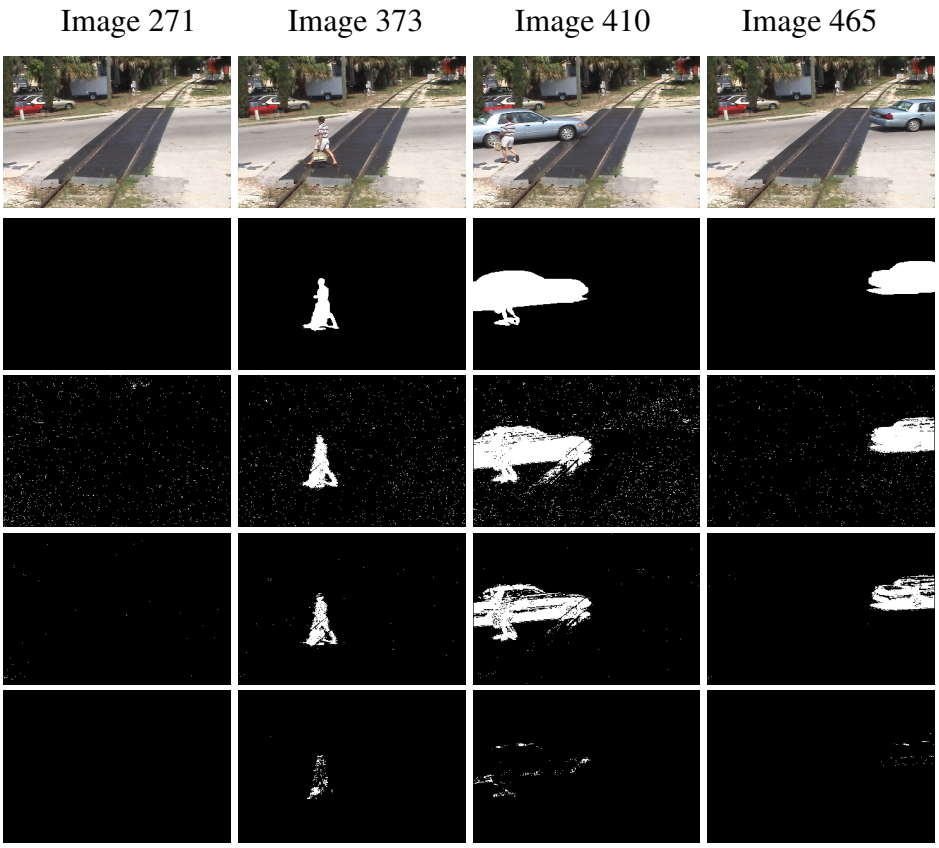


FIGURE 3.15 – Cas des MOG de KaewTraKulPong et Bowden [267] : La première ligne montre des images originales de la séquence de Shah et al. [461]. La deuxième ligne présente les vérités terrain. La troisième ligne montre les images segmentées par cette variante des MOG. La quatrième et la cinquième lignes illustrent les résultats obtenus par l'algorithme T2-FMOG adapté à l'algorithme des MOG de KaewTraKulPong et Bowden [267] sous ses deux versions T2-FMOG-UM et T2-MOG-UV respectivement.

— **Cas des MOG de KaewTraKulPong et Bowden [267] :** La Figure 3.15 montre des images segmentées avec les MOG de KaewTraKulPong et Bowden [267] ainsi que les résultats obtenus avec l'algorithme T2-FMOG qui est adapté à cette version. Rappelons que cette version a amélioré l'étape de l'initialisation des paramètres des MOG en utilisant l'algorithme EM. A première vue, les images segmentées avec cette version des MOG ($3^{ème}\ ligne$) par comparaison au MOG classiques de la Figure 3.15 sont meilleures ; les objets d'intérêt sont bien détectés avec moins de faux positifs (FP). Cependant, cet algorithme ne résout pas tous les problèmes de luminosité qui affecte toujours la scène. T2-FMOG-UM et T2-FMOG-UV sont tous les deux robustes à ces problèmes. T2-FMOG-UM donne une très bonne performance au niveau de la qualité de la détection d'objets d'intérêt sur toute la séquence avec quasiment pas de faux positifs (FP). T2-FMOG-UV perd presque la trace des objets d'intérêts et n'est pas fiable pour la détection.

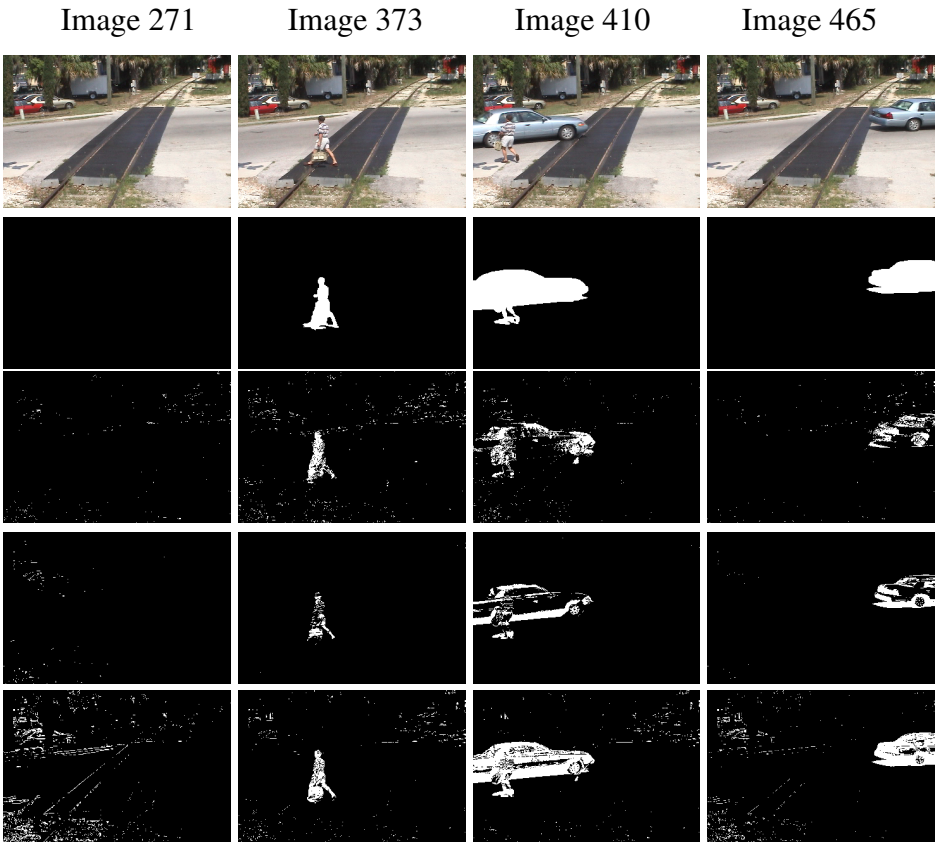


FIGURE 3.16 – Cas des MOG de Zivkovic [649] : La première ligne montre des images originales de la séquence de Shah et al. [461]. La deuxième ligne présente les vérités terrain. La troisième ligne montre les images segmentées par cette variante des MOG. La quatrième et la cinquième lignes illustrent les résultats obtenus par l'algorithme T2-FMOG adapté à l'algorithme des MOG de Zivkovic [649] sous ses deux versions T2-FMOG-UM et T2-MOG-UV respectivement.

— **Cas des MOG de Zivkovic [649] :** La Figure 3.16 montre les images segmentées avec les MOG de Zivkovic [649] ainsi que les résultats obtenus avec l'algorithme T2-FMOG qui est adapté à cette version. Cette version détermine automatiquement le nombre de Gaussiennes qui est fixé entre 3 et 5 dans les versions précédentes. Les résultats obtenus avec cette version des MOG montrent une meilleure adaptation aux problèmes de luminosité et aux fonds dynamiques que les autres versions des MOG (moins de FP). Cette adaptation reste trop limitée par rapport à celle assurée par T2-FMOG-UM et les objets d'intérêt ne sont relativement pas bien détectés en comparaison aux deux autres algorithmes. Entre les trois algorithmes, T2-FMOG-UV détecte les objets d'intérêt avec la plus grande précision. Cependant, plus de faux positifs sont générées avec cet algorithme.

Méthode	Type d'erreurs	Image 271	Image 373	Image 410	Image 465	Total d'erreurs	Moyenne d'erreurs	Moyenne d'erreurs
MOG	FN	0	1120	4818	2050			
Stauffer et Grimson [490]	FP	2093	4124	2782	1589	18576	4644	12512
T2-FMOG-UM	FN	0	1414	6043	2520			
El Baf et al. [46]	FP	203	153	252	46	10631	2657	1400
T2-FMOG-UV	FN	0	957	2217	1069			
El Baf et al. [46]	FP	3069	1081	1119	1158	10670	2667	2624

TABLE 3.4 – Calculs des faux positifs (FP) et des faux négatifs (FN) des MOG de Stauffer et Grimson [490] et de leurs versions floues [46] sur la séquence de Shah et al. [461]. En gras, les approches développées par Fida El Baf sous ma direction scientifique.

Méthode	Type d'erreurs	Image 271	Image 373	Image 410	Image 465	Total d'erreurs	Moyenne d'erreurs	Moyenne d'erreurs
MOG	FN	0	265	637	413			
Kaew. et Bowden [267]	FP	1034	1359	3308	814	7830	1957	2061
T2-FMOG-UM	FN	0	522	2179	1251			
El Baf et al. [46]	FP	37	287	787	122	5185	1296	685
T2-FMOG-UV	FN	0	757	4130	1818			
El Baf et al. [46]	FP	0	162	252	21	7140	1785	1560

TABLE 3.5 – Calculs des faux positifs (FP) et des faux négatifs (FN) des MOG de KaewTraKulPong et Bowden [267] et de leurs versions floues [46] sur la séquence de Shah et al. [461] . En gras, les approches développées par Fida El Baf sous ma direction scientifique.

Méthode	Type d'erreurs	Image 271	Image 373	Image 410	Image 465	Total d'erreurs	Moyenne d'erreurs	Moyenne d'erreurs
MOG	FN	0	1152	6688	3009			
Zivkovic [649]	FP	341	1404	1977	310	13981	3495	1949
T2-FMOG-UM	FN	8	1414	6043	2520			
El Baf et al. [46]	FP	204	154	252	48	10643	2660	1400
T2-FMOG-UV	FN	0	957	2216	1068			
El Baf et al. [46]	FP	3072	1082	1119	1166	10680	2670	2626

TABLE 3.6 – Calculs des faux positifs (FP) et des faux négatifs (FN) des MOG de Zivkovic [649] et de leurs versions floues [46] sur la séquence de Shah et al. [461]. En gras, les approches développées par Fida El Baf sous ma direction scientifique.

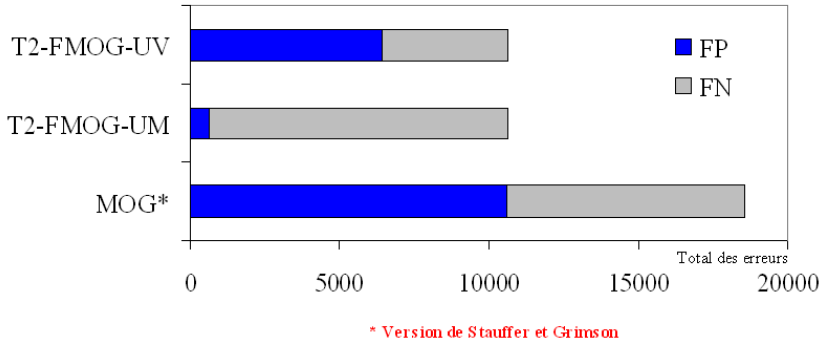


FIGURE 3.17 – Proportion des faux positifs (FP) et des faux négatifs (FN) pour les MOG de Stauffer et Grimson [490], T2-FMOG-UM et T2-FMOG-UV sur les images 271, 373, 410 et 465.

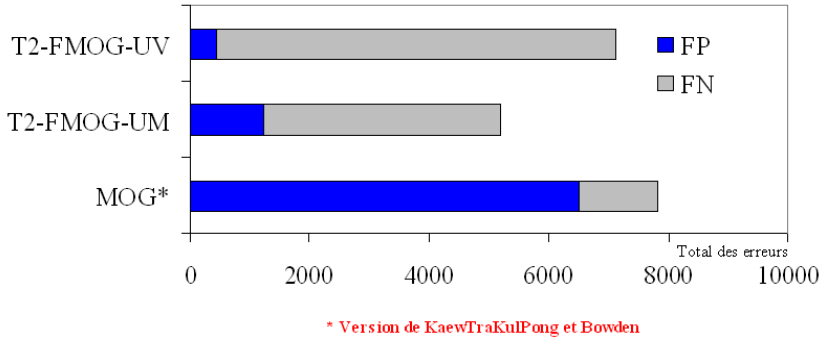


FIGURE 3.18 – Proportion des faux positifs (FP) et des faux négatifs (FN) pour les MOG de KaewTraKulPong et Bowden [267], T2-FMOG-UM et T2-FMOG-UV sur les images 271, 373, 410 et 465.

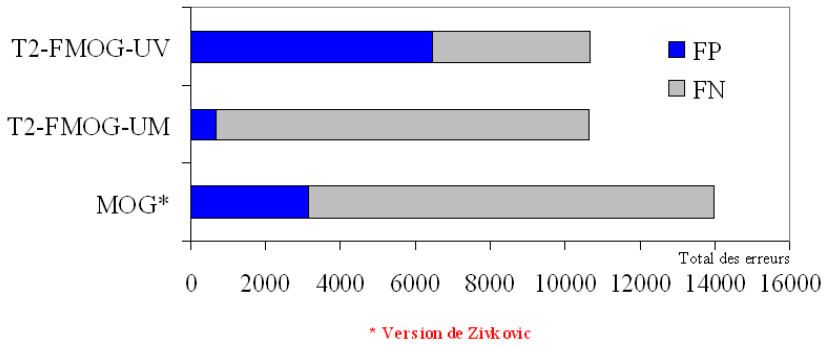


FIGURE 3.19 – Proportion des faux positifs (FP) et des faux négatifs (FN) pour les MOG de Zivkovic [649], T2-FMOG-UM et T2-FMOG-UV sur les images 271, 373, 410 et 465.

6.2 Classification floue des pixels en fond/objet par l'intégrale de Choquet

Dans cette partie, nous comparons, pour différentes séquences vidéo, l'algorithme proposé qui utilise l'intégrale de Choquet avec celui proposé par Zhang et Xu [620] qui utilise l'intégrale de Sugeno. Pour cette évaluation, nous comparons dans un premier temps l'agrégation des composantes couleurs entre-elles et ensuite l'agrégation entre un espace couleur et la caractéristique texture LBP.

1- Agrégation de différentes composantes couleurs

Nous avons utilisé une des séquences de "PETS 2006" pour évaluer notre algorithme de détection qui utilise l'opérateur flou de Choquet pour la fusion des caractéristiques couleurs. Cette séquence est filmée à l'intérieur d'une gare où la scène subit des variations lumineuses et des problèmes d'ombres. Les images sont de taille 720×576. Nous avons comparé les résultats obtenus par l'intégrale de Choquet aux résultats obtenus par les deux opérateurs ; l'intégrale de Sugeno et l'opérateur logique "OU". Dans le tableau 3.7, nous montrons, à chaque fois, les résultats de ces trois opérateurs pour trois espaces couleur différents : YCrCb, Ohta et HSV. Les expérimentations montrent que l'intégrale de Choquet est mieux adaptée pour la fusion des caractéristiques couleurs que l'intégrale de Sugeno ou encore que l'opérateur logique, "OU" quelque soit l'espace couleur choisi. De plus, l'espace YCrCb enregistre une meilleure performance par rapport aux espaces Ohta et HSV. Les évaluations quantitatives réalisées en termes de faux positifs (FP) et de faux négatifs (FN) confirment ces résultats. En effet, l'opérateur logique "OU" génère plus de fausses détections que les opérateurs flous (Sugeno et Choquet) car une fausse détection générée dans une composante est forcément amenée dans le résultat final à cause de la nature de la fusion réalisée avec cet opérateur. A la différence, les opérateurs flous (Sugeno et Choquet) donnent plus de flexibilité au mécanisme de fusion, car ils définissent une certaine pondération pas uniquement à chaque composante mais aussi à toutes les coalitions possibles des composantes. Par conséquent, une fausse détection générée dans une composante n'a pas forcément de répercussion dans le résultat final. Cela permet de mieux gérer les fausses détections et d'améliorer par la suite la qualité de la détection d'objets mobiles. Néanmoins, l'intégrale de Choquet que nous proposons semble être plus adaptée pour la détection que l'intégrale de Sugeno. Notre algorithme qui utilise l'intégrale de Choquet génère le moins d'erreurs avec tous les espaces couleur testés 6065 avec YCrCb, 7197 avec Ohta et 11353 avec HSV (voir le tableau 3.9 et la figure 3.20). Ces évaluations montrent qu'après la performance enregistrée par l'opérateur proposé, vient celle l'intégrale de Sugeno suivie de celle de l'opérateur logique "OU". Vu que les données sont d'ordre cardinal (la différence entre les valeurs est bien réelle), cela explique la meilleure adaptation de l'intégrale de Choquet par rapport à l'intégrale de Sugeno à ce type de données. Les mesures de la similarité de Li [311] (ou coefficient de Jaquard) calculées à partir des erreurs dans la dernière colonne du tableau 3.9 confirment ce résultat en mettant en évi-

dence la meilleure performance de l'intégrale de Choquet avec l'espace couleur YCrCb avec la valeur de la similarité la plus élevée (0.797). Nous notons que les résultats de chaque algorithme sont donnés avec le meilleur réglage du seuil T et des 3 mesures floues attribuées à chaque critère. Les valeurs de T et des mesures floues choisies pour chaque algorithme selon les espaces couleur sont fournies dans le tableau 3.8.

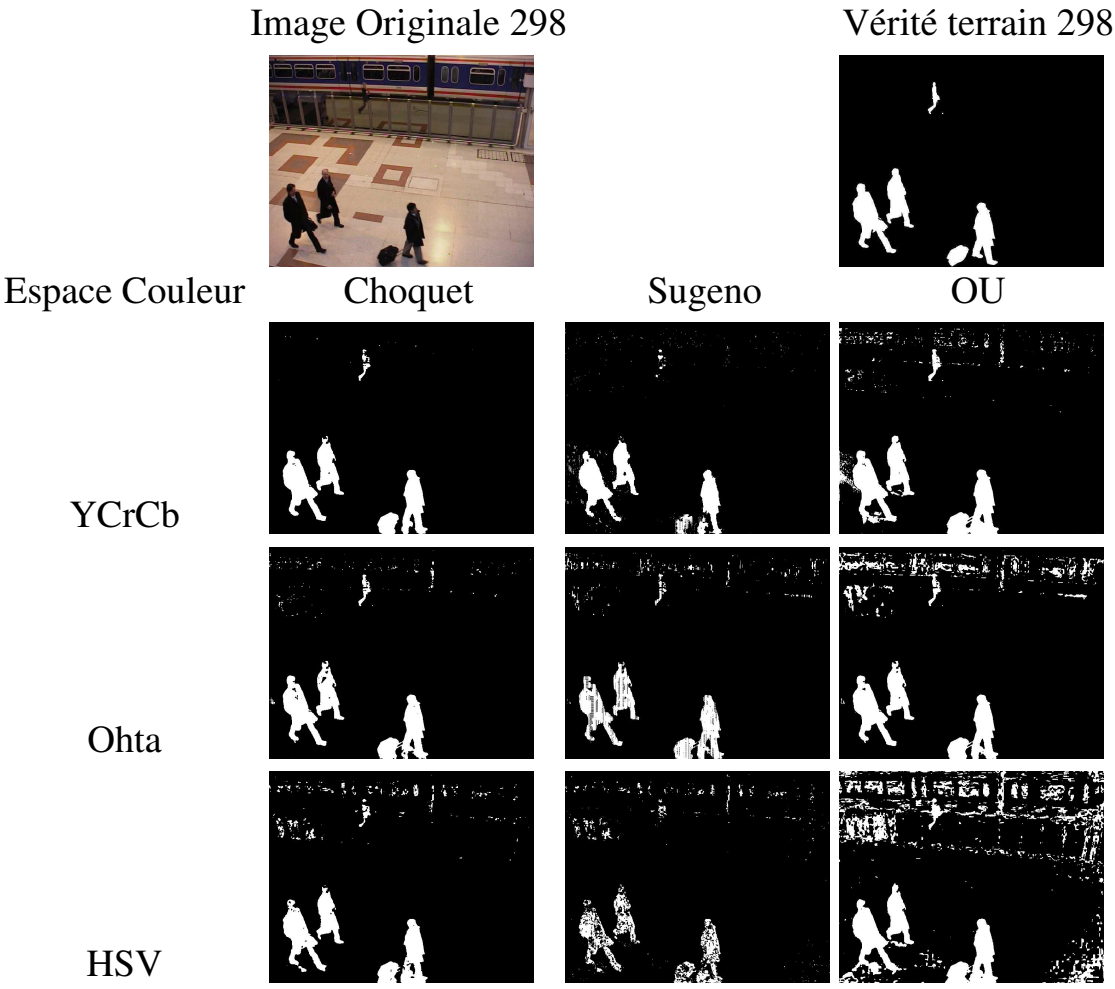


TABLE 3.7 – Fusion des composantes couleur (YCrCb, Ohta et HSV) par Choquet, Sugeno et par l'opérateur "OU" à l'image 298 d'une des séquences de "PETS 2006"

2- Agrégation de caractéristiques couleur et texture

L'évaluation de l'intégrale de Choquet en utilisant les caractéristiques couleur et texture a été effectuée sur les séquences du projet "Aqu@thèque" [40] pour une application en reconnaissance de forme. Celle-ci a été faite en termes de faux positifs (FP) et de faux négatifs (FN). Nous avons également généré les courbes ROC pour comparer la performance de notre algorithme par rapport à une autre approche en testant plusieurs espaces couleur.

Espaces Couleur	Méthodes	Th	x_1	x_2	x_3	x_1, x_2	x_1, x_3	x_2, x_3	X
YCrCb	Choquet	0.8	0.4	0.3	0.3	0.7	0.7	0.6	1
	Sugeno	0.4	0.4	0.3	0.3	0.7	0.7	0.6	1
	OR	0.8	-	-	-	-	-	-	-
Ohta	Choquet	0.7	0.0199	0.98	0.0001	0.9999	0.02	0.9801	1
	Sugeno	0.7	0.0199	0.98	0.0001	0.9999	0.02	0.9801	1
	OR	0.7	-	-	-	-	-	-	-
HSV	Choquet	0.62	0.04	0.82	0.15	0.86	0.19	0.97	1
	Sugeno	0.65	0.04	0.82	0.15	0.86	0.19	0.97	1
	OR	0.62	-	-	-	-	-	-	-

TABLE 3.8 – Paramètres utilisés : Seuil et mesures floues.

Espaces Couleur	Méthodes	FN	FP	Erreur Totale	$S(A, B)$
YCrCb	**Choquet**	1395	4670	6065	**0.797**
	Sugeno	8387	1815	10202	0.664
	OR	3383	8376	11759	0.681
Ohta	**Choquet**	4391	2806	7197	**0.770**
	Sugeno	7668	3959	11627	0.642
	OR	4065	14091	18156	0.574
HSV	**Choquet**	5464	5889	11353	**0.670**
	Sugeno	14166	4650	18816	0.433
	OR	3811	12209	16020	0.349

TABLE 3.9 – Evaluation quantitative des algorithmes de détection par Choquet, Sugeno et OR selon les espaces couleur. En gras, l'approche développée par Fida El Baf sous ma direction scientifique.

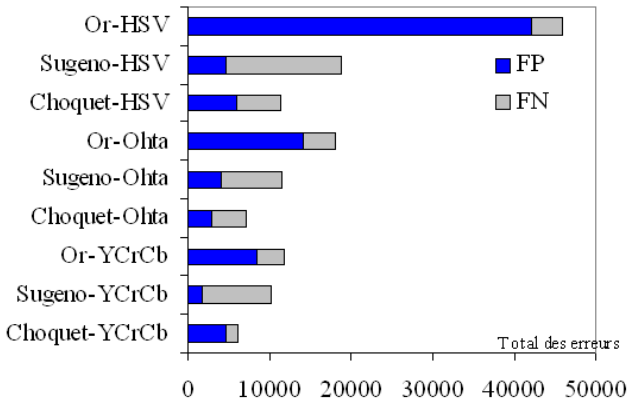


FIGURE 3.20 – Performance des algorithmes de détection par Choquet, Sugeno et OU selon les espaces couleur. Pour un espace couleur donné, l'approche proposée utilisant l'intégrale de Choquet donne le moins de fausses détections.

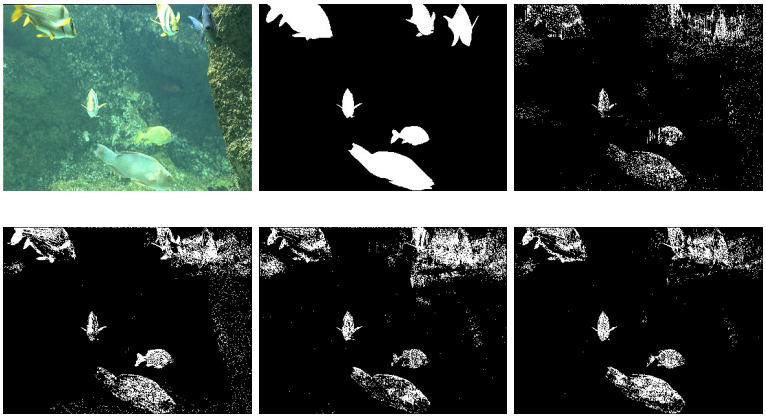


FIGURE 3.21 – Première ligne : Image courante, vérité terrain, Sugeno (Ohta, LBP) développée par Zhang et Xu [620]). Seconde ligne : **Choquet (Ohta-LBP)** [45], **Choquet (HSV-LBP)** [45], et **Choquet (YCrCb, LBP)** [45]. En gras, les approches développées par Fida El Baf sous ma direction scientifique.

$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_1, x_3\}$	$\{X\}$
0.6	0.3	0.1	0.9	0.7	0.4	1
0.5	0.4	0.1	0.9	0.6	0.5	1
0.5	0.3	0.2	0.8	0.7	0.5	1
0.50	0.39	0.11	0.89	0.61	0.5	1
0.53	0.34	0.13	0.87	0.66	0.47	1

TABLE 3.10 – Mesures floues.

Le meilleur résultat est obtenu avec YCrCb. La figure 3.21 montre les résultats obtenus. Le tableau 3.10 indique les valeurs de densité floues que nous avons testé. Les meilleurs résultats ont été obtenus avec $\{0.53, 0.34, 0.13\}$. On peut remarquer que les résultats obtenus par l'intégrale de Choquet sont meilleurs que ceux obtenus avec l'intégrale de Sugeno [620] avec le même espace couleur (Ohta). Les meilleurs résultats pour l'intégrale de Choquet sont obtenus pour d'autres espaces couleur, HSV et YcrCb comme le montre également les courbes ROC correspondantes à la figure 3.22 et la figure 3.23.

Intégrale	Sugeno (Ohta, LBP) Zhang et Xu [620]	Choquet (Ohta, LBP El Baf et al. [45]	Choquet (HSV, LBP) El Baf et al. [45]	**Choquet** (HSV, LBP) **El Baf et al.** [45]
$S(A, B)$	0.27	0.44	0.34	**0.46**

TABLE 3.11 – Mesures de similarité de Li [311] (ou coefficient de Jaquard) sur la séquence Aqu@thèque. En gras, l'approche développée par Fida El Baf sous ma direction scientifique.

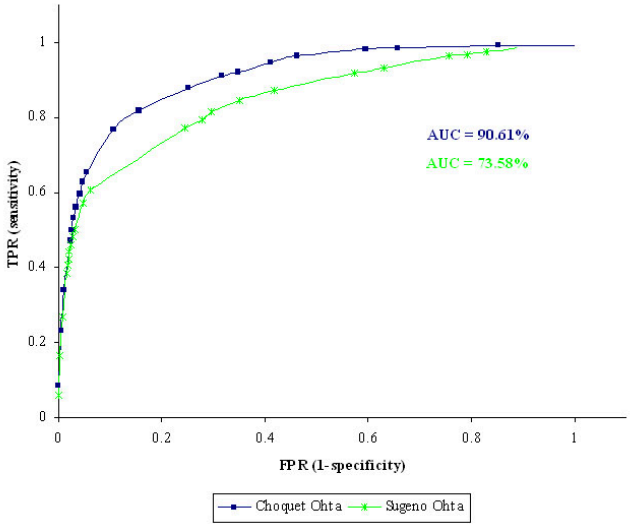


FIGURE 3.22 – Courbes ROC comparant les résultats obtenus avec l'intégrale de Sugeno et ceux obtenus avec de l'intégrale de Choquet avec l'espace couleur Ohta.

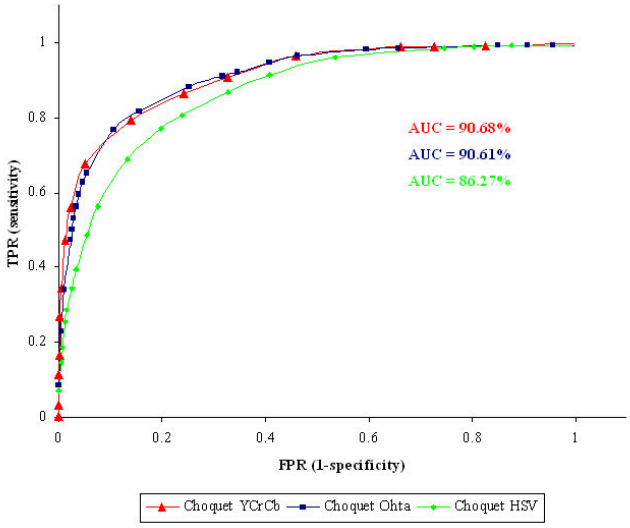


FIGURE 3.23 – Courbes ROC comparant l'influence des espaces couleur Ohta, HSV et YCrCb sur la détection en utilisant l'intégrale de Choquet.

FIGURE 3.24 – PETS 2003 : Image 340. De gauche à droite : Image originale, résultat obtenu par l'intégrale de Sugeno (Ohta, LBP) [620], résultat obtenu par l'intégrale de Choquet (YCrCb, LBP).

Nous avons également testé notre algorithme sur d'autres types de séquences tant dans le domaine du visible que de l'infrarouge comme suit :

— **PETS 2003 :** La scène filmée montre des joueurs sur un terrain en plein match de foot. Les images sont de taille 720×576. La figure 3.24 montre les résultats obtenus par les deux algorithmes de détection de Sugeno et de Choquet avec respectivement les espaces couleur Ohta et YCrCb. Par comparaison avec l'algorithme proposé par Zhang et Xu [620], les résultats de la détection de notre algorithme (Choquet - YCrCb) donne une meilleure performance. Les silhouettes sont mieux détectées et l'algorithme est moins sensible aux variations lumineuses qui se produisent dans certaines régions, notamment, sur la ligne blanche du terrain de foot.

— **PETS 2000 :** Il s'agit d'une application en vidéosurveillance. La scène filmée montre une route traversée par des voitures où des variations lumineuses peuvent apparaître. Les images sont de taille 720×576. Dans ces évaluations, nous avons utilisé les caractéristiques couleurs et texture (LBP) pour les deux algorithmes de détection de Sugeno et de Choquet avec respectivement les espaces couleur Ohta et YCrCb. La figure 3.25 montre les résultats obtenus par les deux algorithmes. Par comparaison avec la méthode de Zhang et Xu [620], les résultats de la détection montrent que notre algorithme est plus robuste. En effet, la voiture est mieux détectée avec Choquet-(YCrCb, LBP).

— **OTCBVS 2006 :** La scène issue de la "Collection 1" montre un carrefour où des voitures et des piétons passent. Les conditions d'éclairage sont perturbées provoquant ainsi beaucoup de bruit. Dans la figure 3.26, nous comparons le résultat de la segmentation obtenue avec l'intégrale de Choquet utilisant les composantes infra-rouge (IR) et LBP au résultat obtenu avec le mélange de Gaussiennes MOG [490]. Ce dernier reste plus sensible au bruit et génère plus de fausses alarmes que l'algorithme proposé.

FIGURE 3.25 – PETS 2000 (Image 172). De gauche à droite : Image originale, résultat obtenu par l'intégrale Sugeno (Ohta, LBP) [620], résultat obtenu par l'intégrale de Choquet (YCrCb, LBP).

FIGURE 3.26 – OTCBVS 2006 - Collection 1. De gauche à droite : Image originale, vérité terrain, résultat obtenu par l'intégrale de Choquet-(intensité IR, LBP), résultat obtenu par le mélange de Gaussiennes [490].

6.3 Mise à jour de l'image du fond

Nous montrons dans cette section les résultats obtenus pour la règle de mise à jour adaptative que nous avons proposée dans la Section 5. Nous rappelons que cette règle de mise à jour intègre le caractère flou dans le mécanisme de la mise à jour car elle utilise le résultat de la classification de pixel en fond ou en objets mobiles faite par l'intégrale de Choquet.

L'évaluation de l'adaptation du système de détection avec cette règle adaptative est réalisée par comparaison avec les règles de mises à jour sélective et non sélective habituellement utilisées dans la littérature. Pour cela, nous avons choisi la séquence "Time Of Day" de la base de données Wallflower [521] où la scène est filmée à l'intérieur. Le choix de cette séquence est justifié par le fait qu'elle présente un changement progressif de luminosité assez fort. En effet, la luminosité au cours du temps varie de celle de l'aube jusqu'à celle de la lumière de pleine journée ensoleillée qui commence à blanchir la salle. Entre temps, une personne rentre dans la salle et agit dans la scène. Ces variations assez fortes demandent une bonne adaptation du modèle aux changements que subit la scène. Les images sont de taille 160×120. Les expérimentations ont été menées sur 20 images de cette séquence.

La figure 3.27 montre les résultats de la détection en mettant en évidence la capacité d'adaptation du système de détection avec une règle floue et adaptative de mise à jour que nous avons proposée pour la mise à jour par comparaison avec les règles de mises à jour sélective et non sélective qui sont habituellement utilisées dans la littérature. En effet, nous pouvons remarquer que la tête de la personne assise est mieux détectée dans la cas de la mise

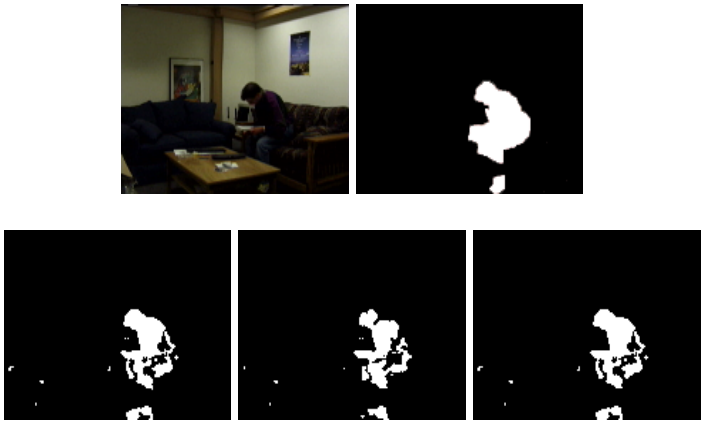


FIGURE 3.27 – Première ligne : Image courante, Vérité terrain. Seconde ligne : Mise à jour non sélective, Mise à jour sélective, Mise à jour adaptative

Mise à jour	Non sélective	Sélective	**Adaptative** [43]
$S(A,B)\,\%$	58.40	57.08	58.96

TABLE 3.12 – Mesures de similarité de Li [311] (ou coefficient de Jaquard) pour la mise à jour non sélective, sélective, et adaptative sur 20 images. En gras, l'approche développée par Fida El Baf sous ma direction scientifique.

à jour adaptative floue que dans le cas de la mise à jour sélective. Par rapport à mise à jour non sélective, la méthode mise à jour adaptative ne donne pas les faux négatifs situé au dessus du livre.

La performance de la règle proposée est calculée, dans le tableau 3.12 à l'aide de la mesure de la similarité de Li [311] (ou coefficient de Jaquard). Dans ces évaluations, les valeurs numériques montrent que la règle de mise à jour proposée donne presque des performances similaires à celles obtenues par les règles non sélectives et sélectives. Par contre, sur toute la séquence, notre règle s'adapte plus vite que les deux autres règles aux changements récents surtout dans le cas des variations lumineuses progressives et rapides comme dans le cas de cette séquence.

6.4 Temps de calcul

Le tableau 3.13 regroupent pour les approches proposées les temps de calcul sur des séquences d'images de résolutions 384*288 en RGB sur un Pentium 1,66GHz avec RAM 1GB. Les implémentations des différents algorithmes ont été effectuées en C avec la libraire OpenCV.

Algorithme	Images par seconde
MOG [490]	20
T2-FMOG-UM [46]	11
T2-FMOG-UV [46]	12
"OU" Logique	40
Intégrale de Sugeno [620]	22
Intégrale de Choquet [45]	31

TABLE 3.13 – Temps de calcul. En gras, les approches développées par Fida El Baf sous ma direction scientifique.

Les algorithmes T2-FMOG-UM et T2-FMOG-UV sont plus lents que les MOG de Stauffer et Grimson [490]. Pour la détection d'"objets mobiles, l'approches basée sur l'intégrale de Choquet se révèle plus rapide que celle basée sur l'intégrale de Sugeno [620]. Tous ces temps de calcul peuvent être améliorés par des implémentations sur GPU.

7 Conclusion

Dans ce chapitre, nous nous sommes focalisés sur l'apport des concepts flous dans la détection d'objets mobiles dans des séquences vidéo. Dans ce cadre, nous avons effectué les trois travaux de recherches suivants :

— **Modélisation floue du fond (Section 3)** : Nous avons proposé une modélisation originale qui applique la théorie des ensembles flous de Type-2 à la modélisation de fonds dynamiques. Nous avons ainsi proposé deux versions de mélange de Gaussiennes floues ; l'une qui gére l'incertitude associée à la moyenne (T2-FMOG-UM) et l'autre introduit le caractère flou à la variance (T2-FMOG-UV). Cette approche s'est montrée particulièrement adaptée aux fonds dynamiques : vibration de la caméra, mouvement d'arbres, écoulement d'eau dans le fond et ondulation de la surface d'eau. Cette modélisation floue est également robuste dans un environnement fermé (scène filmée à l'intérieur) face aux ombres et aux variations lumineuses. Les évaluations qualitatives et quantitatives montrent la pertinence de notre approche par comparaison à la version précise des MOG qui a été développée par Stauffer et Grimson [490] ou encore à deux autres versions améliorées des MOG ([267][649]).

— **Classification floue des pixels en fond/objet (Section 4)** : Pour cette étape, nous avons proposé une méthode de détection d'objets mobiles qui utilisent la caractéristique couleur et une caractéristique texture (LBP). Les résultats provenant de chacune des caractéristiques ont été fusionnés par l'intégrale floue de Choquet permettant ainsi

à la méthode développée d'être plus robuste aux variations lumineuses.

— **Mise à jour de l'image du fond (Section 5)** : Finalement, nous avons développé une méthode floue de mise à jour qui est adaptative à partir du degré d'appartenance du pixel à la classe "fond" et à la classe "objet mobile". Une perspective possible est que le coefficient d'apprentissage soit déterminé par une fonction plus évoluée des degrés d'appartenance pour être plus adaptatif comme une fonction sigmoide ou triangle [565].

Finalement, cette recherche a montré que les concepts flous permettent donc de prendre en compte l'imprécision et l'incertitude au niveau de la modélisation du fond, de la classification fond/objet et de la mise à jour du fond en présence de fonds dynamiques et de variations lumineuses. De plus, ce travail a suscité chez d'autres chercheurs des travaux dans la continuité de ces approches floues développées par Fida El Baf sous ma direction scientifique. Ces autres travaux ont été recensés dans un état de l'art publié en 2012 sur plus de 50 articles sous la proposition de Lucia Maddalena (ICAR, Italie) [71]. Les méthodes SOBS [350] et SC-SOBS [354] qui utilisent une mise à jour basée sur les concepts flous sont des méthodes leaders du domaine sur la base de données ChangeDetection.net montrant donc ainsi de nouveau la pertinence de l'utilisation des concepts flous dans ce domaine. Le lecteur pourra trouver les publications récentes utilisant les concepts flous sur notre site "Background Subtraction Web Site".

Chapitre 4

Approche discriminative et mixte par réduction de dimension pour la modélisation de fond

1 Introduction

Dans ce chapitre [1], nous nous focalisons sur les approches de réduction de dimension pour la modélisation de fond. **Ces travaux ont été menés sous ma direction scientifique dans le cadre du stage ERASMUS de Master 2 de Diana Farcas et Cristina Marghes (Univ. Timisoara, Roumanie).**

Les approches de réduction de dimension sont classées en deux grandes catégories par Skocaj et al. [483][484] : reconstructives et discriminatives. Avec les méthodes reconstructives, l'idée est d'être aussi informatif que possible pour rapprocher au mieux les données originales en considérant le reste comme bruit. Leur but principal est de prendre en compte la variabilité des données de manière non supervisées et ces représentations ne sont pas dépendantes de l'application. Par contre, les méthodes discriminatives fournissent une meilleure reconstruction car ces représentations sont supervisées. Ces méthodes sont dépendantes de l'application, mais sont par ailleurs plus efficaces d'un point de vue implémentation. De plus, elles permettent une meilleure classification que les méthodes reconstructives. Ainsi appliquées à la détection d'objets mobiles, les méthodes reconstructives s'appliquent plus à reconstruire de façon non supervisée une image de fond que de fournir une bonne classification des pixels dans l'étape de détection. De plus, ces méthodes considèrent que les objets mobiles ont une contribution faible lors de l'initialisation, alors que cette hypothèse est seulement vérifiée si les objets mobiles à détecter sont petits par rapport à la taille de l'image ou loin de la caméra. Le seul avantage de modéliser le fond avec une approche reconstructive est l'aspect non-supervisé. D'où notre motivation à nous intéresser à l'apport d'une approche discrimina-

1. Ce chapitre est basé sur nos publications [177] et [359].

tive et d'une approche mixte par réduction de dimension dans la détection d'objets mobiles dans des séquences vidéo. Plus particulièrement, nous avons proposé, dans un premier temps, la modélisation du fond par une approche discriminative de réduction de dimension. Cette approche discriminative permet une initialisation supervisée robuste de l'image du fond et une classification robuste des pixels. Dans un second temps, nous avons proposé l'application d'une approche mixte de réduction de dimension. L'objectif étaient d'obtenir des méthodes robustes aux variations lumineuses lors de la détection d'objets mobiles. Les contributions de cette partie sont donc les suivantes :

— **Un état de l'art sur les approches reconstructives de réduction de dimension (Section 2)** : Dans cette section, l'Analyse en Composante Principales (ACP), l'Analyse en Composantes Indépendantes (ACI) seront brièvement rappelées dans le contexte de la modélisation de fond. Nous discuterons de leurs avantages et inconvénients ainsi que de leur robustesse dans le cas de fonds comportant des variations lumineuses.
Publication relative : 1 journal RPCS [69])

— **Modélisation de fond par une approche discriminative de réduction de dimension (Section 3)** : Les approches reconstructives ont été jusqu'alors les seules utilisées pour modéliser le fond car elles permettent d'extraire les composantes principales de façon robuste et parce qu'elles ne nécessitent aucune connaissance *a priori* sur les données. Cependant, elles ne sont pas adaptées à la classification nécessaire lors de la détection et émettent une hypothèse sur la taille des objets mobiles. Dans ce contexte, nous avons proposé d'utiliser une approche discriminative de réduction de dimension appelée Incremental Maximum Margin Criterion. Cette approche permet de s'affranchir de la taille des objets mobiles par son aspect discriminatif et d'être employée de façon effective par son aspect incrémental.
Publications relatives : 1 journal (Machine Vision and Applications [177]), 2 conférences (ACCV [358], IVPCV [176])

— **Modélisation de fond par une approche mixte de réduction de dimension (Section 4)** : Dans cette section, nous proposons de combiner une approche reconstructive (ACP) avec une approche discriminative (LDA). L'objectif est, dans un premier temps, de modéliser de façon robuste le fond avec l'avantage d'une méthode non-supervisée et, dans un second temps, de classer les pixels en fond ou objet mobile avec une méthode adaptée à la classification.
Publication relative : 1 Conférence (IPCV [359]).

Pour la partie sur l'approche discriminative, un site web a été mis en ligne : IMMC[1]. Ce site permet de visualiser les résultats obtenus.

1. http://sites.google.com/site/immcforegrounddetection/

2 Approche reconstructive de réduction de dimension pour la modélisation de fond : Une vue d'ensemble

Dans la littérature, seulement les approches reconstructives de réduction de dimension étaient employées pour le modélisation de fond avant nos travaux. La première approche développée par [387] consiste à appliquer l'Analyse en Composantes Principales (ACP) sur N images pour construire un modèle de fond qui est constitué d'une image correspondant à l'image moyenne et d'une matrice de projection comprenant les p vecteurs propres de l'ACP. De cette façon, la détection des objets mobiles est effectuée en calculant la différence entre l'image courante et l'image reconstruite. Suivant la même idée, Yamazaki et al. [579] et Tsai et al. [523] emploient une Analyse en Composantes Indépendantes (ACI).

2.1 Réduction de dimension par ACP (RD-ACP)

Oliver et al. [387] ont proposé de modéliser le fond en utilisant l'Analyse en Composante Principale (ACP). A partir de N images, Oliver et al. [387] ont calculé une image moyenne μ_B and la matrice de covariance associée C_B. Cette matrice est ensuite diagonalisée de la façon suivante :

$$L_B = \Phi_B C_B \Phi_B^T \tag{4.1}$$

où Φ_B est la matrice contenant les vecteurs propres et L_B est la matrice diagonale comportant les valeurs propres. Pour réduire la dimension, seulement M vecteurs propres ($M < N$) sont conservés par l'ACP. Les M plus grandes valeurs propres sont stockées dans la matrice L_M et les M vecteurs correspondant dans la matrice Φ_M.

Ensuite, l'image courante I_t est approximée par l'image moyeenne et par la somme pondérée des fonds "propres" Φ_M. Ainsi, les coordonnées dans l'espace des images "propres" de l'image courante I_t est calculée de la façon suivante :

$$w_t = \left(I_t - \mu_B\right)^T \Phi_M \tag{4.2}$$

Quand w_t est projeté sur l'espace image, une image du fond est ainsi reconstruite comme suit :

$$B_t = \Phi_M w_t^T + \mu_B \tag{4.3}$$

Ensuite, la classification fond/objet est effectué de la façon suivante :

$$|I_t - B_t| > T \tag{4.4}$$

où T est un seuil constant.

Ce modèle que nous avons appelé RD-ACP fournit un modèle robuste de la fonction de distribution de la probabilité du fond mais pas des objets mobiles et cela tant que les objets mobiles n'ont pas une contribution significative dans la séquence d'apprentissage. Ainsi, la première limitation de ce modèle réside dans la taille des objets mobiles qui doit être petite

par rapport à la taille de l'image et qui de plus ne doivent pas apparaitre sur une longue période dans la séquence d'apprentissage. La seconde limitation apparait lors de la mise à jour du modèle. En effet, il est très coûteux en temps de calcul de mettre à jour l'ACP en mode "batch". De plus, sans un mécanisme d'analyse robuste, les données aberrantes (les objets mobiles) peuvent être absorbées dans le modèle du fond. La troisième limitation est l'application limitée quasi exclusivement à l'intensité du pixel en niveau de gris car il n'est pas aisé de l'étendre aux cas multi-canaux comme l'espace couleur RGB. Enfin, une autre limitation concerne la représentation unimodale qui ne permet pas de prendre en compte les variations lumineuses multimodales. Différentes solutions ont été apportées dans la littérature pour résoudre ces différentes limitations et peuvent être classées comme suit [69] :

1. **Robustesse à la taille des objets :** Xu et al. [574] ont proposé d'appliquer de façon récursive un processus de compensation d'erreur (*Recursive Error Compensation (REC)*) qui réduit l'influence des objets mobiles sur le modèle de fond. Un seuil adaptatif est par ailleurs employé lors de la classification fond/objets mobiles. Cette méthode est plus robuste que la méthode originale développée par Oliver et al. [387] lors de la présence d'objets mobiles de grande taille. Une autre approche développée par Kawabata et al. [275] consiste en une projection itérative (*Iterative Optimal Projection (IOP)*) pour estimer le fond en temps réel dans des scènes comportant des fonds dynamiques avec des objets mobiles. Dans un premier temps, les images de fond sont collectées et sont stockées sous forme d'eigen-structure dans une base. Après cette étape d'initialisation, une nouvelle image est acquise et projetée pour estimer le fond. Comme le fond ainsi estimé est contaminé par les objets mobiles, les objets mobiles sont obtenus par différence avec l'image de fond précédente. Alors, les zones de l'image estimées et contaminées par les objets mobiles sont remplacées par les zones correspondantes de l'image de fond précédente. Kawabata et al. [275] ont démontré que ce processus itératif converge vers une image de fond correcte. D'une autre façon, Quivy et Kumazawa [421] ont généré les images de fond en employant un algorithme appelé Nelder-Mead Simplex et un processus de masque dynamique. Cette approche remplace l'étape de projection/reconstruction de la version développée par Oliver et al. [387] par une génération directe de l'image de fond. Les expérimentations montrent que cette méthode est plus robuste que les méthodes précédentes, c'est-à-dire, la version originale [387], REC [574] et IOP [421], pour des objets mobiles de grande taille et des objets mobiles rapides.

2. **Robustesse aux outliers :** Pour la mise à jour, plusieurs auteurs ont proposés des versions incrémentales de PCA plus ou moins robustes aux outliers. La version incrémentale développée par Rymel et al. [441] nécessite peu de temps mais l'image de fond est contaminée lors de la mise à jour par les objets mobiles. Pour résoudre cela, Li et al. [317] ont proposé une version incrémentale plus robuste aux outliers.

Cependant, il assigne le même poids à toutes les images lors de la mise à jour du modèle. Ainsi, des images avec ou sans objets mobiles contribuent de la même façon à la reconstruction de l'image de fond. La conséquence est une contamination progressive de l'image de fond. Dans ce contexte, Skocaj et al. [482] ont proposé un algorithme avec des poids incrémental et robuste. Les poids assignés sont différents suivant l'image courante pour obtenir une meilleure image de fond. Cependant, tous les poids sont appliqués de la même façon à tous les pixels de l'image courante. Ce qui a pour conséquence que chaque partie ou pixel de l'image courante contribue avec la même importance à la construction de l'image de fond. Pour obtenir une précision de l'ordre du pixel pour les poids, Zhang et Zhuang [622] ont adopté une sélection adaptive des poids. Cette méthode est plus précise car elle assigne un poids différent à chaque pixel de l'image courante. Cette méthode est plus robuste que IPCA [441]. Wang et al. [541] ont employé une méthode similaire avec un algorithme séquentiel de Karhunen-Loeve. Zhang et al. [621] ont quant à eux amélioré cette approche en proposant une approche adaptative. Toutes ces approches incrémentales évitent la décomposition en vecteurs propres et valeurs propres de la matrice de covariance compléte lors de la mise à jour de celle-ci en effectuant une plus petite décomposition avec un temps de calcul moindre. Cependant, ces méthodes maintiennent quand même toute cette structure. Dans ce contexte, Li et al. [313] ont développé une mise à jour rapide et récursive en évitant d'avoir à stocker toute cette structure. Cette méthode, bien qu'approximative de la version exacte, donne de meilleurs résultats que IPCA [441] et IRPCA [317] avec un nombre de frames par seconde plus élevé. La figure 4.1 montre l'ordre d'adaptabilité de certaines de ces méthodes.

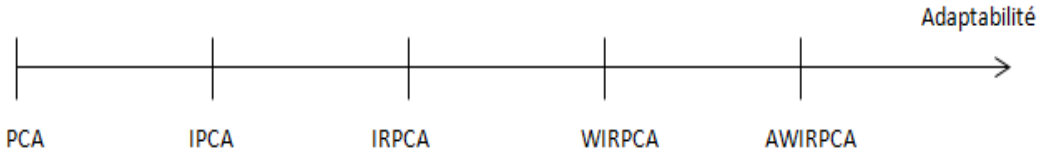


FIGURE 4.1 – Vue de l'ordre d'adaptabilité de certaines versions incrémentales ACP.

3. **Contraintes "niveau de gris" et "pixel" :** Wu et al. [566] ont combiné le modèle ACP et le modèle SG [562]. Le premier permet d'être robuste aux variations lumineuses et les second permet d'avoir des informations couleur pour chaque pixel. Ainsi, les variations de chromaticité sont détectés et permettent ainsi de supprimer les ombres. D'une autre manière, Han et Jain [234] ont employé un algorithme incrémental appelé *Weighted Incremental 2-Dimensional Principal Component Analysis*. Cette approche 2DPCA est plus efficace que l'approche standard. De plus, l'aspect incrémental permet de mettre à jour les vecteurs propres et les valeurs propres pour

prendre en compte les variations du fond. Cet algorithme peut employé 3 canaux (RGB) ou 4 canaux (RGB+IR). Les expérimentations [234] ont montré des améliorations importantes de la robustesse en présence de fonds dynamiques et d'ombres. Pour supprimer la contrainte du niveau "pixel", Zhao et al. [634] ont utilisé des blocs spatio-temporels. De plus, leur méthode emploie un algorithme appelé *Candid Covariance Free Incremental Principal Components Analysis algorithm (CCIPCA)* qui converge relativement rapidement avec une complexité moindre que l'algorithme IPCA [441]. Les expérimentations [634] ont montré la robustesse de la méthode en présence de bruit et de variations lumineuses rapides.

4. **Robustesse aux variations lumineuses :** Dong et al. [161] ont proposé d'employer une méthode de réduction de dimension avec plusieurs sous-espaces pour prendre en compte différentes variations lumineuses. L'espace des caractéristiques est organisé sous forme de clusters qui représentent différents conditions d'illumination. Une ACP locale (*Local Principle Component Analysis (LPCA)*) est employée pour apprendre séparément chaque cluster. Quand une nouvelle image arrive, l'algorithme sélectionne automatiquement le sous-espace qui présente la condition d'illumination la plus proche de l'image courante. Les expérimentations montrent que cette méthode est plus robuste que la méthode originale [387] et que les MOG [490] en présence notamment de variations lumineuses soudaines. De façon similaire, Kawanishi et al. [276] ont généré des images de fond qui sont représentatives des différentes conditions lumineuses de la scène au cours du temps. Cette méthode collecte ainsi un grand nombre d'images représentatives observées sur de longues séquences, les classifient en termes de leur apparition au cours de la journée, et applique l'ACP sur les images adéquates pour retrouver l'image de fond.

Le tableau 4.1 montre une vue d'ensemble de ces différentes versions de PCA développées pour la détection d'objets mobiles.

Remarque : Il est à noter que ces méthodes ACP robustes sont robustes par suppression des outliers au fur et à mesure de leur apparition alors que celles développées dans le chapitre 5 le sont par séparation de la matrice des observations en une matrice de rang faible (qui contient le fond) et une matrice parcimonieuse (qui contient les outliers).

2.2 Réduction de dimension par ACI (RD-ACI)

L'Analyse en Composantes Indépendantes (ACI) généralise la technique d'Analyse en Composantes Principales (ACP) et a montré son efficacité dans l'extraction de caractéristiques. Quand un mélange de signaux indépendants est observé, l'ACI décompose les signaux sources du mélange observé sans connaissances *a priori* sur la façon dont les signaux

Améliorations	**Méthodes**	**Auteurs - dates**
Robustesse	Recursive Error Compensation (REC)	Xu et al. (2008) [574]
à la	Iterative Optimal Projection (IOP)	Kawabata et al. (2006) [275]
taille des objets	Simplex Algorithm (SA)	Quivy and Kumazawa (2011) [421]
Robustesse	Incremental PCA (IPCA)	Rymel et al. (2004) [441]
aux	Incremental and robust PCA (IRPCA)	Li et al. (2003) [317]
outliers	Weighted Incremental and Robust PCA (WIRPCA)	Skocaj et al. (2003) [482]
	Adaptive Weight Selection for Incremental and Robust PCA (AWIRPCA)	Zhang and Zhuang (2007) [622]
	Sequential Karhunen-Loeve algorithm (SKL)	Wang et al. (2006) [541]
	Adaptive Sequential Karhunen-Loeve algorithm (ASKL)	Zhang et al. (2009) [621]
	Fast Recursive Maintenance (FRM)	Li et al. (2006) [313]
Suppression	PCA - Single Gaussian(PCA-SG)	Wu et al. (2009) [566]
de la contrainte	Weighted Incremental 2PCA (WI2DPCA)	Han and Jain (2007) [234]
niveau de gris et pixel	Candid Covariance Incremental PCA(CCIPCA)	Zhao et al. (2008) [634]
Robustesse aux	Local Principle Component Analysis on Clusters (LPCA-C)	Dong et al. (2010) [161]
variations lumineuses	Local Principle Component Analysis on Separated Sequences(LPCA-SS)	Kawanishi et al. (2009) [276]

TABLE 4.1 – Vue d'ensemble des différentes versions de PCA développées pour la détection d'objets mobiles.

sont mélangés. L'hypothèse émise est que les vecteurs d'observation $X = (x_1, x_2, ..., x_M)^T$ peuvent être décomposés en une superposition linéaire de vecteurs indépendants inconnus $S = (s_1, s_2, ..., s_N)^T$:

$$X = AS \tag{4.5}$$

où E est une matrice appelée matrice de mélange ($M \times N$). L'objectif d'ACI est de trouver la matrice W telle que les vecteurs résultats :

$$Y = WX \tag{4.6}$$

donne les vecteurs indépendants S, permutés de façon probabiliste et renormés. W est la matrice inverse de A.

Appliqué à la modélisation de fond, le modèle par RD-ACI se décrit de la façon suivante :

$$Y = WX_t \tag{4.7}$$

$X_t = (x_1, x_2)^T$ est le vecteur de mélange de taille $2 * K$ pour lequel $K = M * N$, où M et N sont respectivement le nombre de colonnes et de lignes des images composant la séquence. $x_1 = (x_{11}, x_{12}, ...x_{1K})$ est la première image de la séquence qui peut contenir ou pas des objets mobiles et $x_2 = (x_{21}, x_{22}, ...x_{2K})$ est la seconde image de la séquence qui contient des objets mobiles. $W = (w_1, w_2)^T$ est la matrice de démixage pour laquelle $w_i = (w_{i1}, w_{i2})$ avec $i = 1, 2$. $Y = (y_1, y_2)^T$ est constitué des signaux sources, c'est-à-dire l'un des signaux

correspondant au fond et l'autre à l'objet mobile avec $y_i = (y_{i1}, y_{i2}, ...y_{ik})$. Plusieurs méthodes peuvent être employées pour déterminer la matrice W. En effet, Yamazaki et al. [579] un algorithme d'apprentissage à base de neurones [56] alors que Tsai and Lai [279] emploie un algorithme appelé "Particle Swarm Optimization" (PSO) [450]. Une fois la matrice W déterminée, il existe deux possibilités pour générer l'image de fond et le masque des objets mobiles :

— La première possibilité pour laquelle x_1 contient des objets mobiles comme dans Yamazaki et al. [579]. Dans ce cas, les masques des objets mobiles pour les images x_1 et x_2 sont obtenus en seuillant respectivement y_1 et y_2. L'image du fond est quant à elle obtenue en remplaçant les régions détectées comme "objets mobiles" dans x_1 par les régions correspondantes classées comme "fond" dans x_2.
— La seconde possibilité pour laquelle x_1 ne contient pas d'objets mobiles comme dans Tsai and Lai [523]. Dans ce cas, le masque des objets mobiles pour l'image x_2 est obtenue en seuillant y_2. L'image de fond est y_1.

Ce modèle a été testé sur des scènes de trafic routier dans [579] montrant une bonne robustesse en présence de variations lumineuses. Dans [523], l'algorithme a été testé sur des scènes d'intérieurs en présence également de variations lumineuses importantes.

2.3 Réduction de dimension par INMF (RD-INMF)

L'objectif de NMF de rang R est de décomposer la matrice de données $V \in \mathbf{R}^{P \times Q}$ en deux matrices $W \in \mathbf{R}^{P \times R}$ et $H \in \mathbf{R}^{R \times Q}$ respectivement appelées matrice de mélange et matrice d'encodage :

$$V \approx WH \tag{4.8}$$

Ainsi, NMF tente de trouver une approximation de la factorisation qui minimise l'erreur de reconstruction. Différentes fonctions coûts sont utilisées dans la litterature comme l'erreur quadratique principalement employée de part sa simplicité :

$$F = ||V - WH||^2 = \sum_{i=1}^{P} \sum_{j=1}^{Q} \left(V_{ij} - (WH)_{ij} \right)^2 \tag{4.9}$$

où les indices ij indiquent le ij^{ieme} élement de la matrice concernée. Dans le cadre de la modélisation de fond, Bucak et al. [83] ont proposé un algorithme incrémental de NMF. Le fond est ainsi initialisé avec N images d'apprentissages. Ensuite, V est le vecteur colonne correspondant à une matrice de taille $(P \times Q) \times N$ où P et Q sont le nombre de colonnes et de lignes. Les matrices W et H sont mises à jour de façon incrémentale. La classification fond/objets mobiles est effectuée en seuillant l'erreur entre le modèle du fond et la projection de l'image courante. RD-INMF montre des performances similaires dans le cas de fonds dynamiques et de variations lumineuses que RD-IPCA [317].

2.4 Réduction de dimension par Incremental Rank-Tensor (RD-IRT)

Les précedentes méthodes de réduction de dimension considèrent l'image comme un vecteur perdant ainsi l'information spatiale locale. Pour résoudre cela, Li et al. [315] ont employé une méthode de réduction de dimension basée sur les tenseurs et appelée "incremental rank-($R1$,$R2$,$R3$)". L'algorithme construit un tenseur de rang inférieur pour lequel la moyenne, les valeurs propres et les vecteurs propres sont mises à jour de façon adaptative incrémentalement. Soit $G = \{BM_q \in \mathbf{R}^{M \times N}\}_{q=1,2,..,t}$ la séquence d'apprentissage avec la q-ième image notée BM_q, et M et N sont le nombre de colonnes et de lignes des images constituant la séquence. Soit $p(x, y)$ le x-ième and y-ième pixel de l'image. Le modèle correspondant à $A = \{BM_q(x, y) \in \mathbf{R}^{I_1 \times I_2 \times t}\}_{q=1,2,...,t}$ ($I_1 = I_2 = 5$ correspondent à un voisinage K du pixel $p(x, y)$ avec $K = I_1 I_2 - 1 = 24$) consiste à l'espace de dimensions (R_1,R_2,R_3) correspondant aux tenseurs des trois modes suivants : le mode-n correspondant à la matrice de projection suivant les colonnes $X^{(n)} \in \mathbf{R}^{I_n \times R_n}$, le mode-3 correspondant à la matrice de projection suivant les lignes $Y^{(3)} \in \mathbf{R}^{I_n I_2 \times R_3}$, la colonne des moyennes $\overline{L}^{(1)}$ et $\overline{L}^{(2)}$, et la ligne des moyennes $\overline{L}^{(3)}$ du mode-3 de la matrice $A_{(3)}$. Soit un voisinage $I_{t+1}(x, y) \in \mathbf{R}^{I_1 \times I_2 \times 1}$ centré sur le x-ième et le y-ième pixel $p(x, y)$ de l'image courante $I_{t+1} \in \mathbf{R}^{M \times N \times 1}$, la distance $RM(x, y)$ déterminée à partir des trois erreurs de reconstruction des trois modes entre $I_{t+1}(x, y)$ et le modèle est ensuite seuillé pour classer les pixels en fond/objets mobiles comme suit :

$p(x, y)$ est classé en "fond" si $exp(\frac{-RM(x,y)^2}{2\sigma^2}) > T$
$p(x, y)$ est classé en "objets mobiles" sinon

où σ est un facteur d'échelle et T est un seuil fixé et constant. Ensuite, la nouvelle image de fond $BM_{t+1}(x, y)$ à l'instant $t + 1$ est définie comme suit :

$BM_{t+1}(x, y) = H(x, y)$ if $p(x, y)$ est classée comme "objets mobiles"
$BM_{t+1}(x, y) = I_{t+1}(x, y)$ sinon

où $H(x, y) = (1 - \alpha)MB_t(x, y) + \alpha I_{t+1}(x, y)$, MB_t est la matrice moyenne de $BM_{1:t}$ à l'instant t et α est le coefficient d'apprentissage. Enfin, le modèle est mis à jour de façon incrémentale et ainsi de suite. RD-IRT donne de meilleures performances en présence de bruit et de variations lumineuses que RD-IPCA [317].

2.5 Discussion

Grâce à ces différentes améliorations au cours du temps, la méthode RD-ACP est la méthode de réduction de dimension la plus utilisée dans la modélisation de fond. Elle est de plus fiable quand des variations lumineuses apparaissent. Enfin, elle peut également être implémentée de façon incrémentale et robuste. RD-ACI et RD-INMF montrent des performances similaires à RD-ACP. Par contre, RD-IRT montre de meilleurs résultats grâce à sa prise en

compte des informations spatiales contenues dans les blocs. Cependant, toutes ces méthodes reconstructives de réduction de dimension présentent plusieurs inconvénients comme expliqué dans la Section 1.

Dans la section suivante, nous développons l'utilisation d'une méthode discriminative de réduction de dimension pour la modélisation de fond.

3 Modélisation de fond par une approche discriminative de réduction de dimension

3.1 Approches discriminatives de réduction de dimension : Une vue d'ensemble

Les trois approches discriminatives de réduction de dimension les plus utilisées sont l'Analyse Discriminante Linéaire (LDA) [508], l'Analyse Canonique des Corrélations (CCA) [238] et l'Analyse suivant le Critère de Marge Maximale (MMC) [309]. La méthode LDA projette les données dans un espace de plus petite dimension de telle façon que le ratio entre la distance inter-classe et la distance intra-classe est maximisé. L'objectif est d'obtenir une discrimination maximale. La méthode CCA est une méthode statistique multivariées qui facilite l'étude des corrélations parmi des ensembles de variables multiples dépendantes et des variables multiples indépendantes. La méthode MMC cherche les axes de projection sur lesquels les données des différentes classes sont loin des autres tout en faisant que les données de la même classe soient proches les unes des autres. Le principal inconvénient de LDA est que cette méthode est sensible quand il apparait des singularités. La méthode MMC est plus robuste dans ce cas [579]. De plus, le temps de calcul de la méthode MMC est inférieure à celle de la méthode LDA parce que la méthode MMC ne demande pas d'effectuer des calculs de matrice inverse. De surcroit, des récentes recherches ont montré que la méthode MMC offre un cadre adaptée à la classification [531][591] et que dans ce cas, elle donne de meilleurs résultats que la méthode LDA. Le vrai seul inconvénient des approches discriminatives est qu'elles nécessitent des algorithmes de type "batch", c'est-à-dire que toutes les données doivent être connues à l'avance et fournies à l'algorithme en une seule fois. Ensuite, Yan et al. [580] ont proposé un algorithme incrémental de la méthode MMC appelé IMMC (Incremental Maximum Margin Criterion). Pour toutes ces raisons, nous avons choisi de développer une méthode de suppression de fond par cette approche discriminative incrémentale.

3.2 Réduction de dimension par la méthode de Critère de Marge Maximale

Nous rappelons brièvement le principe de l'algorithme développé par Yan et al. [580]. L'algorithme IMMC est utilisé pour l'initialisation et la mise à jour du fond dans la Section 3.3. Soient les vecteurs de données $u(1), u(2), ..., u(N)$ de dimension d, et U une matrice

constituée de $u(i)$ à la colonne i. La méthode MMC cherche les axes de projection sur lesquels les données des différentes classes sont loin des unes des autres tout en faisant que les données de la même classe soient proches les unes des autres. L'objectif est de réaliser la discrimination maximale et le vecteur de dimension inférieure est calculé comme suit $y = W^T u$ où $W \in \mathbf{R}^{d \times p}$ est la matrice de projection de l'espace d'origine de dimension d vers l'espace de dimension inférieure dont la dimension est p. Ainsi, MMC cherche à maximiser le critère suivant :

$$J(W) = W^T(S_b - S_w)W \tag{4.10}$$

où

$$S_b = \sum_{i=1}^{c} p_i(m_i - m)(m_i - m)^T \tag{4.11}$$

$$S_w = \sum_{i=1}^{c} p_i E(u_i - m_i)(u_i - m_i)^T \tag{4.12}$$

sont respectivement la matrice inter-classe et la matrice intra-classe. c est le nombre de classes, m est la moyenne sur toutes les données, m_i est la moyenne sur les données appartenant à la classe i et p_i est la probabilité *a priori* d'une donnée d'appartenir à la classe i. La matrice de projection W est obtenue en résolvant la décomposition en vecteurs propres :

$$(S_b - S_w)w = \lambda w \tag{4.13}$$

Pour maximiser de façon incrémentale le critère MMC, Yan et al. [580] ajoute la contrainte que les vecteurs de W soient des vecteurs unitaires, c'est-à-dire que $W = [w_1, w_2, ...w_p]$ et $w_k^T w_k = 1$. Ainsi, l'optimisation de $J(W)$ est transformée comme suit :

$$max \sum_{k=1}^{p} w_k^T(S_b - S_w)w_k \tag{4.14}$$

avec la contrainte que $w_k^t w_k = 1$ avec $k = 1, 2, ..., p$. W est constitué des k premiers vecteurs propres de la matrice $S_b - S_w$ et les vecteurs colonnes de W sont orthogonaux l'un à l'autre. Ainsi, le problème consiste à mettre à jour les p vecteurs propres de $S_b - S_w$ de façon incrémentale.

1- Mise à jour des p premiers vecteurs propres

Soit $C = S_b + S_w$ la matrice de covariance, alors nous avons $J(W) = W^T(2S_b - C)W$, $W \in \mathbf{R}^{d \times p}$. Maximiser $J(W)$ signifie trouver les p premiers vecteurs propres de $2S_b - C$. La matrice inter-classe à l'itération n après la mise à jour des n premières données peut être écrite comme suit :

$$S_b(n) = \sum_{j=1}^{c} p_j(n)(m_j - m(n))(m_j(n) - m(n))^T \tag{4.15}$$

et

$$S_b = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} S_b(i) \tag{4.16}$$

D'autre part,

$$C = E(u(n) - m)(u(n) - m)^T \tag{4.17}$$

$$= \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (u(n) - m(n))(u(n) - m(n))^T \tag{4.18}$$

$2S_b - C$ doit avoir les mêmes vecteurs propres que $2S_b - C + \theta I$ où θ est nombre réel positif et $I \in R^{d\times d}$. De () et () , on obtient :

$$2S_b - C + \theta I = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} A(i) = A \tag{4.19}$$

où $A(i) = 2S_b(i) - (u(i) - m(i))(u(i) - m(i))^T + \theta I$, $A = 2S_b - C + \theta I$. La forme générale des vecteurs propres est $Ax = \lambda x$, où x est le vecteur propre de la matrice A correspondant à la valeur propre λ. En remplaçant la matrice A avec la matrice MMC à l'itération n, on obtient un vecteur propre approximé avec $\nu = \lambda x$.

$$\begin{aligned} \nu(n) = \frac{1}{n} \sum_{i=1}^{n} (2 \sum_{j=1}^{c} p_j(i)\Phi_j(i)\Phi_j(i)^T \\ - (u(i) - m(i))(u(i) - m(i))^T + \theta I)x(i) \end{aligned} \tag{4.20}$$

où $\Phi_j\,(i) = m_j\,(i) - m\,(i)$, $v\,(n)$ est la $n^{iéme}$ estimation de v et $x\,(n)$ est la $n^{iéme}$ estimation de x. Une fois l'estimation de ν obtenue, le vecteur propre x peut être directement calculé comme suit : $x = \nu/||\nu||$. Soit $x\,(i) = \nu\,(i-1)\,/||\nu\,(i-1)\,||$, alors la formulation incrémentale est la suivante :

$$\nu(n) = \frac{n-1}{n}\nu(n-1) + \frac{1}{n}(2 \sum_{j=1}^{c} p_j(n)\alpha_j(n)\Phi_j(n) - \beta(n)(u(n) - m(n)) + \theta\nu(n-1))/||\nu(n-1)|| \tag{4.21}$$

où $\alpha_j(n) = \phi_j(n)^T\nu(n-1)$ et $\beta(n) = (u(n) - m(n))^T\nu(n-1)$, $j = 1, 2, ..., c$. Pour l'initialisation, $\nu(0)$ est égale à la première donnée.

2- Mise à jour des autres vecteurs propres

Pour calculer le $(j+1)^{iéme}$ vecteur propre, sa projection est soustraite de la $j^{iéme}$ estimation du vecteur propre à partir des données,

$$u_{1_n}^{j+1}(n) = u_{1_n}^{j}(n) - (u_{1_n}^{j}(n)^T\nu^j(n))\nu^j(n) \tag{4.22}$$

où $u_{1_n}^1(n) = u_{1_n}(n)$. Le même principe est utilisé pour mettre à jour $m_i^j(n)$ et $m^j(n)$, $i = 1, 2, ..., c$. Comme $m_i^j(n)$ et $m^j(n)$ sont des combinaisons linéaires de $x_{l_i}^j(i)$, où $i = 1, 2, ..., k$, et $l_i \in \mathbf{1}, \mathbf{2}, ..., \mathbf{C}$. Φ_i sont des combinaisons linéaires de m_i and m, par convention, seulement Φ est mis à jour à chaque itération comme suit :

$$\Phi_{l_n}^{j+1}(n) = \Phi_{l_n}^j(n) - (\Phi_{l_n}^j(n)^T \nu^j(n))\nu^j(n) \tag{4.23}$$

De cette façon, le coût de calcul de l'orthonormalisation est évitée et l'orthogonalité est renforcée quand la convergence est atteinte.

3.3 Modélisation de fond via Incremental Maximum Margin Criterion

1- Description générale de la méthode basée sur IMMC

La figure 4.2 montre une vue générale de la méthode proposée. La méthode de suppression de fond qui est basée sur IMMC comporte les étapes suivantes : (1) Initialisation de l'image de fond avec MMC à partir de N images. (2) Classification des pixels comme "fond" ou "objets mobiles". (3) Mise à jour de l'image du fond via IMMC. Les étapes (2) et (3) sont répétées au cours du temps. Ces différentes étapes sont développées dans les sections suivantes et l'algorithme est donné en détail dans la Section 3.4.

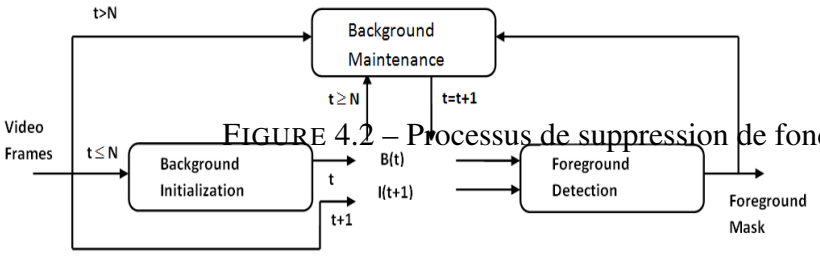


FIGURE 4.2 – Processus de suppression de fond par IMMC

2- Initialisation du fond

Soit la séquence vidéo d'apprentissage $S = \{I_1, ...I_N\}$ où I_t est l'image à l'instant t et N est le nombre d'images d'apprentissage. Chaque pixel (x, y) est caractérisé par son intensité en niveau de gris et considérons que les images vérité terrain sont connus pour la séquence d'apprentissage, c'est-à-dire que nous connaissons pour chaque pixel la classe à laquelle il appartient. Ces images vérités terrain peuvent être obtenues par une segmentation manuelle ou par l'utilisation d'une méthode de base comme la moyenne ou la médiane. Dès lors, nous avons :

$$S_b = \sum_{i=1}^{c} p_i(\overline{I}_i - \overline{I})(\overline{I}_i - \overline{I})^T, S_w = \sum_{i=1}^{c} p_i E(I_i - \overline{I}_i)(I_i - \overline{I}_i)^T \tag{4.24}$$

où $c = 2$, $\overline{I}$ est la moyenne des intensités du pixel x,y sur la séquence d'apprentissage et $\overline{I_i}$ est la moyenne des pixels appartenant à la classe i et p_i est la probabilité *a priori* des pixels à appartenir à la classe i avec $i \in \{Background, Foreground\}$. Les probabilités *a priori* sont obtenues à partir de la carte des classes, c'est-à-dire que les pixels identifiés comme "fond" ont une grande probabilité initiale $p_{Background}$ alors que les pixels identifiés comme "objets mobiles" ont une grande probabilité initiale $p_{Foreground}$. En pratique, $p_{Background}$ et $p_{Foreground}$ sont initialisés respectivement à 0.9 et 0.1 pour les pixels identifiés comme "fond" et réciproquement pour les pixels identifiés comme "objets mobiles". Ensuite, nous pouvons utiliser l'algorithme "batch" pour obtenir les premiers vecteurs propres qui correspondent au fond. Les valeurs propres correspondantes sont stockées dans la matrice L_M et les vecteurs propres dans ma matrice Φ_M. Une fois que les vecteurs propres stockés dans la matrice Φ_M sont obtenus et que la moyenne $\overline{I}$ aussi, l'image courante I_t peut être approximée comme l'image moyenne et la somme pondérée des vecteurs propres Φ_M. Ainsi, les coordonnées dans l'espace des vecteurs propres de l'image courante I_t sont calculées comme suit :

$$w_t = (I_t - \overline{I})^T \Phi_M \tag{4.25}$$

Quand w_t est projetée sur l'espace image, l'image du fond est obtenue comme suit :

$$B_t = \Phi_M w_t^T + \overline{I} \tag{4.26}$$

B_N est la première image de fond obtenue à partir des N images de la séquence d'apprentissage. Dans les expérimentations, nous avons choisi N compris 30 et 100 pour la plupart des scènes. L'influence de N est étudiée qualitativement dans la Section 3.5 et quantitativement dans la Section 5.

3- Classification des pixels

Après la reconstruction de l'image du fond, la classification des pixels peut être effectuée. Elle consiste à classer les pixels en "fond" ou "objet mobile" en comparant l'image de fond et l'image courante. En général, une simple soustraction entre ces deux images est effectuée. En suivant ce principe, la classification des pixels est effectuée comme suit :

$$F_t(x,y) = \begin{cases} 1, & \text{si } d\left(I_t(x,y), B_t(x,y)\right) > T \\ 0, & \text{sinon} \end{cases} \tag{4.27}$$

où $I_t(x,y)$, $B_t(x,y)$ et $F_t(x,y)$ sont respectivement l'image courante, l'image de fond et le masque des objets mobiles à l'instant t. T est un seuil constant qui est déterminé manuellement. $d(.,.)$ est une distance. $F_t(x,y) = 1$ signifie que le pixel est identifié comme objet mobile alors que $F_t(x,y) = 0$ que le pixel est identifié comme "fond".

4- Mise à jour de l'image du fond

Une fois la classification des pixels effectuée, l'image du fond doit être mise à jour pour prendre en compte les changements qui apparaissent dans la scène filmée. Pour cela, nous utilisons l'algorithme IMMC pour mettre à jour les équations 4.21 et 4.23. La classe à laquelle appartient chaque pixel est obtenu à partir du masque courant des objets mobiles. Ce processus est similaire à une mise à jour adaptative. En effet, seuls les pixels identifiés comme "fond" contribuent à la nouvelle image de fond. L'inconvénient est que quand le masque des objets mobiles présente des faux positifs ou des faux négatifs, le processus de mise à jour est affecté par ceux-ci. Pour éviter en partie cela, nous appliquons des opérateurs morphologiques pour éliminer les pixels isolés.

Nous avons remarqué que l'algorithme IMMC peut être appliqué à l'instant $t = 1$ en commençant par la première image et en utilisant sa carte des classes mais cette approche est moins robuste que celle proposée avec N images pour l'inconvénient de la mise à jour expliqué ci-dessus. D'autre part, le processus de mis à jour peut ne pas être appliqué à chaque image mais toutes les M images suivant la fréquence des changements.

3.4 Algorithme SL-IMMC pour la suppression de fond

En combinant l'algorithme MMC pour l'initialisation du fond avec l'algorithme IMMC pour la mise à jour du fond, nous avons proposé l'algorithme SL-IMMC qui est décrit comme suit.

Algorithme de suppression de fond

Initialisation du fond

Etape 1. Calcul de $S_b - S_w$ à partir des N images d'apprentissage et des images vérité terrain correspondantes.

Etape 2. Résoudre la décomposition en vecteur propres dans l'équation 4.13 et stocker les p premiers vecteurs propres et valeurs propres dans Φ_M et L_M.

Etape 3. Générer la première image de fond à partir Φ_M et l'image moyenne $\overline{I}$ (Equations 4.25 et 4.26).

Pour chaque image I_t faire

Classification Fond/Objets mobiles

Etape 4. Calcul du masque des objets mobiles F_t à partir de l'équation 4.27.

Mise à jour du fond

Etape 5. Mettre à jour de façon incrémentale les vecteurs propres correspondants à $2S_b - C$ à partir de l'équation 21 et stocker les p premiers vecteurs propres et valeurs propres dans Φ_M et L_M.

Etape 6. Calcul de la nouvelle image à partir de Φ_M et l'image moyenne $\overline{I}$ (Equations 4.25 et 4.26).

Etape 7. Aller à l'étape 4 jusque la fin de la vidéo.

End

Cet algorithme a été implémenté en C++ et sa performance en termes de complexité et de mémoire est analysée dans notre publication [177] et dans la section 5.2.

3.5 Illustration

Les figures 4.3 et 4.4 montrent respectivement les résultats de la méthode proposée SL-IMMC pour $N = 30$ et $N = 100$. La séquence testée vient d'une caméra fixée sur un poteau et fournie par Sheikh et Shah [461]. Le vent fait bouger le poteau et occasionne donc des mouvements lors de la prise de vue. Dans chaque figure, la première ligne montre les différentes images courantes correspondant aux images 271, 373 et 410. La seconde et la troisième ligne montrent respectivement les images $\overline{I}_{Back}$ et $\overline{I}_{Fore}$ qui correspondent respectivement à l'image moyenne des pixels appartenant à la classe "Background" et à la classe "Foreground". La quatrième ligne montre l'image de fond générée. Finalement, la cinquième ligne présente les détections obtenues à partir de l'équation 4.27. Pour l'initialisation du fond, nous pouvons constater que dans la seconde et la troisième ligne que les images $\overline{I}_{Back}$ contiennent seulement les pixels identifiés comme "fond" et que les images $\overline{I}_{Foreg}$ contiennent seulement les pixels identifiés comme "objets mobiles". La méthode MMC cherche les axes de projection sur lesquels les données de classes différentes sont loin des unes des autres tout en faisant que les données de la même classe soient proches les unes des autres. Ainsi, cette approche discriminative permet d'extraire seulement les vecteurs propres correspondant à l'image $\overline{I}_{Back}$ ou $\overline{I}_{Fore}$ alors qu'une approche reconstructive utilise toutes les données sans séparer les données correspondant au fond et les données correspondant aux objets mobiles.

Influence de N : Pour l'initialisation du fond, nous pouvons constater que la reconstruction de l'image du fond est plus précise quand $N = 100$ que $N = 30$. En particulier pour l'image 410 avec $N = 30$, l'image $\overline{I}_{Back}$ présente un halo noir à gauche de l'image parce que le fond n'est pas visible pendant la période d'apprentissage. Comme la voiture reste à la même place dans l'image courante, il n'y pas de conséquence sur la détection. Pour l'image 271, $N = 30$ est suffisant parce qu'il n'y a pas d'objet mobile présent dans la séquence vidéo d'apprentissage. Finalement, nous pouvons voir que la détection est meilleure quand $N = 100$ pour les images 373 et 410 avec moins de fausses détections positives.

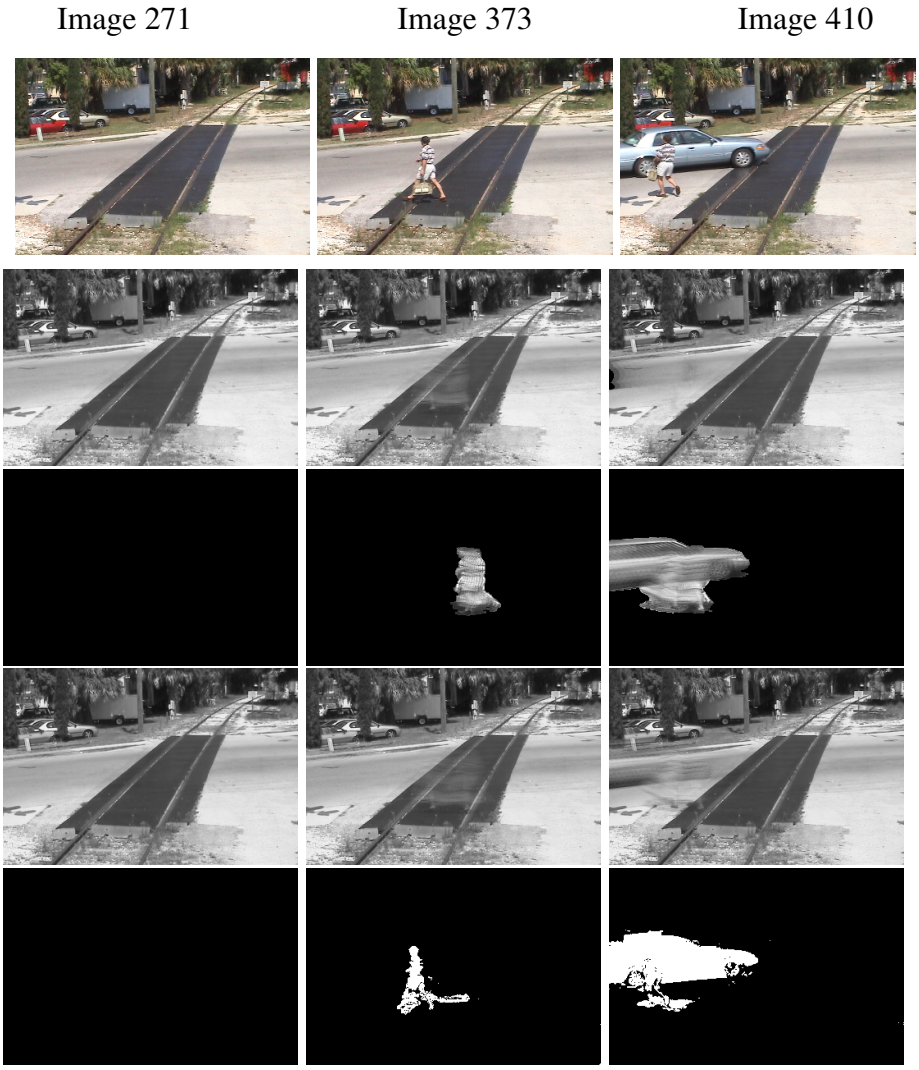


FIGURE 4.3 – Cas N=30 : La première ligne montre les images courantes choisies, la seconde ligne montre les images $\overline{I}_{Back}$, la troisième ligne montre les images $\overline{I}_{Fore}$, la quatrième ligne montre les images du fond et la cinquième ligne montre les résultats de la détection.

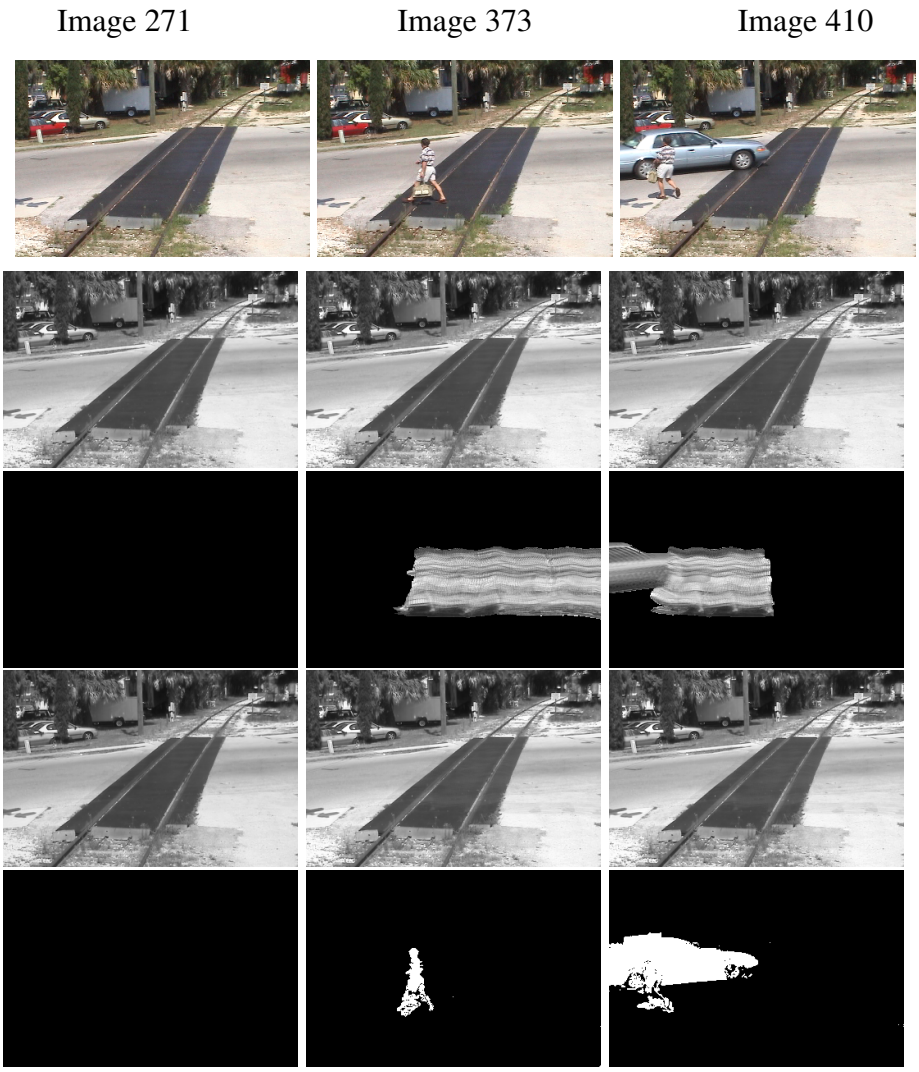


FIGURE 4.4 – Cas N=100 : La première ligne montre les images courantes choisies, la seconde ligne montre les images $\overline{I}_{Back}$, la troisième ligne montre les images $\overline{I}_{Fore}$, la quatrième ligne montre les images du fond et la cinquième ligne montre les résultats de la détection.

4 Approche mixte de réduction de dimension pour la modélisation de fond

4.1 Approches mixtes de réduction de dimension

Les approches mixtes de réduction de dimension ont été utilisées avec succès dans la détection de classes d'objet, la reconnaissance d'objets et la classification d'images. Récemment, Uray et al. [528] ont appliqué avec succès une méthode reconstructive et discriminative à la reconnaissance de visage. Cette méthode combine une approche reconstructive (ACP) avec une approche discriminative (LDA). Dans cette section, nous avons utilisé et validé cette méthode pour la modélisation de fond. Dans ce cadre, nous avons utilisé l'ACP pour modéliser la distribution primaire des pixels, et nous avons émis l'hypothèse que celle-ci correspondait au fond. Ensuite, cette méthode considère que les vecteurs principaux de rang faible contiennent les informations discriminatives. Ainsi, nous appliquons sur ceux-ci l'algorithme LDA pour classer les pixels en fond ou objet mobile.

4.2 Définition du problème

Soit N le nombre d'images d'apprentissage, chacune de ces images contient m pixels stockés en colonnes dans la matrice $X \epsilon R^{m \times N}$, soit $\mu \epsilon R^m$ l'image moyenne, et c le nombre de classes auquel le pixel appartient. Dans notre cas, nous avons deux classes : la classe "fond" et la classe "objets mobiles". Donc $c = 2$. L'objectif est donc de trouver un sous-espace qui contient les pixels des images d'apprentissage de sorte qu'on puisse obtenir une classification pertinente des pixels de l'image courante.

4.3 Modélisation de fond

L'algorithme mixte calcule la nouvelle image de fond à partir de l'image de fond courante et de l'image courante. L'idée est de modéliser le fond à partir de N images de la séquence d'apprentissage en conservant les informations reconstructives et discriminatives dans le modèle. Les aspects reconstructifs sont obtenus par ACP. Comme la plupart des informations correspondant au fond sont contenus dans les p premiers vecteurs (avec $p < N$), seulement ceux-ci sont conservés. L'ACP calcule ainsi les premiers p vecteurs qui contiennent l'information la plus représentative, alors que l'information discriminative n'est pas stockée. Pour garder cette information discriminative, le sous-espace principal tronqué est augmenté avec $c-1$ vecteurs additionnels qui apportent l'information nécessaire pour appliquer l'algorithme LDA. Supposons qu'un sous-espace augmenté ACP (AACP) a déjà été construit pour les N premières images. Le modèle courant du fond consiste alors en les vecteurs $U^{(N)} \epsilon \mathbb{R}^{m \times (k+c-1)}$, le vecteur moyenne $\mu \epsilon \mathbb{R}^m$ et les vecteurs coefficients $A^{(N)} \epsilon \mathbb{R}^{(k+c-1) \times n}$. A l'étape $N+1$, un nouvel sous-espace AACP est calculé à partir de la représentation (vecteurs coefficients) des N premières images et de l'image courante. La mise à jour du sous-espace commence par la

projection de l'image $x^{(N+1)}$ dans le sous-espace courant :

$$a = U^{(N)T}(x^{(N+1)} - \mu^{(N)}) \tag{4.28}$$

L'étape suivante consiste à calculer la reconstruction y de l'image courante et du vecteur résiduel r :

$$y = U^{(N)}a + \mu^{(N)} \tag{4.29}$$

$$r = x^{(N+1)} - y \tag{4.30}$$

La nouvelle base prend en compte le vecteur résiduel en l'utilisant dans le sous-espace. Ainsi, les coefficients de la nouvelle base deviennent :

$$A' = \begin{bmatrix} A^{(n)} & a \\ 0 & ||r|| \end{bmatrix}$$

Ensuite, l'ACP est appliquée à la matrice A'. Les nouveaux vecteurs propres U'' et la valeur moyenne μ' sont ainsi obtenus. L'étape suivante consiste à projeter les vecteurs coefficients A' sur la nouvelle base U'' [580] et à appliquer une rotation de la base courante pour correspondre à la nouvelle base [580] :

$$A = U''^{T}(A' - \mu' 1_{1\times(n+1)}) \tag{4.31}$$

$$U = U'U'' \tag{4.32}$$

Une fois que l'image courante du fond a été mise à jour, l'algorithme LDA est appliqué aux vecteurs coefficients $A \epsilon \mathbb{R}^{(k+c)\times(n+1)}$. La représentation discriminative est ainsi contenue dans les vecteurs LDA $V \epsilon \mathbb{R}^{(k+c)\times(c-1)}$. A ce point, les informations reconstructives et discriminatives sont conservées dans le modèle. Cependant, la taille du sous-espace AACP a été augmenté de un. Pour conserver la dimension du modèle, la dimension des matrices U,A et V sont réduites de un.

4.4 Classification

La détection des objets mobiles est obtenue à partir de la comparaison entre l'image de fond reconstruite et l'image courante :

$$|y - x^{(n+1)}| > T_{threshold} \tag{4.33}$$

4.5 Mise à jour du fond

La classification est effectuée en deux étapes. Premièrement, la nouvelle image est projetée dans la base de l'ACP incrémentale et ensuite le vecteur coefficient obtenu est projeté sur les vecteurs LDA. Ainsi, la fonction classification est mise à jour comme suit :

$$g(x) = V^T U^T (x - \mu^{(n+1)}) \tag{4.34}$$

Pour réduire la dimension du sous-espace de un, les matrices U, A et V sont divisées en sous-matrices contenant les p premières dimensions qui doivent être conservées et les c dimensions qui doivent être réduites de un :

$$U = \left[\begin{array}{cc} U_k & U_c \end{array} \right]$$

$$A = \left[\begin{array}{c} A_k \\ A_c \end{array} \right]$$

$$V = \left[\begin{array}{c} V_k \\ V_c \end{array} \right]$$

La moyenne doit être aussi recalculée suivant la nouvelle base :

$$\mu^{(n+1)} = \mu^{(n)} + U' \mu' \tag{4.35}$$

L'étape suivante consiste à orthonormaliser V_c et mettre à jour la base AACP, les coefficients et les vecteurs LDA comme suit :

$$U^{(n+1)} = \left[\begin{array}{cc} U_k & U_c V_c \end{array} \right]$$

$$A^{(n+1)} = \left[\begin{array}{c} A_k \\ V_c^T A_c \end{array} \right]$$

$$V^{(n+1)} = \left[\begin{array}{c} V_k \\ (V_c^T V_c)^{1/2} \end{array} \right]$$

5 Expérimentations et évaluations

Nous avons comparé l'approche discriminative IMMC et l'approche mixte IPCA-LDA que nous proposons dans ce chapitre avec les algorithmes conventionnels comme les modèles Gaussiens (SG [562], MOG [490] et KDE [167]) et les approches reconstructives (SL-PCA [387], SL-INMF [83] and SL-IRT [315]). Cette évaluation a été menée à l'époque sur la base de données Wallflower [521].

Le tableau 4.3 montre les résultats obtenus par les différents algorithmes sur chaque séquence. Les résultats des algorithmes SG, MOG et SL-PCA proviennent de l'article [521], et ceux de l'algorithme SL-INMF ont été fournis directement par Bucak et al. [83]. Pour les algorithmes KDE, SL-IRT et SL-IMMC, les résultats proviennent de nos expérimentations. Pour notre algorithme SL-IMMC, nous présentons les résultats obtenus pour $N = 30$ et $N = 100$ en les notant respectivement SL-IMMC (1) et SL-IMMC(2). Les scores correspondants en faux positifs (FP) et faux négatifs (FN) sont regroupés dans le tableau 4.2. En premier, nous pouvons comparer les performances des modèles Gaussiens avec celles des modèles de réduction de dimension. Ensuite, nous comparons les approches proposées avec les approches reconstructives.

5.1 Evaluation en termes de détection

1- Modèles Gaussiens vs modèles de réduction de dimension

A partir du tableau 4.2, on peut constater que les modèles Gaussiens ont un nombre TE plus élevé (>20 000) que ceux des modèles de réduction de dimension (TE<20 000). Cela reste essentiellement dû à leur faible robustesse dans le cas de variations lumineuses globales sur la séquence "Light Switch" (LS). Par contre, le modèle des MOG ainsi que celui des KDE semblent être mieux adaptés aux fonds dynamiques comme le montrent les résultats sur la séquence "Waving Trees" (WT). Cela confirme le constat que les modèles Gaussiens sont plus adaptés aux fonds dynamiques alors que les méthodes de réduction de dimension sont plus robustes aux variations lumineuses.

2- Modèles discriminatifs/mixtes vs modèles reconstructifs

Les méthodes proposées IMMC et IPCA-LDA présentent les termes TE les plus faibles suivi dans l'ordre par les approches reconstructives suivantes : SL-IRT, SL-PCA et SL-INMF. Ces performances sont à interpréter avec précaution dans la mesure où une performance faible sur une seule séquence affecte le terme TE fortement. La seule interprétation fiable de cette évaluation est que certains modèles sont plus ou moins adaptés à telle ou telle situation. Par exemple, seul l'algorithme SL-PCA donne des faux positifs (FP) pour la séquence "Moved Object" (MO) parce que ce modèle n'est pas mis à jour au cours du temps. L'algorithme

		Challenges							
Algorithme	Type Erreur	MO	TD	LS	WT	C	B	FA	Erreurs Totales (TE)
SG	false neg	0	949	1857	3110	4101	2215	3464	
Wren et al.[562]	false pos	0	535	15123	357	2040	92	1290	35133
MOG	false neg	0	1008	1633	1323	398	1874	2442	
Stauffer et al.[490]	false pos	0	20	14169	341	3098	217	530	27053
KDE	false neg	0	1298	760	170	238	1755	2413	
Elgammal et al.[167]	false pos	0	125	14153	589	3392	933	624	26450
PCA	false neg	0	879	962	1027	350	304	2441	
Oliver et al.[387]	false pos	1065	16	362	2057	1548	6129	537	17677
ICA	false neg	0	1199	1557	3372	3054	2560	2721	
Tsai et Lai[523]	false pos	0	0	210	148	43	16	428	15308
INMF	false neg	0	724	1593	3317	6626	1401	3412	
Bucak et al.[83]	false pos	0	481	303	652	234	190	165	19098
IRT	false neg	0	1282	2822	4525	1491	1734	2438	
Li et al.[315]	false pos	0	159	389	7	114	2080	12	17053
IMMC (30)	false neg	0	1336	2707	4307	1169	2677	2640	
Farcas et al. [177]	false pos	0	11	16	6	136	506	203	15714
IMMC (100)	false neg	0	626	711	4106	1167	2175	2320	
Farcas et al. [177]	false pos	0	10	15	5	135	503	201	11974
IPCA-LDA	false neg	0	443	1107	1426	922	2380	2025	
Marghes et al. [359]	false pos	0	103	49	958	204	142	576	10335

TABLE 4.2 – Evaluation sur Wallflower dataset [521]. En gras, les approches développées par Diana Farcas et Cristina Marghes sous ma direction scientifique.

SL-INMF donnent le terme erreur le plus grand à cause de ses résultats sur la séquence "Camouflage" (C). SL-IMMC donnent les meilleurs résultats quand les objets sont proches de la caméra comme on peut le voir sur la séquence "Time of Day" (TD), "Light Switch" (LS) et "Camouflage" (C). Quand les objets sont loin de la caméra comme dans la séquence "Bootstrap" (B), les résultats restent acceptables. Cela confirme le fait que les approches discriminatives et mixtes sont adaptées quand la taille des objets est importante.

En résumé, les approches discriminatives et mixtes donnent de meilleurs résultats que les approches reconstructives. Cependant, l'approche discriminative nécessite la vérité-terrain sur la séquence d'apprentissage. Une façon de pallier cela consiste à obtenir une vérité-terrain approximée à partir des résultats obtenus par une méthode de base telle que la moyenne ou la médiane.

TABLE 4.3 – Résultats obtenus avec le Wallflower dataset [521]. De gauche à droite : Séquence "Moved Object"' (985), Séquence "Time of Day" (1850), Séquence "Light Switch" (1865), Séquence "Waving Trees" (247), Séquence "Camouflage" (251), Séquence "Bootstrap" (2832), Séquence "Foreground Aperture" (449). En gras, les approches développées par Diana Farcas et Cristina Marghes sous ma direction scientifique.

5.2 Evaluation en termes de complexité et de temps de calcul

Ici, nous comparons les performances des différentes méthodes en termes de nombre de paramètres, de complexité et de mémoire requise comme suit :

— **Paramètres :** Le paramètre important pour les méthodes statistiques est le nombre N d'images d'apprentissage pour les modèles. Pour chacune des méthodes, il y a un compromis entre à faire entre d'une part la précision obtenue avec un nombre élevé N, et d'autre part l'espace mémoire requis et la vitesse de convergence de l'algorithme. D'autres paramètres comme la moyenne et la variance pour le modèle SG et le modèle des MOG sont nécessaires pour l'apprentissage. Le seuil est déterminé quant à lui de façon empirique pour le modèle des MOG et des KDE. Le paramètre le plus important pour les MOG est K qui détermine le nombre de modes du modèle. Pour les modèles de réduction de dimension reconstructifs comme pour l'approche discriminative proposée, deux paramètres sont importants : le nombre N d'images d'apprentissage et le seuil T dans l'étape de classification des pixels en "fond" ou "objets mobiles".

— **Complexité :** L'algorithme SG est le plus rapide parce que la classification utilise juste un seuil et que la mise à jour adapte uniquement la moyenne et la variance. Sa complexité dépende de N uniquement lors de l'initialisation. Le modèle des MOG a une complexité de $O(NK)$ où K est le nombre de Gaussiennes. Pour la mise à jour, le modèle KDE calcule ces valeurs dans les noyaux Gaussiens sur les n images précédentes et a donc une complexité en $O(n)$ avec n égal typiquement à 100. Quant aux modèles de réduction de dimension reconstructifs, leur complexité dépend des opérations nécessaires pour stocker et mettre à jour comme la matrice des observations, et les matrice contenant les vecteurs propres et les valeurs propres. Par exemple, le modèle IRT nécessite $O(I_1 I_2(R_1 + R_2 + R_3))$ opérations [315]. Pour la classification, les modèles de réduction de dimension reconstructifs ont une complexité estimée par pixel de $O(P)$, où P est le nombre de vecteurs propres employés pour la reconstruction de l'image de fond. Pour la mise à jour, la complexité dépend de M qui est le nombre d'images employées pour mettre à jour le modèle. Finalement, la complexité de l'approche discriminative proposée IMMC est en $O(Nc)$ où c est le nombre de classes. Dans notre cas, $c = 2$ car il y a seulement deux classes "fond et "objet mobile".

— **Mémoire requise :** Pour les modèles statistiques, l'espace mémoire utilisé est directement lié à la complexité [77]. Pour l'étape de classification, les modèles de réduction de dimension reconstructifs nécessitent un espace mémoire par pixel en $O(P)$ où P est le nombre de vecteurs propres employés pour la reconstruction de l'image de fond. En outre, pour l'étape d'initialisation, ces méthodes nécessitent N images. L'approche discriminative proposée IMMC nécessite quant à elle en plus les images vérité terrain

Algorithme	Initialisation du fond	Mise à jour du fond	Classification	Images par seconde (fps)
SG [562]	O(N)	O(1)	O(1)	71
MOG [490]	O(NK)	O(K)	O(K)	33
KDE [167]	O(N)	O(n)	O(1)	11
PCA [387]	O(N)	O(N+M)	O(P)	3
INMF [83]	O(N)	O(M)	O(P)	-
IRT [315]	O(N)	O(M)	O(P)	5
IMMC	O(2N)	O(M)	O(P)	7

TABLE 4.4 – Complexité et images traitées par seconde (fps)

pour pouvoir estimer les matrices inter-classes et intra-classes. Si les images sont en niveau de gris, $2N$ valeurs doivent être stockées par pixel.

Le tableau 4.4 montre la complexité par pixel des différents algorithmes à chaque étape. Le lecteur trouvera plus de détails sur cette complexité dans les publications correspondantes. Bien que la méthode discriminative proposée ait une complexité plus importante et demande un espace mémoire plus important pour l'étape de l'initialisation, ses performances restent correctes lors du traitement.

Pour le temps de calcul par image, les expérimentations ont été effectuées sur un PC avec 4GB RAM, Intel Pentium D et une carte 128MB ATI Radeon X1300 GPU. Pour les modèles SG, MOG, KDE et PCA, nous avons utilisés les implémentations de Parks and Fels [395]. Pour les modèles IRT et IMMC, nous les avons implémentés en C++ avec OpenCV. Finalement, les résultats obtenus par le modèle INMF ont été fournis par les auteurs [83]. Le tableau 4.4 montre la performance des algorithmes en images traitées par secondes (fps) pour des images de taille 240×360. Les résultats obtenus par la méthode discriminative proposée IMMC sont similaires à ceux obtenus par le modèle IRT [315]. Bien que la méthode discriminative proposée soit plus beaucoup plus lente que les algorithmes des modèles Gaussiens, sa vitesse peut être améliorée par une implémentation GPU par exemple.

6 Conclusion

Dans ce chapitre, nous nous sommes focalisés sur l'utilisation et l'évaluation des méthodes de réduction de dimension dans la détection d'objets mobiles dans des séquences vidéo. Dans ce cadre, nous avons effectué les trois travaux de recherches suivants :

— **Un état de l'art sur les approches de réduction de dimension (Section 2)** qui a montré que les avantages et inconvénients des méthodes reconstructives utilisées jus-

qu'alors. Leur avantage principal est leur robustesse dans le cas de fonds comportant des variations lumineuses. Leur inconvénient majeur est du à l'hypothèse émise sur la taille des objets.

— **Un modèle de fond par une approche discriminative de réduction de dimension (Section 3)** adapté à la classification nécessaire lors de la détection. Cette approche discriminative appelée IMMC est incrémentale et permet effectivement de s'affranchir de la taille des objets mobiles par son aspect discriminatif. L'inconvénient principal est la nécessité d'avoir la vérité-terrain sur la séquence d'apprentissage. Une solution peut consister à obtenir une vérité-terrain approximée à partir des résultats obtenus par une méthode de base telle que la moyenne ou la médiane.

— **Un modèle de fond par une approche mixte de réduction de dimension (Section 4)** combinant une approche reconstructive (ACP) avec une approche discriminative (LDA). Dans un premier temps, le fond est modélisé de façon robuste avec l'avantage d'une méthode non-supervisée et, dans un second temps, les pixels sont classés en "fond" ou "objet mobile" avec une méthode adaptée à la classification.

Pour toutes ces méthodes de réduction de dimension, la détection des objets en mouvement nécessite un seuillage de l'image de différence entre l'image courante et celle du fond obtenue par apprentissage. Le seuil requis est la plupart du temps déterminé de façon expérimentale en fonction des conditions d'acquisition et de l'environnement filmé. Il est de plus identique et fixe pour tous les pixels et pour toute la séquence. Il serait donc intéressant d'investiguer un seuillage adaptatif. Pour l'instant, seuls Xu et al. [574] ont proposé un tel seuillage adaptatif qui combine deux seuils : 1) un seuil local basé sur la variance au niveau pixel pour la partie apprentisage et un 2) seuil global déterminé à partir de l'histogramme cumulatif de la différence pour la partie fonctionnement.

De plus, toutes ces méthodes permettent de prendre la décision de classification fond/objets mobiles pixel par pixel sans aucune régularisation spatiale ou temporelle. Il serait donc intéressant d'appliquer une stratégie de régularisation comme les champs de Markov.

Finalement, cette recherche a montré qu'une approche discriminative de réduction de dimension est plus performante pour l'étape de classification des pixels que les approches reconstructives. Cependant, une approche mixte telle que développée dans ce chapitre parait plus adéquate car elle permet de combiner les avantages des approches reconstructives et discriminatives. Pour rendre plus applicable ces deux approches, il reste à développer des versions temps-réel de ces méthodes. De plus, d'autres approches discriminatives peuvent être testées comme l'analyse discriminante linéaire (Linear Discriminant Analysis (LDA)) [508] et l'analyse canonique des corrélations (Canonical Correlation Analysis (CCA) [238]). Il existe d'ailleurs plusieurs versions incrémentales de LDA [109][285]. Par ailleurs, les moindres carrés partiels (Partial Least Squares (PLS)] [438] offrent aussi un bon potentiel

pour modéliser et mettre à jour le fond.

Chapitre 5

Approche par ACP robuste basée sur la décomposition en une matrice de rang faible et une matrice parcimonieuse

1 Introduction

Dans ce chapitre [1], nous nous focalisons sur l'apport des nouvelles avancées par ACP robuste basée sur la décomposition en matrices de rang faible et parcimonieuse dans la détection d'objets mobiles dans des séquences vidéo. Dans ces approches, ma matrice d'observation A est ainsi décomposée tel que $A = L + S$ où L est la matrice de rang faible et S est la matrice parcimonieuse contenant une faible partie de ses coefficients qui sont non nuls. Le fond est ainsi modélisé par la matrice L alors que les objets mobiles par la matrice parcimonieuse. **Les travaux de cette partie ont été menés sous ma direction scientifique dans le cadre de la thèse de Charles Guyon.**

Nous développons ainsi comment les problèmes apparaissant lors de la détection d'objets mobiles peuvent être résolus en utilisant ces nouvelles avancées. L'objectif est d'obtenir des méthodes plus robustes en présence de variations lumineuses et de fonds dynamiques au niveau de la modélisation du fond. Les contributions de cette partie sont donc les suivantes :

— **Un état de l'art pour une analyse systématique (Section 2)** : Dans cette section, notre contribution consiste en une revue des différentes approches par décomposition en matrices de rang faible et parcimonieuse qui sont classées en différentes catégories [78] : ACP robustes (RPCA), NMF robustes (RNMF), apprentissage robuste de sous-espaces (Robust Subspace Learning), suivi de sous-espaces (Subspace Tracking), et minimisation en rang faible (Low Rank Minimization). De plus, ces méthodes ont été évaluées sur une base de données récente du domaine : Background Models Chal-

1. Ce chapitre est basé sur nos publications [79], [78], [226] et [228].

lenge (BMC) [530].
Publications relatives : 1 Chapitre de livre INTECH [227], 1 Journal CVIU [79], 1 Journal en rédaction COSREV [78], 3 conférences (ICIP [221], ICIAR [222], ISVC [225]).

— **ACP robuste par "Iteratively Reweighted Least Squares" (IRLS) (Section 3)** : Dans cette seconde section, nous développons un nouvel algorithme pour la détection d'objets mobiles basé sur la factorisation robuste de matrice. La matrice A est considérée comme approximativement de rang faible avec une faible partie corrompue par la présence des objets mobiles. L'objectif est de pallier une limitation de la méthode originale de décomposition par Principal Component Pursuit (PCP)[92] en prenant en compte l'aspect spatial pour obtenir une estimation robuste de la matrice L de rang faible et de la matrice parcimonieuse S. Notre contribution se résume comme suit : 1) L'addition de la contrainte spatiale dans le processus de minimisation et 2) l'optimisation à partir de l'algorithme IRLS pour pondérer les paramètres de la norme $||.||_{\alpha,\beta}$ dans la décomposition.
Publications relatives : 3 conférences (ISVC [226], ICPR [224], ACCV workshop on Background Model Challenges (BMC) [223]).

— **ACP robuste polynomiale (Section 4)** : Les problèmes des variations lumineuses soudaines, des variations des paramètres de la caméra (auto-ajustements) et du bruit lors de l'acquisition engendrent des effets non linéaires qui apparaissent localement dans le voisinage des intensités des pixels. Pour cela, nous avons proposé de modéliser le fond à l'aide d'une ACP polynomiale. Cette méthode permet de plus de prendre en compte l'aspect additif de certaines variations et l'aspect multiplicatif d'autres variations. De plus, lors de la minimisation, nous utilisons comme précédemment la norme $||.||_{\alpha,\beta}$ et l'algorithme IRLS.
Publication relative : 1 journal en rédaction [228].

Pour la partie état de l'art correspondant au chapitre [227], un site web a été mis en ligne : IRLS[1]. Ce site permet de visualiser les résultats obtenus ainsi que de télécharger les différents codes.

2 Approche par décomposition en une matrice de rang faible et une matrice parcimonieuse : Une vue d'ensemble

Depuis les travaux de Candès et al. [92] en 2009 sur les ACP robustes basée sur la décomposition en matrices de rang faible et parcimonieuse, un grand nombre de publications

1. http://sites.google.com/site/rpcaforegrounddetection/

(plus de 160 articles recensés dans notre publication [78]) se sont focalisées sur ce type de décomposition en montrant des applications sur la détection d'objets mobiles. Ces différentes approches peuvent se classer en cinq grands groupes comme suit :

1. Approches ACP robustes (RPCA).
2. Approches NMF robustes (RNMF).
3. Apprentissages robustes de sous-espaces (Robust Subspace Learning).
4. Suivis de sous-espaces (Subspace Tracking).
5. Minimisations en rang faible (Low Rank Minimization).

Nous examinons ces cinq groupes dans les sous-sections suivantes dans l'objectif de donner une vue globale, synthétique et la plus exhaustive possible pour évaluer le potentiel de ces méthodes et en dégager des pistes pour l'application à la modélisation de fond et la détection d'objets mobiles. Le tableau 5.1 donne une vue d'ensemble de ces méthodes.

2.1 Approches ACP robustes

Les récentes avancées en ACP robustes différent par la décomposition et le problème d'optimisation associé comme le montre les tableaux 5.7 et 5.8. Différents solveurs existent également comme le montre de façon synthétique les tableaux 5.12, 5.13, 5.14 et 5.15. Ces différentes approches sont présentées de façon détaillée dans nos articles [79] [78] et peuvent être classées comme suit.

1- RPCA via Principal Component Pursuit (RPCA-PCP)

La méthode RPCA-PCP proposée par Candès et al. [92] en 2009 est à l'heure actuelle la méthode la plus investiguée [79]. Dans ces premiers travaux concernant la méthode RPCA-PCP, Candès et al. [92] ont proposé d'employer une optimisation convexe pour effectuer une ACP robuste à des outliers. La matrice d'observation A est ainsi décomposée comme suit :

$$A = L + S \tag{5.1}$$

où L est une matrice de rang faible et S est une matrice parcimonieuse contenant une faible partie de ces coefficients qui sont non nuls. La formulation adéquate est basée sur l'emploi de la norme l_0 pour minimiser la fonction d'énergie suivante :

$$\min_{L,S} \ rank(L) + \lambda ||S||_0 \quad \text{avec la contrainte} \quad A - L - S = 0 \tag{5.2}$$

où $\lambda > 0$ est un paramètre arbitraire permettant de pondérer l'aspect rang faible et l'aspect parcimonieux. Mais comme ce problème est NP-complexe, une solution courante est de trouver L avec l'optimisation suivante :

$$\min_{L,S} \ ||L||_* + \lambda ||S||_1 \quad \text{avec la contrainte} \quad A - L - S = 0 \tag{5.3}$$

Décomposition en matrices de rang faible et parcimonieuse	Catégories	Sous-catégories	Auteurs - Dates
Robust Principal Components Analysis (RPCA)	Principal Component Pursuit	PCP	Candès et al. (2009) [92]
(**Survey Guyon et al. [227]**)	(**Survey Bouwmans and Zahzah[79]**)	Stable PCP	Zhou et al. (2010) [647]
		Quantized PCP	Becker et al. (2011) [55]
		Block based PCP	Tang and Nehorai (2011) [509]
		Local PCP	Wohlberg et al. (2012) [560]
	Outlier Pursuit	OP	Xu et al. (2010) [570]
		SOP	Xu et al. (2010) [570]
	Sparsity Control	SpaCtrl	Mateos et Giannakis (2010) [361]
		SpaCorr	Hsu et al. (2011) [252]
	Log-sum heuristic Recovery	LHR	Deng et al. (2012) [150]
	Iteratively Reweighted Least Square	IRLS	Guyon et al. (2012) [226]
		Spatial IRLS	Guyon et al. (2012) [224]
		Spatio-temporal IRLS	Guyon et al. (2012) [223]
	Stochastic Optimization	RPCA-SO	Goes et al. (2014) [198]
		OR-RPCA	Feng et al. (2013) [181]
	Bayesian RPCA	Bayesian RPCA	Ding et al. (2011) [156]
		Variational Bayesian RPCA	Babacan et al. (2012) [36]
		Factorized Variational Bayesian RPCA	Aicher (2013) [10]
	Approximated RPCA	GoDec	Zhou and Tao (2011) [643]
		Semi-Soft GoDec	Zhou and Tao (2011) [643]
	Sparse Additive Matrix Factorization	SAMF	Babacan et al. (2012) [35]
	Variational Bayesian Sparse Estimator	VBSE	Chen et al. (2014) [112]
Robust Non-negative Matrix Factorization (RNMF)	Manhattan Non-negative Matrix Factorization	MahNMF	Guan et al. (2012) [212]
	Near-separable Non-negative Matrix Factorization	NS-NMF	Kumar and Sindhwani (2013) [295]
Robust Subspace Learning	Robust Subspace Discovery	RSD	Wang et al. (2013) [548]
	Robust Subspace Recovery via Bi-Sparsity	RoSuRe	Bian and Krim (2014) [64]
	Robust Orthonomal Subspace Learning	ROSL	Xu et al. (2014) [475]
		ROSL+ (Random Sampling)	Xu et al. (2014) [475]
Subspace Tracking (ST)	Grassmannian Robust Adaptive Subspace Tracking Algorithm	GRASTA	He et al. (2011) [240]
		t-GRASTA	He et al. (2013) [242]
	l_p-norm Robust Online Subspace Tracking	pROST	Hage and Kleinsteuber [230]
		Real Time pROST	Hage and Kleinsteuber [451]
	Grassmannian Online Subspace Updates with Structured-sparsity	GOSUS	Xu et al. (2013) [571]
	Fast Adaptive Robust Subspace Tracking	FARST	Ahn (2014) [9]
Low Rank Minimization (LRM)	Contiguous Outlier Detection	DECOLOR	Zhou et al. (2011) [646]
	Direct Robust Matrix Factorization	DRMF	Xiong et al. (2011) [569]
		DRMF-R	Xiong et al. (2011) [569]
	Probabilistic Robust Matrix Factorization	PRMF	Wang et al. (2012) [542]
	Bayesian Robust Matrix Factorization	BRMF	Wang et al. (2013) [543]
		Markov BRMF (MBRMF)	Wang et al. (2013) [543]
	Practical Low-Rank Matrix Factorization	PLRMF	Zheng et al. (2012) [636]
	Low Rank Matrix Factorization with MoG noise	LRMF-MOG	Meng et al. (2013) [365]
	Unifying Nuclear Norm and Bilinear Factorization	UNN-BF	Cabral et al. (2013) [85]
	Robust Rank Factorization	RRF	Sheng et al. (2014) [465]

TABLE 5.1 – Vue d'ensemble des méthodes de décomposition en matrices de rang faible et parcimonieuse. La première colonne indique les différentes méthodes et la seconde colonne montre les différentes catégories de chaque cadre mathématique. Le troisième colonne indique les différentes méthodes de chaque catégorie ainsi que son acronyme entre parenthèses. La quatrième colonne donne le nom des auteurs et la date de la première publication de la méthode concernée.

FIGURE 5.1 – De gauche à droite : Image originale(309), matrice de rang faible L (Fond), matrice de parcimonieuse S (objets mobiles), masque de détection.

où $||.||_*$ et $||.||_1$ sont respectivement la norme nucléaire (qui est la norme l_1 des valeurs propres de la matrice A) et la norme l_1. $\lambda > 0$ a le même rôle que dans l'équation 5.2. Couramment, $\lambda = \frac{1}{\sqrt{\max(m,n)}}$. Sous certaines hypothèses, cette approche appelée Principal Component Pursuit (PCP) retrouve parfaitement la matrice de rang faible et la matrice parcimonieuse. Le fond est ainsi modélisé par un sous-espace de rang faible qui peut changer progressivement au cours du temps, alors que les objets mobiles sont quant à eux les données atypiques bien éloignées de la masse des données correspondant au fond et de plus distribuées de façon parcimonieuse. Ainsi, Candès et al. [92] ont montré des résultats visuels intéressants sur la détection d'objets mobiles.

RPCA-PCP et modélisation de fond :

La minimisation en rang faible de la matrice L offre un cadre adéquate à la modélisation de fond à cause de la corrélation entre les images consécutives. Ainsi, la décomposition de la matice A en les matrices L et S supposent que le fond est approximé par la composante de rang faible qui peut changer graduellement au cours du temps, alors que les objets mobiles constituent les outliers contenus dans la matrice S.

Pour obtenir le masque de détection, la matrice S est seuillée par un seuil déterminé de façon expérimentale. $rank(L)$ influence le nombre de mode qui peuvent être contenu dans la matrice L : Si $rank(L)$ est trop élevé, le modèle va contenir des objets mobiles dans sa représentation ; si $rank(L)$ est trop faible, la modèle pourra représenter moins de modes et la multi-modalité qui apparait en présence de fonds dynamiques ne sera pas prise en compte dans la représentation. La qualité de la séparation entre le fond et les objets mobiles est donc directement liée aux hypothèses de rang faible et de parcimonie émises respectivement pour le fond et les objets mobiles. Si ces hypothèses sont respectées, la meilleure décomposition est obtenue quand l'algorithme d'optimisation converge. La figure 5.1 montre un exemple de résultats obtenus.

Cependant, l'algorithme PCP original présente plusieurs limitations : La première limitation est que ce type d'algorithmes est souvent très coûteux en temps de calcul par les solveurs mis en oeuvre ainsi que par le nombre de SVD nécessaire à chaque itération. La deuxième limitation réside dans l'aspect "batch" de l'algorithme qui nécessite d'avoir toutes les données pour faire la décomposition alors que les données apparaissent de façon continue au cours du temps dans l'application de la suppression de fond. La troisième limitation est que les données sont stockées sous forme de vecteurs et exploitées comme tels excluant ainsi l'information spatiale et temporelle. La quatrième limitation concerne l'hypothèse émise sur l'aspect rang faible et parcimonieux des matrices car celle-ci n'est pas strictement respectée dans le cas de la détection d'objets mobiles lors de présence de bruit par exemple. La cinquième limitation quant-à-elle concerne l'hypothèse que les matrices peuvent être complètement retrouvées à partir des observations et que la distribution de la perturbation est parcimonieuse et sans bruit.

Ces hypothèses sont rarement vérifiées dans le cas de la détection d'objets mobiles parce que 1) seulement une fraction des données sont observables dans certains environnements (Bootstrap), 2) les observations sont perturbées par du bruit, 3) les outliers qui sont les objets mobiles sont localisés spatialement.

Dans ce cadre, plusieurs améliorations ont donc été apportées à la version originale du PCP. Les tableaux 5.2 et 5.3 montrent une vue d'ensemble de ces améliorations concernant la décomposition ainsi que de leur solveur [325][326][86][604][584][464][564][373][331][327][348][199].

De plus, quelques auteurs ont proposé des versions incrémentales en mettant à jour les matrices de rang faible et parcimonieuse à l'apparition de chaque nouvelle image [415][417][416][418]. Des implémentations temps-réel ont aussi été conçues [20][406][218].

D'autres auteurs [425][194] ont proposé de déterminer la valeur de λ optimale. Ramirez and Shapiro [425] détermine cette valeur de façon expérimentale en faisant varier λ dans un intervalle et en conservant celle qui donne la meilleure décomposition au sens d'un critère appelé MDL (*Minimum Description Length*). Une autre solution proposée par Gao et al. [194] consiste à adapter λ en fonction du flot optique [194]. Dans les deux cas, la valeur employée est différente de celle prise par Candès et al. [92], c'est-à-dire $\lambda = \frac{1}{\sqrt{\max(m,n)}}$.

Enfin, d'autres auteurs [332][333] ont proposé des algorithmes rapides de calcul de SVD.

Toutes ces différentes approches sont présentées de façon détaillée dans notre article [78].

Principal Component Pursuit	Méthodes	Auteurs - Dates
Décomposition	1) original PCP	Candès et al. (2009) [92]
	2) modified-PCP (Fixed Rank)	Leow et al. (2013) [308]
	3) modified-PCP (Nuclear Norm Free)	Yuan et al. (2013) [603]
	4) modified-PCP (Capped Norms)	Sun et al. (2013) [498]
	5) modified-PCP (Inductive)	Bao et al. (2012) [50]
	6) modified-PCP (Partial Subspace Knowledge)	Zhan and Vaswani (2014) [617]
	7) p,q-PCP (Schattern-p Norm, l_q Norm Regularization)	Wang et al. (2014) [538]
	8) modified p,q-PCP (Schattern-p Norm, l_q Seminorm)	Shao et al. (2014) [459]
	9) modified PCP (2D-PCA)	Sun et al. (2013) [499]
	10) modified PCP (Rank-N Soft Constraint)	Oh (2012) [386]
	11) Joint Video Frame Set Division RPCA (JVFSD-RPCA)	Wen (2014) [558]
Solveur	1) Basic solvers	
	Singular Values Decomposition (SVT)	Cai et al. (2008) [86]
	Iterative Thresholding (IT)	Wright et al. (2009) [564]
	Accelerated Proximal Gradient (APG)	Lin et al.(2009) [326]
	Dual Method (DM)	Lin et al.(2009) [326]
	Exacted Augmented Lagrangian Method (EALM)	Lin et al. (2009) [325]
	Inexact Augmented Lagrangian Method (IALM)	Lin et al. (2009) [325]
	Alternating Direction Method (ADM)	Yuan and Yang (2009) [604]
	Symmetric Alternating Direction Method (SADM)	Goldfarb et al. (2010) [199]
	Non Convex Splitting ADM (NCSADM)	Chartrand (2012) [104]
	Douglas-Rachford Splitting Method (DRSM)	Gandy and Yamada (2010) [190]
	Variant of Douglas-Rachford Splitting Method (VDRSM)	Zhang and Liu (2013) [619]
	Proximity Point Algorithm (PPA)	Zhu et al. (2014) [648]
	Proximal Iteratively Reweighted Algorithm (PIRA)	Wang et al. (2014) [538]
	Alternating Rectified Gradient Method (ARGM)	Kim et al. (2014) [278]
	Parallel Direction Method of Multipliers (PDMM)	Wang et al. (2014) [533]
	2) Linearized solvers	
	Linearized Alternating Direction Method (LADM)	Yang and Yuan (2011) [584]
	Linearized Alternating Direction Method with Adaptive Penalty (LADMAP)	Lin et al. (2011) [327]
	Linearized Symmetric Alternating Direction Method (LSADM)	Goldfarb et al. (2010) [199]
	Fast Linearized Symmetric Alternating Direction Method (Fast-LSADM)	Goldfarb et al. (2010) [199]
	Linearized Inexact Alternating Direction Based Contraction Methods (LIADCM)	Gu et al. (2013) [210]
	3) Fast solvers	
	Randomized Projection for ALM (RPALM)	Mu et al. (2011) [373]
	l_1 filtering (LF)	Liu et al. (2011) [331]
	Block Lanczos with Warm Start	Lin and Wei(2010) [328]
	Exact Fast Robust Principal Component Analysis (EFRPCA)	Abdel-Hakim and El-Saban (30002012) [550]
	Inexact Fast Robust Principal Component Analysis (IFRPCA)	Abdel-Hakim and El-Saban (30002012) [550]
	PRoximal Iterative SMoothing Algorithm (PRISMA)	Orabona et al. (2012) [388]
	Fast Alterning Minimization (FAM)	Rodriguez and Wohlberg (2013) [434]
	Fast Alternating Direction Method of Multipliers(FADMM)	Yang and Wang (2014) [587]
	4) Online solvers	
	Online Alternating Direction Method (OADM)	Wang and Banerjee (2013) [532]
	5) Non convex solvers	
	Difference of Convex (DC)	Sun et al. (2013) [498]
	Fast Alternating Difference of Convex (FADC)	Sun et al. (2013) [498]
	6) 2D solvers	
	Iterative method for Bi-directional Decomposition (IMBD)	Sun et al. (2013) [499]

TABLE 5.2 – Vue d'ensemble des décompositions et des solveurs pour PCP (Partie 1). La première colonne indique les différentes catégories. La seconde colonne indique les différentes méthodes de chaque catégorie ainsi que son acronyme entre parenthèses. La troisième colonne donne le nom des auteurs et la date de la première publication de la méthode concernée.

Principal Component Pursuit	Méthodes	Auteurs - Dates
Algorithme Incrémental	Recursive Robust PCP (RR-PCP)	Qiu and Vaswani (2010) [415]
	Recursive Project Compressive Sensing (ReProCS)	Qiu and Vaswani (2011 [416]
	Support-Predicted Modified-CS RR-PCP)	Qiu and Vaswani (2011) [417]
	Support-Predicted Modified-CS)	Qiu and Vaswani (2011) [416]
	Automated ReProCS	Qiu and Vaswani (2012) [418]
	Practical ReProCS (Prac-ReProCS)	Guo et al. (2013) [216]
	Incremental Low-Rank (iLR) Algorithm	Wei et al. (2014) [557]
	Incremental PCP (incPCP)	Rodriguez and Wohlberg (2014) [435]
Implémentation Temps Réel	CAQR	Anderson et al. (2010) [20]
	Real-Time PCP	Pope et al. (2011) [406]
	LR Submatrix Recovery/Reconstruction (LRSRR)	Guo et al. (2014) [218]
Solution Optimale	Minimum Description Length (MDL)	Ramirez and Shapiro (2012) [425]
	Saliency Measure	Gao et al. (2012) [194]
Algorithme SVD	Full SVD	-
	Partial SVD	-
	Linear Time SVD	Yang and An (2013) [586]
	Limited Memory SVD (LMSVD)	Liu et al. (2013) [332]
	Symmetric Low-Rank Product-Gauss-Newton (SLRPGN)	Liu et al. (2014) [333]

TABLE 5.3 – Vue d'ensemble des décompositions et des solveurs pour PCP (Partie 2). La première colonne indique les différentes catégories. La seconde colonne indique les différentes méthodes de chaque catégorie ainsi que son acronyme entre parenthèses. La troisième colonne donne le nom des auteurs et la date de la première publication de la méthode concernée.

De plus, certains travaux se sont attachés à résoudre certains problèmes spécifiques apparaissant dans la détection d'objets mobiles comme les suivants :

1. **Présence de bruit :** La présence de bruit est courante dans les vidéos issues d'une caméra d'entrée de gamme ou de la compression avec perte.
2. **Quantification en pixels :** La quantification en pixels induit au plus une erreur de 0.5 dans la valeur du pixel.
3. **Aspect spatial et temporel :** Les objets mobiles forment des régions où les pixels sont spatiallement connexes. De plus, ceux-ci présentent un mouvement continu engendrant une continuité spatiale entre images consécutives.
4. **Variation locale du fond :** Le fond peut présenter des mouvements dus à une vibration de la caméra par exemple.

Pour résoudre (1), Zhou et al. [647] ont proposé une version stable de PCP (SPCP) qui garantit une décomposition stable et précise en présence de bruit. Becker et al. [55] ont proposé quant à eux une version qui prend en compte l'erreur induite par la quantification dans la valeur des pixels (2). Pour le point (3), Tang et Nehorai [509] ont développé une approche PCP qui renforce l'aspect rang faible sur la matrice de rang faible et l'aspect spatial sur la

matrice parcimonieuse. Pour le point (4), Wohlberg et al. [560] ont proposé un modèle de décomposition basé sur l'union de sous-espace de dimension inférieure pour prendre en compte les variations locales. Ces quatres différentes déclinaisons de PCP peuvent être brièvement expliquées comme suit :

1. **PCP prenant en compte la présence de bruit :** Zhou et al. [647] ont proposé une version stable de PCP (SPCP) qui décompose la matrice d'observation A comme suit :

$$A = L + S + E \tag{5.4}$$

où E est un terme "bruit" (c'est-à-dire qu'il contient le bruit de la matrice A) et $||E||_F < \delta$ pour $\delta > 0$. Pour déterminer L and S, Zhou et al. [136] résolvent le problème d'optimisation convexe suivant :

$$\min_{L,S} \ ||L||_* + \lambda||S||_1 \quad \text{subj} \quad ||A - L - S||_F < \delta \tag{5.5}$$

où $||.||_F$ est la norme de Frobenius et $\lambda = \frac{1}{\sqrt{n}}$. Pour résoudre ce problème convexe, Zhou et al. [647] ont proposé une version modifiée du solveur employé pour le PCP. Plusieurs améliorations ont été apportées à cette version stable du PCP et le tableau 5.4 montre une vue d'ensemble de ces améliorations ainsi que de leur solveur. De plus, quelques auteurs ont proposé des versions incrémentales [489] ou des implémentations temps réel [349]. Ces différentes approches sont présentées de façon détaillée dans notre article [78].

2. **PCP prenant en compte la quantification :** Becker et al. [55] ont proposé une version avec une contrainte basée sur une inégalité pour prendre en compte l'erreur de quantification de la valeur de l'intensité des pixels. En effet, chaque pixel a une valeur comprise entre $0, 1, 2, \ldots, 255$. Cette valeur est une valeur tronquée de la valeur réelle qui est dans l'intervalle $[0, 255]$. Ainsi, l'idée est d'appliquer une ACP robuste sur les valeurs réelles des observations au lieu des valeurs tronquées par la quantification. Dans ce but, Becker et al. [55] ont supposé que la somme $L + S$ n'est pas strictement égale à A, mais plutôt que $L + S$ est proche de A à un facteur de quantification près. La quantification introduit une erreur maximale de 0.5 sur la valeur de l'intensité. La décomposition est donc formulée de la façon suivante :

$$A = L + S + Q \tag{5.6}$$

où Q est l'erreur due à la quantification. Le problème d'optimisation convexe est formulé comme suit :

$$\min_{L,S} \ ||L||_* + \lambda||S||_1 \quad \text{subj} \quad \|A - L - S\|_\infty \leq 0.5 \tag{5.7}$$

La norme ℓ_∞ permet de prendre en compte l'erreur due à la quantification. Pour résoudre ce problème d'optimisation, Becker et al. [55] ont proposé dans un cadre plus

Stable Principal Component Pursuit	Méthodes	Auteurs - Dates
Décomposition	1) Original SPCP	Zhou et al. (2010) [647]
	2) Modified-SPCP (Bilateral Projection)	Zhou and Tao (2013) [644]
	3) Modified-SPCP (Nuclear-Norm Free)	Yuan et al. (2013) [603]
	4) Modified-SPCP (Nuclear-Norm Free for blur in noisy video)	Yuan et al. (2013) [603]
	5) Modified-SPCP (Undercomplete Dictionary)	Sprechman et al. (2012) [489]
	6) Variational SPCP (Huber Penalty)	Aravkin et al. (2014) [25]
Solveur	Alternating Splitting Augmented Lagrangian method (ASALM)	Tao and Huan (2011) [512]
	Variational ASALM (VASALM)	Tao and Huan (2011) [512]
	Parallel ASALM (PSALM)	Tao and Huan (2011) [512]
	Non Smooth Augmented Lagrangian Algorithm (NSA)	Aybat et al. (2011) [30]
	Partially Smooth Proximal Gradient (PSPG)	Aybat et al. (2012) [31]
	Alternating Direction Method - Increasing Penalty (ADMIP)	Aybat et al. (2012) [32]
	Greedy Bilateral Smoothing (GreBsmo)	Zhou and Tao (2013) [644]
	Bilinear Generalized Approximate Message Passing (BiG-AMP)	Parker and Schniter (2012) [394]
Algorithme Incrémental	Fast Trainable Encoders	Sprechman et al. (2012) [489]
Implémentation Temps-réel	DFC-PROJ	Mackey et al. (2011) [349]
	DFC-PROJ-ENS	Mackey et al. (2011) [349]
	DFC-NYS	Mackey et al. (2011) [349]
	DFC-NYS-ENS	Mackey et al. (2011) [349]

TABLE 5.4 – Vue d'ensemble des décompositions et des solveurs pour SPCP. La première colonne indique les différentes catégories. La seconde colonne indique les différentes méthodes de chaque catégorie ainsi que son acronyme entre parenthèses. La troisième colonne donne le nom des auteurs et la date de la première publication de la méthode concernée.

général que les solveurs PCP une série de solveurs appelés *Templates for First-Order Conic Solvers (TFOCS)*.

3. **PCP prenant en compte la présence d'outliers sur des colonnes entières :** Tang et Nehorai [509] ont proposé une décomposition PCP basée "bloc" (BPCP) qui renforce l'aspect rang faible sur la matrice de rang faible et l'aspect spatial sur la matrice parcimonieuse. Cette décomposition est similaire à l'originale sauf que S est contrainte à un aspect spatial. e problème d'optimisation convexe est formulée comme suit [509] :

$$\min_{L,S} \; ||L||_* + \kappa(1-\lambda)||L||_{2,1} + \kappa\lambda||S||_{2,1}$$

$$\text{subj} \quad A - L - S = 0 \tag{5.8}$$

où $||.||_*$ est $||.||_{2,1}$ sont respectivement la norme nucléaire et la norme $l_{2,1}$. La norme $l_{2,1}$ corresponds à la norme l_1 des vecteurs obtenus à partir de la norme l_2 des valeurs contenues en colonne. Le terme $\kappa(1-\lambda)||L||_{2,1}$ assure que la matrice L contient aucune colonne avec des outliers.

4. **PCP prenant en compte l'aspect local :** La décomposition PCP est très efficace mais pas le modèle sous-jacent n'est pas adapté quand les données d'observations ne peuvent pas être modélisée par un sous-espace unique. Pour cela, Wohlberg et al. [560] ont proposé une décomposition plus générale basée sur un modèle sous-jacent consistant en une union de sous-espaces comme suit :

$$A = AU + S \tag{5.9}$$

Le problème d'optimisation convexe correspondant est formulé comme suit :

$$\min_{U,S} \; \alpha||U||_* + \beta||U||_{2,1} + \beta||S||_1 \quad \text{subj} \quad A - AU - S = 0 \tag{5.10}$$

La norme $l_{2,1}$ encourage les colonnes de U à contenir des zéros, sans décourager la présence de valeurs non nulles parmi les entrées des colonnes non nulles. La norme l_1 encourage la présence de valeurs nulles dans les colonnes non nulles de S. Wohlberg et al. [560] résolvent l'équation (5.10) en utilisant un algorithme appelé *Split Bregman Algorithm (SBA)*.

De plus, RPCA-PCP peut être aussi appliqué dans le domaine des mesures plutôt que dans le domaine des pixels pour être employé en *compressive sensing* [556][265]. Les expérimentations ont montré que les objets mobiles peuvent être retrouvés avec un nombre de mesures plus petit. Nous nous sommes limités dans ce chapitre ainsi que dans notre évaluation comparative [79] aux méthodes dans le domaine du pixel pour pouvoir les comparer avec les méthodes classiques.

2- RPCA via Outlier Pursuit (RPCA-OP)

Xu et al. [570] ont introduit une méthode ACP robuste via Outlier Pursuit pour obtenir une décomposition quand tous les coefficients d'une colonne ou de plusieurs colonnes sont corrompus, c'est-à-dire quant un objet prend pratiquement toute une colonne. De plus, Xu et al. [570] ont amélioré cette approche avec une version stable OP (SOP) qui garantit une décomposition stable et précise en présence de bruit.

3- RPCA via Sparsity Control (RPCA-SpaCtrl)

Mateos et Giannakis [361] ont proposé une méthode ACP robuste dans laquelle des paramètres ajustables permettent de contrôler la parcimonie de la matrice estimée et le nombre d'outliers.

4- RPCA via Sparse Corruptions (RPCA-SpaCorr)

Même si la matrice A est exactement la somme d'une matrice de rang faible L et d'une matrice parcimonieuse S, il se peut qu'on ne puisse identifier les caractéristiques de cette somme. En effet, la matrice parcimonieuse S peut être aussi de rang faible, ou bien la matrice de rang faible L peut être aussi parcimonieuse. Pour résoudre ce cas, Hsu et al. [252] ont imposé des conditions sur les caractéristiques de rang faible et de parcimonie pour garantir leur identification à partir de A.

5- RPCA via Log-sum heuristic Recovery (RPCA-LHR)

Quand la matrice présente un rang élevé ou une quantité de corruption dense, les approches convexes ne sont pas très performantes. Pour résoudre cela, Deng et al. [150] ont employé une approche basée sur une heurisitque log-sum qui permet d'identifier la composante de rang faible.

6- Bayesian RPCA (BRPCA)

Ding et al. [156] ont proposé une ACP robuste dans le cadre Bayésien basée une représentation statistique du bruit et permettant d'agir simultanément les composantes de rang faible et parcimonieuse. De plus, les champs de Markov sont employés pour prendre en compte l'aspect spatial et temporel entre les lignes ou les colonnes consécutives correspondant aux images. Cette méthode donne de bons résultats en présence de fond dynamiques. Babacan et al. [36] ont amélioré celle-ci dans un cadre variationnel.

D'autres améliorations ont été apportées à la version originale de BRPCA et le tableau 5.5 montre une vue d'ensemble de ces améliorations ainsi que de leur solveur. Ces différentes approches sont présentées de façon détaillée dans notre article [78].

Bayesian Robust Principal Component Analysis	Méthodes	Auteurs - Dates
Décompositions	1) Original BRPCA	Ding et al. (2011) [156]
	2) Variational BRPCA (VBRPCA)	Babacan et al. (2011) [36]
	3) Factorized Variational BRPCA (FVBRPCA)	Aicher (2013) [11]
	4) MOG-RPCA	Zhao et al. (2014) [633]
Solveurs	Markov chain Monte Carlo (MCMC)	Robert and Cassela (2004) [432]
	Variational Bayesian Inference (VB)	Beal (2003) [54]
	Approximate Bayesian Inference (AB)	Beal (2003) [54]

TABLE 5.5 – Vue d'ensemble des décompositions et des solveurs pour BPCP. La première colonne indique les différentes catégories. La seconde colonne indique les différentes méthodes de chaque catégorie ainsi que son acronyme entre parenthèses. La troisième colonne donne le nom des auteurs et la date de la première publication de la méthode concernée.

7- Approximated RPCA

Zhou et Tao [643] ont proposé d'approximer la décomposition dans leur méthode appelée Go Decomposition (GoDec). Cette méthode permet ainsi de trouver une décomposition approchée avec un rang prédéfini pour la matrice de rang faible et une cardinalité prédéfinie pour la matrice parcimonieuse quand une décomposition exacte n'est pas possible à cause de la présence d'un bruit additif. De plus, GoDec peut être accéléré de façon significative par une méthode dite de "bilateral random projection". Enfin, Zhou et Tao [643] ont proposé une version Semi-Soft GoDec qui utilise un seuil adaptatif pour le seuillage de la matrice S, alors que GoDec impose un seuil commun pour toutes les valeurs de S.

2.2 Approches NMF robustes

Les méthodes Non-negative Matrix Factorization (NMF) approxime une matrice X non-négative par un produit de deux matrices de rang faible W et H. NMF et ses extensions mimise la divergence de Kullback-Leibler ou la distance Euclidienne entre X et $W^T H$ en considérant que le bruit présente soit une distribution de Poisson soit une distribution Gaussienne. En pratique, ces conditions sur la distribution du bruit ne sont pas toujours vérifiées et cela particulièrement dans la cas de la séparation fond/objets mobiles.

Dans ce cadre, Guan et al. [212] ont donc proposé une méthode robuste appelée Manhattan NMF (MahNMF) qui minimise la distance de Manhattan entre X et $W^T H$. De façon similaire aux méthodes RPCA, MahNMF a pour objectif d'estimer de façon robuste la partie rang-faible et la partie parcimonieuse de la matrice X. Cette méthode donne des résultats qualitatifs similaires aux méthodes RPCA résolue via EALM [92] et GoDec [643].

En suivant la même démarche, Kumar et Sindhwani [295] ont présenté une approche similaire mais en utilisant la divergence de Bregman. Cette approche est appelée *Near-separable Non-negative Matrix Factorization* (NS-NMF). Les résultats présentés sous forme de courbes ROC montre une meilleure robustesse aux bruits que la méthode RPCA résolue via EALM [92].

2.3 Apprentissages robustes de sous-espaces

Dans cette catégorie se trouvent les méthodes robustes autres que les méthodes de type RPCA et RNMF. Ces méthodes sont les suivantes :

— **Robust Subspace Discovery (RSD) :** Wang et al. [548] ont introduit une connaissance *a priori* et ont utilisé un algorithme de type ALM pour minimiser la problème de la décomposition.

— **Robust Subspace Recovery via Bi-Sparsity :** Bian et Krim [64] ont proposé une modèle bi-parcimonieux pour mieux prendre en compte la présence des outliers. Ainsi, Bian et Krim [64] autorisent l'aspect parcimonieux aussi sur la matrice de rang faible et fournissent un algorithme appelé *Robust Subspace Recovery via bi-sparsity pursuit (RoSuRe)*. Les résults obtenus montrent de bonnes performances dans le cas de la vibration de caméra.

— **Robust Orthogonal Subspace Learning :** Shu et al. [475] ont proposé une méthode appelée *Robust Orthogonal Subspace Learning* (ROSL) pour pouvoir atteindre des perfomances proches du temps-réel. Pour cela, une mesure de rang sur la matrice de rang faible impose une parcimonie localisée pour ces coefficients dans le cas d'un sous-espace orthonormal. De plus, un algorithme de minimisation de cette mesure permet d'obtenir la matrice de rang faible avec une complexité quadratique en fonction de sa taille. Les résultats de la précision de l'algorithme RPCA résolue via IALM [92] sont obtenus avec des temps de l'ordre de 35 secondes pour une séquence de 1600 images de taille 160×128. Une version plus rapide appelée ROSL+ basée sur le principe du *sub-sampling* obtient des temps de l'ordre de 3 secondes. ROSL et ROSL+ sont respectivement 10 fois et 92 fois plus rapide que le RPCA-PCP résolu via IALM.

2.4 Suivis de sous-espaces

Le suivi de sous-espaces (Subspace Tracking) est un modèle adapté à la modélisation de fond. He et al. [240] ont développé une méthode de descente de gradient incrémental sur les variétés grassmanniennes dont les points correspondent des d sous-espaces vectoriels d'un espace vectoriel fixé d. Cet algorithme appelé GRASTA (Grassmannian Robust Adaptive

Subspace Tracking Algorithm) utilise une fonction robuste de coût basée sur la norme l_1 pour estimer et suivre les sous-espaces non-stationnaires à chaque nouveau vecteur arrivant correspondant à une image et comportant des objets mobiles. Cet algorithme permet de séparer le fond et les objets mobiles en ligne. GRASTA donne de résultats qualitatifs et quantitatifs intéressants en présence de fonds dynamiques et de variations lumineuses. Ensuite, He et al. [242] ont amélioré GRASTA en proposant l'algorithme t-GRASTA (transformed-GRASTA) qui permet de prendre en compte une transformation comme une rotation ou une translation. t-GRASTA est quatre fois plus rapide que les algorithmes standards, requiert moitié moins de mémoire et permet donc de gérer les scènes acquises lorsque la caméra vibre à cause du vent. Xu et al. [571] ont proposé une version plus générale dont GRASTA est un cas particulier. Cette méthode s'appelle GOSUS *Grassmannian Online Subspace Updates with Structured-sparsity* et permet d'apprendre de façon incrémentale un sous-espace quand les observations sont considérées issues d'un mélange de deux composantes. Bien que la norme l_1 dans GRASTA et GOSUS soit adaptée à l'optimisation, il est préférable d'employer la norme l_0 car celle-ci permet une reconstruction là où la norme l_1 échoue. Dans ce contexte, Hage et Kleinsteuber [230] ont amélioré GRASTA en utilisant la norme l_0 et en utilisant la méthode du gradient conjugué en tant que solveur. Cette méthode appelée pROST [451] est plus robuste que GRASTA en présence de fonds dynamiques. Les tableaux 5.9 et 5.16 montrent une vue concise de ces différentes approches de suivi de sous-espaces ainsi que des solveurs employés en terme de complexité. Les performances de temps de calcul de ces méthodes sont plus intéressantes que celles des méthodes ACP robustes avec 11 fps pour GRASTA, 2 fps pour GOSUS et 17 fps pour pROST pour des images de taille 144×176.

2.5 Minimisations en rang faible

La minimisation en rang faible (Low Rank Minimization) offre un cadre adapté pour la modélisation de fond. Zhou et al. [646] ont développé un cadre appelé DECOLOR (Detecting Contiguous Outliers in the Low-Rank Representation) qui permet la détection des outliers dans une représentation de rang faible dans laquelle la matrice de rang faible et le support des outliers sont estimés. Cette approche effectue la modélisation du fond et la détection des objets mobiles en un seul processus en résolvant la décomposition à l'aide un algorithme alternatif. D'une autre façon, Xiong et al. [646] ont proposé une méthode de factorisation robuste appelée DRMF (Direct Robust Matrix Factorization) en faisant l'hypothèse que seulement une petite partie de la matrice A est corrompue par les outliers. L'objectif est ainsi d'obtenir une estimation fiable de la matrice de rang faible et d'en identifier ainsi les outliers. Pour cela, ils sont exclus de l'estimation du modèle. Enfin, Xiong et al. [569] ont proposé une extension de DRMF quand les outliers occupent toute une colonne. Cette méthode est appelée DRMF-Row (DRMF-R). Différentes versions de LRM ont été employées et le tableau 5.6 montre une vue d'ensemble de ces versions ainsi que de leur solveur. Ces différentes approches sont présentées de façon détaillée dans notre article [78]. De plus, les tableaux 5.9 et 5.17 montrent une vue concise de ces différentes approches de minimisation

Low Rank Minimization	Categories	Authors - Dates
Décomposition	Contiguous Outlier Detection (DECOLOR)	Zhou et al. (2011) [646]
	Direct Robust Matrix Factorization (DRMF)	Xiong et al. (2011) [569]
	Direct Robust Matrix Factorization-Rows (DRMF-R)	Xiong et al. (2011) [569]
	Probabilistic Robust Matrix Factorization (PRMF)	Wang et al. (2012) [542]
	Bayesian Robust Matrix Factorization (BRMF)	Wang et al. (2013) [543]
	Practical Low-Rank Matrix Factorization (PLRMF)	Zheng et al. (2012) [636]
	Low Rank Matrix Factorization with MoG noise (LRMF-MOG)	Meng et al. (2013) [365]
	Unifying Nuclear Norm and Bilinear Factorization (UNN-BF)	Cabral et al. (2013) [85]
	Robust Rank Factorization (RRF)	Sheng et al. (2014) [465]
Solveurs	Alternating Algorithm (AA)	Zhou et al. (2011) [646]
	Block Coordinate Descent Strategy (BCDS)	Xiong et al. (2011) [569]
	Conditional EM Algorithm (CEM)	Jebara and Pentland (1999) [263]
	Augmented Lagrangian Multiplier (ALM)	Zheng et al. (2012) [636]
	Alternative Direction Descent Algorithm (ADDA)	Sheng et al. (2014) [465]

TABLE 5.6 – Vue d'ensemble des décompositions et des solveurs pour LRM. La première colonne indique les différentes catégories. La seconde colonne indique les différentes méthodes de chaque catégorie ainsi que son acronyme entre parenthèses. La troisième colonne donne le nom des auteurs et la date de la première publication de la méthode concernée.

en rang faible ainsi que des solveurs employés en terme de complexité.

2.6 Application à la détection d'objets mobiles

L'objectif de cette sous-section est de donner une idée concise du potentiel de ces méthodes et en dégager des pistes pour l'application à la modélisation de fond et la détection d'objets mobiles. Pour cela, nous présentons dans différents tableaux une vue synthétique des caractéristiques clés de ces différentes décompositions en RPCA, suivi de sous-espaces et de minimisation en rang faible. Dans cette idée, les tableaux 5.7, 5.8 et 5.9 montrent une vue synthétique des différentes décompositions en matrices de rang faible et parcimonieuse en termes de décomposition, de minimisation, de contraintes et de convexité.

Pour chaque méthode, nous avons étudié leur robustesse, leurs contraintes spatiales et temporelles, l'existence ou non d'une version incrémentale, d'une version temps-réel ou d'une évaluation quantitative. Les tableaux 5.10 et 5.11 montrent une vue synthétique relative à ces questions et le constat suivant peut être effectué :

1. **Robustesse :** La décomposition via PCP se base sur l'hypothèse que les coefficients de la matrice A peuvent être retrouvés exactement à partir des observations et sur l'hypothèse que la distribution des données corrompues est parcimonieuse et aléa-

toire sans bruit. Ces hypothèses ne sont que rarement vérifiées dans les applications réelles parce que la plupart du temps seulement une partie des entrées de la matrice A peut être observées et les observations sont corrompues par un bruit Gaussien ou autre. La robustesse de la décomposition via PCP peut être améliorée comme dans le cas de l'algorithme SPCP [647] pour le cas des données bruitées, par l'algorithme QPCP [55] pour prendre en compte la quantification des pixels, et par l'algorithme BPCP [509] dans le cas où un outlier occupe toute une colonne ou plusieurs colonnes. Les autres améliorations concernent le contrôle de la parcimonie, la garantie de reconstruction et les données bruitées.

2. **Contraintes spatiales et/ou temporelles :** Ces contraintes sont prises en compte dans les méthodes suivantes : BPCP [509], BRPCA [156] et DECOLOR [646]. Seul, BRPCA prend en compte les deux contraintes.

3. **Algorithmes incrémentaux :** Ces algorithmes existent pour les décompositions suivantes : RPCA-PCP [415][417][416][418], RPCA-SPCP [489], RPCA-SpaCtrl [361], GRASTA [240], t-GRASTA [242] et pROST [230][451].

4. **Implémentations temps-réel :** Il en existe que pour PCP [20][510] et pour SPCP [349].

Toutes ces points doivent être pris en compte pour pouvoir appliquer de façon robuste les décomposition en matrices de rang faible et parcimonieuse à la modélisation de fond et la détection d'objets mobiles en vidéo-surveillance.

2.7 Discussion

Ces récentes avancées en ACP robustes (RPCA), suivi de sous-espaces (Subspace Tracking) et minimisation en rang faible (Low Rank Minimization) sont donc fondamentales et offrent un potentiel important pour la modélisation de fond et la détection d'objets mobiles en vidéo-surveillance à l'image du mélange de Gaussiennes à ses débuts. Cependant, aucun algorithme n'émerge et ne permet de pallier simultanément les principaux challenges rencontrés dans les vidéos. Cela est en partie dû aux causes suivantes :

1. Absence de recherches vraiment dédiées à l'application de séparation fond/objets mobiles. La plupart des auteurs sont de la communauté "mathématique" et montrent l'application possible de leur modèle à différentes applications en traitement des images et des vidéos comme le débruitage [211], la détection de salience [578], le codage [631][219][108], l'extraction de *key frames* [144], le traitement de vidéo dans le do-

maine hyperspectral [197] et bien sûr la séparation fond/objets mobiles [19][389].

2. Absence d'une évaluation quantitative rigoureuse sur une base de données large pour couvrir tous les challenges possibles et possédant la vérité-terrain pour toutes les images de chaque séquence.

En effet, les auteurs comparent usuellement de façon qualitative leur méthode à l'algorithme RSL [517] ou PCP [92]. Des évaluations quantitatives récentes basées sur les métriques du domaine ont été effectuées mais sont limitées à un algorithme. Par exemple, Wang et al. [551] ont seulement évalué la méthode RPCA-PCP résolue par le solveur APG [326]. Xue et al. [578] et Yang [585] ont quant-à-eux évalué la méthode RPCA-PCP résolue par le solveur IALM [325]. De notre coté, nous avons évalué la méthode RPCA-PCP résolue par le solveur LADMAP [266] ainsi que la méthode BPCP [509] pour la détection d'objets mobiles [222][221]. Les résultats ont montré que la méthode BPCP est plus robuste que la méthode PCP.

Dans une première étude comparative [227], nous avons comparé cinq algorithmes : RSL [517], RPCA-PCP résolue par le solveur EALM [325], RPCA-PCP résolue par le solveur IALM [325], QPCP [55] et BRPCA [156] sur les bases de données "Wallflower" [521], d'I2R [311] et de Carnegie Mellon [461]. Les résultats ont montré que la méthode BRPCA qui prend en compte les contraintes spatiales et temporelles est la plus robuste. Une étude similaire menée par Rueda et al. [440] a comparé RPCA-PCP résolue par le solveur EALM, BRPCA et GoDec. Cette comparaison a aussi mis en évidence la robustesse de la méthode BRPCA. Cependant dans les deux cas, l'évaluation était limitée à quelques méthodes et ne portait pas sur une base de données suffisamment large pour couvrir tous les challenges possibles et possédant la vérité-terrain pour toutes les images de chaque séquence.

Considérant tout cela, nous avons d'abord effectué une évaluation complète de 19 algorithmes parmi ceux de l'ACP robustes, du suivi de sous-espaces et de la minimisation en rang faible sur la base de données BMC 2012. Celle-ci est présentée dans la Section 5.1.

Ensuite, nous avons développé deux méthodes spécifiques pour la détection d'objets mobiles qui sont l'ACP robuste par "Iteratively Reweighted Least Squares" (IRLS) (Section 3) et l'ACP polynomiale (Section 4).

L'objectif de l'ACP robuste par "Iteratively Reweighted Least Squares" (IRLS) est de pallier une limitation de la méthode présentée par Candès et al. [92] en prenant en compte l'aspect spatial (objets mobiles) pour obtenir une estimation robuste de la matrice L de rang faible et de la matrice parcimonieuse S.

L'approche basée l'ACP polynomiale a quant à elle l'objectif de prendre en compte les effets non linéaires qui apparaissent localement dans le voisinage des intensités des pixels, comme ceux qu'engendrent des variations lumineuses soudaines, des variations des paramètres de la caméra (auto-ajustements) et le bruit lors de l'acquisition. Cette méthode permet

de plus de prendre en compte l'aspect additif de certaines variations et l'aspect multiplicatif d'autres variations.

Catégories	Méthodes	Décomposition	Minimisation	Contraintes	Convexité
RPCA-PCP	**PCP** **Candès et al. [92]**	$A = L + S$	$\min_{L,S} \ ||L||_* + \lambda||S||_1$	$A - L - S = 0$	Yes
	PCP Leow et al. [308]	$A = L + S$	$\min_{L,S} \ ||L||_* + \lambda||S||_1$	$rank(L) = known \ \ r$	Yes
	PCP Yuan et al. [603]	$A = u1^T + S$	$\min_{u} ||A - u1^T||_1$	$rank(u1^T) = 1$	Yes
	PCP Bao et al. [50]	$A = PA + S$	$\min_{P,S} ||P||_* + \lambda||S||_1$	$A - PA - S = 0$	Yes
	SPCP **Zhou et al. [647]**	$A = L + S + E$	$\min_{L,S} \ ||L||_* + \lambda||S||_1$	$||A - L - S||_F < \delta$	Yes
	SPCP Zhou and Tao [477]	$A = UV + S + E$	$\min_{U,V,S} \ \lambda||S||_1 + ||A - UV - S||_F^2$	$rank(U) = rank(S) \leq r$	Yes
	SPCP Li et al. [603]	$A = U1^T + S + E$	$\min_{S \in \mathbf{R}^{m \times n}, u \in \mathbf{R}^m} ||S||_1 + \frac{\mu}{2}||A - u1^T - S||_F^2$	$rank(U1^T) = 1$	Yes
	SPCP Li et al. [603]	$A = U1^T + S + E$	$\min_{S \in \mathbf{R}^{m \times n}, u \in \mathbf{R}^m} ||S||_1 + \frac{\mu}{2}||A - H(u1^T + S)||_F^2$	$rank(U1^T) = 1$	Yes
	SPCP Sprechman et al. [489]	$A = UT + S + E$	$\min_{U,T,S} \lambda||S||_1 + \frac{\lambda_1}{2}(||U||_F^2 + ||T||_F^2) + \frac{1}{2}||A - UT - S)||_F^2$	$rank(UT) \leq r$	Yes
	QPCP **Becker et al. [55]**	$A = L + S$	$\min_{L,S} \ ||L||_* + \lambda||S||_1$	$||A - L - S||_\infty < 0.5$	Yes
	BPCP **Tang and Nehorai [509]**	$A = L + S$	$\min_{L,S} \ ||L||_* + \kappa(1 - \lambda)||L||_{2,1} + \kappa\lambda||S||_{2,1}$	$A - L - S = 0$	Yes
	LPCP **Wohlberg et al. [560]**	$A = AU + S$	$\min_{U,S} \ \alpha||U||_* + \beta||U||_{2,1} + \beta||S||_1$	$A - AU + S = 0$	Yes

TABLE 5.7 – Vue comparative des décompositions en matrices de rang faible et parcimonieuse par RPCA-PCP en termes de minimisation, de contraintes et de convexité.

Catégories	Méthodes	Décomposition	Minimisation	Contraintes	Convexité
RPCA-OP	OP Xu et al. [570]	$A = L + S$	$\min_{L,S} \ ||L||_* + \lambda ||S||_{1,2}$	$A - L - S = 0$	Yes
	SOP Xu et al. [570]	$A = L + S + E$	$\min_{L,S} \ ||L||_* + \lambda ||S||_{1,2}$	$||A - L - S||_F < \delta$	Yes
RPCA-SpaCtrl	Sparsity Control Mateos et Giannakis [361]	$A = M + UT + S + E$	$\min_{U,S} \ ||X + 1_N M^{'} - TU^{'} - S||_F^2 + \lambda ||S||_{2,r}$	$UU^{'} = I_q$	Yes
RPCA-SpaCorr	Sparse Corruptions (case 1) Hsu et al. [252]	$A = L + S$	$\min_{L,S} ||L||_* + \lambda ||S||_1$	$||A - L - S||_1 \leq \epsilon_1$ $||A - L - S||_* \leq \epsilon_*$ $||L||_\infty \leq b$	Yes
	Sparse Corruptions (case 2) Hsu et al. [252]	$A = L + S$	$\min_{L,S} ||L||_* + \lambda ||S||_1 + \frac{1}{2\mu} ||A - L - S||_2^2$	$||A - L - S||_1 \leq \epsilon_1$ $||A - L - S||_* \leq \epsilon_*$ $||A - S||_\infty \leq b$	Yes
RPCA-LHR	LHR Deng et al. [150]	$A = L + S$	$\min_{\hat{X} \in \hat{D}} \frac{1}{2}(||Diag(Y)||_+ ||Diag(Z)||_1 + \lambda ||S||_1$	$\hat{X} = \{Y, Z, L, S\}$ $\hat{D} = \left\{(Y, Z, L, S) : \begin{pmatrix} Y & L \\ L^T & Z \end{pmatrix} \geq 0, (L, S) \in C\right\}$	No
Bayesian RPCA	BRPCA Ding et al. [156]	$A = D(ZG)W + BX + E$	$-logp(\Theta \mid A, H)$	Distribution constraints	-
	VBRPCA Babacan et al. [36]	$A = DB^T + S + E$	$p(A, D, B, S)$	Distribution constraints	-
Approximated RPCA	GoDec Zhou and Tao [643]	$A = L + S + E$	$\min_{L,S} \ ||A - L - S||_F^2$	$rank(L) \leq e, card(S) \leq k$	No
	Semi-Soft GoDec Zhou and Tao [643]	$A = L + S + E$	$\min_{L,S} \ ||A - L - S||_F^2$	$rank(L) \leq e, card(S) \leq \tau$ τ is a soft threshold	No

TABLE 5.8 – Vue comparative des décompositions en matrices de rang faible et parcimonieuse par RPCA en termes de minimisation, de contraintes et de convexité

Catégories	Méthodes	Décomposition	Minimisation	Contraintes	Convexité
Subspace Tracking	GRASTA He et al. [240]	$A = UW + S + E$	$\min_{W} \ \|\|UW - A\|\|_1$	$A - UW - S = 0$	No
	t-GRASTA He et al. [242]	$A = UW + S + E$	$\min_{W} \ \|\|UW - A\|\|_1$	$A - UW - S = 0$	No
	pROST Hage and Kleinsteuber [230][451]	$A = L + S$	$\min_{Rank(L) \leq k} \ \|\|L - A\|\|_0$	$A - L = 0$	No
	GOSUS Xu et al. [571]	$A = UW + S$	$\min_{U^T U = I_{d,w,x}} \ \sum_{i=1}^{l} \mu_i \|\|D^i x\|\|_2 + \frac{\lambda}{2} \|\|Uw + x - a\|\|_2^2$	$A - UW - S = 0$	No
Low-Rank Minimization	Contiguous Outlier Detection (DECOLOR) Zhou et al. [646]	$A = L + S + E$	$\min_{L,S} \ \frac{1}{2} \|\|P_{\bar{F}}(A - L)\|\|_F^2$ $+\alpha \|\|L\|\|_* + \beta \|\|F\|\|_1 + \gamma \|\|Cvec(F)\|\|_1$	$rank(L) \leq K$	No
	Direct Robust Matrix Factorization (DRMF) Xiong et al. [569]	$A = L + S$	$\min_{L,S} \|\|A - S - L\|\|_F$	$rank(L) \leq K, \|\|S\|\|_0 \leq e$	No Original formulation PCP
	Direct Robust Matrix Factorization-Row (DRMF-R) Xiong et al. [569]	$A = L + S$	$\min_{L,S} \|\|A - S - L\|\|_F$	$rank(L) \leq K, \|\|S\|\|_{2,0} \leq e$	No Original formulation OP
	Probabilistic Robust Matrix Factorization (PRMF) Wang et al. [542]	$A = U\acute{V} + S$	$logp(U, V \| A, \lambda, \lambda_U, \lambda_V)$	Distribution Constraints	No
	Bayesian Robust Matrix Factorization (BRMF) Wang et al. [543]	$A = U\acute{V} + S$	$logp(U, V \| A, \lambda, \lambda_U, \lambda_V)$	Bayesian Distribution Constraints	No

TABLE 5.9 – Vue comparative des décompositions par suivi de sous-espaces et minimisation de rang faible en termes de minimisation, de contraintes et de convexité.

Catégories	Méthodes	Robustesse	Contraintes spatiales et temporelles	Version incrementale	Implémentation temps réel	Validation
RPCA-PCP	PCP Candès et al. [92]	Sparse noise	-	Recursive Robust PCP (RR-PCP[1]) [415] Recursive Project Compressive Sensing (ReProCS[1]) [416][419] Support-Predicted Modified-CS RR-PCP[1] [417] Support-Predicted Modified-CS[1] [416] Automated ReProCS[2] [405] Practical ReProCS (Prac-ReProCS) [420]	CAQR[3] [20] Real-Time PCP[4] [406]	APG [551] IALM [578], IALM [585] EALM, IALM [227] LADMAP [222] LSADM [225]
	SPCP Zhou et al. [136]	Entry-wise noise	-	Fast Trainable Encoders [489]	DFC-PROJ[5] [27] DFC-PROJ-ENS[5] [349] DFC-NYS[5] [349] DFC-NYS-ENS[5] [27]	-
	QPCP Becker et al. [55]	Quantization error	-	-	-	Becker et al. [55] Guyon et al. [227]
	BPCP Tang and Nehorai [509]	Outliers in entire columns	Spatial	-	-	Guyon et al. [221]
	LPCP Wohlberg et al. [560]	Multimodal backgrounds	-	-	-	Wohlberg et al. [560]
RPCA-OP	OP Xu et al. [570]	Outliers in entire columns	-	-	-	-
	SOP Xu et al. [570]	Outliers in entire columns Entry-wise noise	-	-	-	-
RPCA-SpaCtrl	Sparsity control Mateos et Giannakis [361]	Sparsity control	-	Yes	-	Mateos et Giannakis [361]
RPCA-SpaCorr	Sparse Corruptions (case 1) Hsu et al. [252]	Recovery guarantees	-	-	-	-
	Sparse Corruptions (case 2) Hsu et al. [252]	Recovery guarantees	-	-	-	-
RPCA-LHR	LHR Deng et al. [150]	High intrinsic rank structure Dense outliers	-	-	-	Deng et al. [150]

TABLE 5.10 – Vue d'ensemble des développements concernant les décompositions en matrice de rang faible et parcimonieuse pour la problèmatique de la séparation fond/objets mobiles en termes de robustesse, de contraintes spatiales et temporelles, de version incrémentale, d'implémentation temps-réel et de validation sur des bases de données.

Catégories	Méthodes	Robustesse	Contraintes spatiales et temporelles	Version incrémentale	Implémentation temps-réel	Validation
Bayesian RPCA	BRPCA Ding et al. [156]	Entry-wise noise	Spatial and temporal	-	-	Ding et al. [156] Guyon et al. [227]
	VBRPCA Babacan et al. [36]	Entry-wise noise	-	-	-	Babacan et al. [36]
Approximated RPCA	GoDec Zhou and Tao [643]	Entry-wise noise	-	-	-	Zhou and Tao [643]
	Semi-Soft GoDec Zhou and Tao [643]	Entry-wise noise	-	-	-	Zhou and Tao [643]
Subspace Tracking	GRASTA He et al. [240]	-	-	Yes	-	He et al. [240]
	t-GRASTA He et al. [242]	-	-	Yes	-	He et al. [242]
	pROST Hage and Kleinsteuber [230]	-	-	Yes	-	Hage and Kleinsteuber [451]
	GOSUS Xu et al. [571]	-	-	Yes	-	Xu et al. [571]
Low Rank Minimization	Contiguous Outlier Detection (DECOLOR) Zhou et al. [646]	Entry-wise noise	Spatial	-	-	Zhou et al. [646]
	Direct Robust Matrix Factorization (DRMF) Xiong et al. [569]	- -	-	-	-	Xiong et al. [569]
	Direct Robust Matrix Factorization-Row (DRMF-R) Xiong et al. [569]	Entry-wise noise	-	-	-	-
	Probabilistic Robust Matrix Factorization (PRMF) Wang et al. [542]	Entry-wise noise	-	Yes	-	Wang et al. [542]

TABLE 5.11 – Vue d'ensemble des développements concernant les décompositions en matrice de rang faible et parcimonieuse pour la problèmatique de la séparation fond/objets mobiles.

1. http://home.engineering.iastate.edu/ chenlu/ReProCS/ReProCS.htm
2. http://home.engineering.iastate.edu/ chenlu/ReProCS/ReProCs-auto/results.htm
3. Available on request by email to the corresponding author
4. http://www.nari.ee.ethz.ch/commth/research/downloads/
5. http://www.cs.berkeley.edu/ ameet/dfc/

Catégorie	Solveur	Complexité
Basic algorithms	Singular Value Threshold (SVT[6]) Cai et al. [86]	$O(mnmin(mn))$
	Iterative Thresholding (IT) Wright et al. [564]	$O(mnmin(mn))$
	Accelerated Proximal Gradient (APG)[6]) Lin et al.[326]	$O(mnmin(mn))$ Full SVD
	Dual Method (DM)[6]) Lin et al.[326]	$O(rmn)$ Partial SVD
	Exacted Augmented Lagrangian Method (EALM) (EALM[6]) Lin et al.[325]	$O(mnmin(mn))$ Full SVD
	Inexact Augmented Lagrangian Method (IALM) (IALM[6]) Lin et al. [325]	$O(rmn)$ Partial SVD, Linear Time SVD [586], Limited Memory SVD [332]
	Alternating Direction Method (ADM) (LRSD[6]) Yuan and Yang [604]	$O(mnmin(mn))$
	Symmetric Alternating Direction Method (SADM[7]) Ma [348], Goldfarb et al. [199]	$O(1/\epsilon)$ Iteration complexity
	Non Convex Splitting ADM (NCSADM) Chartrand [104]	- -
	Variant of Douglas-Rachford Splitting Method (VDRSM) Zhang and Liu [619]	- -
	Proximity Point Algorithm (PPA) Zhu et al. (2014) [648]	- -
	Proximal Iteratively Reweighted Algorithm (PIRA) Wang et al. (2014) [538]	- -
	Alternating Rectified Gradient Method (ARGM) Kim et al. (2014) [278]	- -
	Parallel Direction Method of Multipliers (PDMM) Wang et al. (2014) [533]	- -

TABLE 5.12 – Vue comparative des solveurs basiques pour la décomposition RPCA-PCP en terme de complexité.

6. http://perception.csl.illinois.edu/matrix-rank/samplecode.html
7. Available on request by email to the corresponding author

Catégorie	Solveur	Complexité
Linearized algorithms	Linearized Augmented Lagrangian Method (LALM) Yang and Yuan [584]	$O(mnmin(mn))$
	Linearized Alternating Direction Method (LADM) Yang and Yuan [584]	$O(mnmin(mn))$
	Linearized Alternating Direction Method with Adaptive Penalty (LADMAP[7]) Lin et al. 2011 [327]	$O(rmn)$ Accelerated version
	Linearized Symmetric Alternating Direction Method (LSADM[7]) Ma [348], Goldfarb et al. [199]	$O(1/\epsilon)$ Iteration complexity
	Fast Linearized Symmetric Alternating Direction Method (Fast-LSADM[7]) Ma [348], Goldfarb et al. [199]	$O(1/\sqrt{\epsilon})$ Iteration complexity
	Linearized Alternating Direction Method (LADM) (LMaFit[8]) Shen et al. [464]	-

TABLE 5.13 – Vue comparative des solveurs linéaires pour la décomposition RPCA-PCP en terme de complexité.

Catégorie	Solveur	Complexité
Fast algorithms	Randomized Projection for ALM (RPALM) Mu et al. [373]	$O(pmn)$
	l_1 filtering (LF[9]) Liu et al. [331]	$O(r(m+n))$
	Block Lanczos with Warm Start Lin and Wei [328]	-
	Exact Fast Robust Principal Component Analysis (EFRPCA) Abdel-Hakim and El-Saban [8]	$O(mk^2)$ with $k \ll n$ Full SVD
	Inexact Fast Robust Principal Component Analysis (IFRPCA) Abdel-Hakim and El-Saban [8]	$O(mk^2)$ with $k \ll n$ Partial SVD
	Matrix Tri-Factorization (MTF) Liu et al. [335]	$O(n^3 + t(r^3 + r^2n + mn^2 + rn^2))$ Iteration Complexity
	Fast Tri-Factorization(FTF) Liu et al. [334]	$O(t(r^3 + r^2(m+n) + r^2m + rmn))$ Iteration Complexity
	PRoximal Iterative SMoothing Algorithm (PRISMA) Orabona et al. [388]	-
	Fast Alterning Minimization (FAM)[10] Rodriguez and Wohlberg [434]	-

TABLE 5.14 – Vue comparative des solveurs rapides pour la décomposition RPCA-PCP en terme de complexité.

8. http://lmafit.blogs.rice.edu/
9. Available on request by email to the corresponding author
10. https://sites.google.com/a/istec.net/prodrig/Home/en/pubs

Méthodes	Solveurs	Complexité
RPCA via Stable Principal Component Pursuit (RPCA-SPCP) Zhou et al. [136]	Alternating Splitting Augmented Lagrangian method (ASALM) Tao and Huan [512]	-
	Variational ASALM (VASALM[11]) Tao and Huan [512]	-
	Parallel ASALM (PSALM) Tao and Huan [512]	-
	Non Smooth Augmented Lagrangian Algorithm (NSA[12]) Aybat et al. [30]	$O(1/\epsilon)$ Iteration complexity
	Partially Smooth Proximal Gradient (PSPG[12]) Aybat et al. [31]	$O(1/\epsilon)$ Iteration complexity
RPCA via Stable Principal Component Pursuit (RPCA-SPCP) Zhou and Tao [644]	Greedy Bilateral Smoothing (GreBsmo) Zhou and Tao [644]	-
	Bilinear Generalized Approximate Message Passing (BiG-AMP) Parker and Schniter [394]	-
RPCA via Quantized PCP (RPCA-QPCP) Becker et al. [55]	Templates for First-Order Conic Solvers (TFOCS[12]) Becker et al. [55]	$O(1/\epsilon)$ Iteration complexity
RPCA via Block based Principal Component Pursuit (RPCA-BPCP) (RPCA-LBD[11]) Tang and Nehorai [509]	Augmented Lagrangian Method (ALM) Tang and Nehorai [509]	$O(mnmin(mn))$
RPCA via Local Principal Component Pursuit (RPCA-LPCP) Wohlberg et al. [560]	Split Bregman Algorithm (SBA) Goldstein and Osher [200]	-
RPCA via Outlier Pursuit (RPCA-OP[14]) Xu et al. [570]	Singular Value Threshold (SVT) Cai et al. [86]	$O(mnmin(mn))$
RPCA via Stable Outlier Pursuit(RPCA-SOP[14]) Xu et al. [570]	Singular Value Threshold (SVT) Cai et al. [86]	$O(mnmin(mn))$
RPCA with Sparsity Control (RPCA-SpaCtrl) Mateos and Giannakis [361]	Alternating Minimization (AM) Zhou et al. [650]	-
RPCA via Sparse Corruptions (RPCA-SpaCorr) Hsu et al. [252]	- -	-
RPCA via Log-sum Heuristic Recovery (RPCA-LHR) Deng et al. [150]	Majorization-Minimization (MM) Fazel [180], Lange et al. [301]	-
Bayesian Robust Component Analysis (B-RPCA[15]) Ding et al. [156]	Markov chain Monte Carlo (MCMC) Robert and Cassela [432]	$O(r(m+n)+mn)$
	Variational Bayesian Inference (VB) Beal [54]	-
Variational Bayesian Robust Component Analysis (VB-RPCA[16]) Babacan et al. [36]	Approximate Bayesian Inference (AB) Beal [54]	$O(min(n^3,r^3)+min(m^3,r^3))$ per iteration
Approximated RPCA (A-RPCA) (GoDec[17]) Zhou and Tao [643]	Naive GoDec Zhou and Tao [643]	-
	Fast Godec via Bilateral Random Projection Zhou and Tao [643]	-

TABLE 5.15 – Vue comparative des solveurs pour les décompositions RPCA en terme de complexité.

11. Available on request by email to the corresponding author
12. http://www2.ie.psu.edu/aybat/codes.html
13. http://www.ugcs.caltech.edu/ srbecker/rpca.shtml
14. http://guppy.mpe.nus.edu.sg/ mpexuh/publication.html
15. http://people.ee.duke.edu/ lcarin/BCS.html
16. http://www.dbabacan.info/software.html
17. http://sites.google.com/site/godecomposition/code

Suivi de sous-espaces	Solveurs	Complexité
GRASTA[18] He et al. [240]	Augmented Lagrangian algorithm with the Grassmannian geodesic gradient descent Boyd et al. [81], Edelman et al. [166]	$O(\|\Omega\| d^3 + Kd\|\Omega\| + nd^2)$
t-GRASTA He et al. [240]	Augmented Lagrangian algorithm with the Grassmannian geodesic gradient descent Boyd et al. [81], Edelman et al. [166]	$O(\|\Omega\| d^3 + Kd\|\Omega\| + nd^2)$
pROST[19] Hage and Kleinstauber [230]	Conjugate Gradient Hage and Kleinstauber [230]	-
GOSUS[20] Xu et al. [571]	Alternating Direction Method of Multipliers (ADMM) Xu et al. [571]	-

TABLE 5.16 – Vue comparative des solveurs pour le suivi de sous-espaces (Subspace tracking (ST)) en terme de complexité.

Minimisation en rang faible	Solveurs	Complexité
Contiguous Outliers Detection (DECOLOR[21]) Zhou et al. [646]	Alternating Algorithm Zhou et al. [646]	-
DRMF[22] Xiong et al. [569]	Block Coordinate Descent strategy Xiong et al. [569]	$O(mn(K + log(e)))$
DRMF-R[22] Xiong et al. [569]	Block Coordinate Descent strategy Xiong et al. [569]	$O(mn(K + log(e)))$
PRMF[23] Wang et al. [542]	Conditional EM Algorithm (CEM) Jebara and Pentland [263]	-
BRMF[24] Wang and Yeung [543]	Conditional EM Algorithm (CEM) Jebara and Pentland [263]	-

TABLE 5.17 – Vue comparative des solveurs pour la minimisation en rang faible (Low Rank Minimization(LRM)) en terme de complexité.

18. http://sites.google.com/site/hejunzz/grasta
19. http://www.gol.ei.tum.de/index.php?id=37 L=1
20. http://pages.cs.wisc.edu/ jiaxu/projects/gosus/
21. http://bioinformatics.ust.hk/decolor/decolor.html
22. http://www.autonlab.org/autonweb/downloads/software.html
23. http://winsty.net/prmf.html
24. http://winsty.net/brmf.html

3 ACP Robuste par Iteratively Reweighted Least Squares

Dans cette section, nous présentons un algorithme pour la détection d'objets mobiles basé sur la factorisation robuste de matrice [225][224][223]. La matrice A est considérée comme approximativement de rang faible avec une faible partie corrompue par la présence des objets mobiles. L'objectif est de pallier une limitation de la méthode présentée par Candès et al. [92] en prenant en compte l'aspect spatial pour obtenir une estimation robuste de la matrice L de rang faible et de la matrice parcimonieuse S.

3.1 Minimisation avec contraintes spatiales

Nous avons modifié la décomposition (5.1) en ajoutant un troisième terme pour modéliser le bruit et ajouter la contrainte spatiale lors de la minimisation. Soit le problème de minimisation suivant où A est considéré comme une matrice dictionnaire et b un vecteur colonne :

$$\underset{x}{\text{argmin}} \ ||Ax - b||_\alpha + \mu||grad(s)\phi(E)||_1, \ \ E = |b - Ax| \tag{5.11}$$

où $grad(s)$ est le gradient spatial. De cette façon, l'erreur doit avoir une forme connexe. Par l'utilisation de la mesure de Variation Totale, celle-ci est minimale. On peut remarquer que le problème est convexe pour $\alpha > 1$ et que l'algorithme IRLS (Iteratively Reweighted Least Squares) est donné par :

$$\underset{x}{\text{argmin}} \ ||Ax - b||_\alpha \quad \left| \quad \begin{array}{l} D^{(i)} = \text{diag}((\varepsilon + |b - Ax^{(i)}|)^{\alpha-2}) \\ x^{(i+1)} = (A^t D^{(i)} A)^{-1} A^t D^{(i)} b \end{array} \right. \tag{5.12}$$

L'algorithme IRLS converge pour $1 < \alpha < 3$. Nous avons apporté une modification pour avoir plus de stabilité numérique. Ainsi, nous fixons $\alpha \in [1, \infty[$ avec λ_{opt} adaptatif :

$$r^{(i)} = b - Ax^{(i)} D = \text{diag}((\epsilon + |r^{(i)}|)^{\alpha-2}) y^{(i)} = (A' D A)^{-1} A' D r^{(i)} x^{(i+1)} = x^{(i)} + (1 + \lambda_{opt}) y^{(i)} \tag{5.13}$$

L'algorithme est iteratif de façon alternative et permet ainsi d'obtenir le x et le λ optimaux à chaque itération.

$$c^{(i)} = A y^{(i)} d^{(i)} = b - A(x^{(i)} + y^{(i)})\lambda_{opt} = \underset{\lambda}{\text{argmin}} \ ||c^{(i)}\lambda - d^{(i)}||_\alpha \tag{5.14}$$

$$\begin{array}{l} \lambda^{(0)} = \Lambda(\alpha) \\ s^{(k)} = d - c\lambda^{(k)} \\ E = \text{diag}((\epsilon + |s^{(k)}|)^{\alpha-2}) \\ \lambda^{(k+1)} = \lambda^{(k)} + (1 + \Lambda(\alpha)) \frac{c^t E s^{(k)}}{c^t E c} \end{array} \tag{5.15}$$

Pour la valeur de λ_{opt}, nous pouvons choisir $\lambda_{opt} = \Lambda(\alpha)$ pour permettre $1 < \alpha < \infty$ et ainsi trois cas peuvent être distingués :

$$\Lambda(\alpha) = \begin{cases} 5/6 & \text{if } \alpha \leq 1 \\ -\frac{2}{3}\alpha + \frac{3}{2} & \text{if } 1 < \alpha < 1 + \frac{3}{4} \\ \frac{1}{\alpha - 1} - 1 & \text{if } \alpha \geq 1 + \frac{3}{4} \end{cases} \tag{5.16}$$

Pour le cas $\alpha > 2$, la convergence est obtenue quand $0 < 1 + \lambda < \frac{2}{\alpha - 1}$. De façon plus générale, nous résolvons le problème de régression avec les deux paramètres de la norme (α, β) et une matrice de poids (W).

$$\min_X \ ||AX - B||_{\substack{\alpha,\beta \\ W}} \quad \text{where} \quad ||M_{ij}||_{\substack{\alpha,\beta \\ W}} = (\sum_{i=1}^{n}(\sum_{j=1}^{m} W_{ij} \ |M_{ij}|^{\beta})^{\frac{\alpha}{\beta}})^{\frac{1}{\alpha}} \tag{5.17}$$

Cette minimisation est résolue par l'algorithme suivant :

Tant que X est stable, répeter sur les k colonnes

$$\left| \begin{array}{rcl} R & \leftarrow & B - AX \\ S & \leftarrow & \varepsilon + |R| \\ D & \leftarrow & diag(S_{ik}^{\beta-2} \circ (\sum_j (S_{ij}^{\beta} \circ W_{ij}))^{\frac{\alpha}{\beta}-1} \circ W_{ik}) \\ X_{ik} & \leftarrow & X_{ik} + (1 + \Lambda(\max(\alpha, \beta)))(A^t D A)^{-1} A^t D R_{ik} \end{array} \right. \tag{5.18}$$

3.2 Détection des objets mobiles par factorisation robuste de la matrice de rang faible

Soit la séquence vidéo d'initialisation $A = \left\{I_1, ...I_m\right\}, A \in \mathbb{R}^{n \times m}$ où I_j correspond à une image stockée sous forme d'un vecteur colonne à l'instant j et m est le nombre d'images pour l'initialisation. Chaque pixel (x, y)est caractérisé par son niveau de gris. La décomposition proposée est donc la suivante :

$$A = L + S = BC + S \tag{5.19}$$

où B est la matrice de rang faible correspondant au fond plus le bruit, et la matrice C permet de reconstruire L par combinaison linéaire. S est la matrice qui correspond aux objets mobiles. L'erreur de reconstruction relative à cette décomposition est définie comme suit :

$$\min_{B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times m}} ||(A - BC) \circ W||_{\alpha,\beta} + \mu ||BC||_* \tag{5.20}$$

où $||.||_*$ est la norme nucléaire. Cette décomposition est séparée en deux parties. Dans un premier temps, nous suivons la décomposition de rang 1 parce que les premiers vecteurs propres correspondent à ce qui apparait le plus souvent, c'est-à-dire le fond.

$$\left| \begin{array}{l} R_1 = A - B_1 C_1 \\ R = A - B_1 C_1 - B_r C_r \end{array} \right| \begin{array}{l} (1) \ \min\limits_{B_1, C_1} ||R_1||_{1,1} \\ (2) \ \min\limits_{B_r, C_r} ||R \circ \phi(R_1)||_{2,1 \to 0} \end{array} \tag{5.21}$$

Dans un second temps, nous utilisons la norme $||.||_{2,1\rightarrow 0}$ à la place de la norme usuelle $||.||_{1,1}$ pour forcer la connexité spatiale. $\beta = (1 \rightarrow 0)$ signifie que le paramètre β diminue à chaque itération. Nous cherchons donc d'abord une solution au problème convexe $||.||_{2,1}$. Ensuite, nous utilisons cette solution comme solution point de départ pour l'optimisation non-convexe $||.||_{2,(1-\epsilon)}$. Finalement, nous trouvons un minimum local de $||.||_{2,0}$ considéré comme proche du minimum global.

Dans le cas où $\alpha = \beta = 2$, la décomposition est résolue via un algorithme de type SVD (Singular Value Decomposition). Cet algorithme SVD peut être considéré comme un algorithme de régression itérative. La méthode proposée détermine ainsi alternativement les coefficients optimaux, c'est-à-dire qu'on cherche C pour B fixé et vice-versa.

$$\begin{aligned} C^{(k+1)} &= (A^t A)^{-1} A^t B^{(k)} \\ \bar{C}^{(k+1)} &= C^{(k+1)} \sqrt{C^{t(k+1)} C^{(k+1)}}^{-1} \\ B^{(k+1)} &= (A^t A)^{-1} A^t \bar{C}^{(k+1)} \end{aligned} \tag{5.22}$$

De plus, cette méthode de régression alternative permet d'utiliser une matrice de poids W qui pondére l'erreur pour chaque coefficient comme suit :

$$\underset{B,C}{min} ||(A - BC) \circ W||_{\alpha,\beta} \tag{5.23}$$

où $\circ$ est le produit de Hadamard. Nous définissons ensuite une fonction ϕ qui a pour objectif de lisser l'erreur comme un filtre médian spatial, et de transformer l'erreur de sorte à la traiter sous forme de régression.

$$W = \exp(-\sigma\, \phi(|A - BC|)) \tag{5.24}$$

En ajoutant une pénalité locale comme une contrainte dans la décomposition, nous augmentons ainsi la coherence locale de l'erreur.

3.3 Discussion

L'algorithme IRLS est conceptuellement simple, facile à implémenter et efficace pour la décomposition en rang faible. De plus, les expérimentations (voir Section 5.2) montrent que cette approche est plus robuste que les méthodes PCP classiques dans la plupart des cas. En effet, l'utilisation d'un terme de contrainte spatiale dans la minimisation renforce cette performance en présence de fonds dynamiques et de variations lumineuses. L'algorithme IRLS est donc une bonne alternative aux approches Lagrangiennes classiques. Finalement, il reste donc en à développer une version incrémentale et temps réel pour être utilisable dans une application comme la vidéo-surveillance.

4 ACP robuste polynomiale

Dans cette section, nous présentons un algorithme pour la détection d'objets mobiles basé une ACP polynomiale. L'aspect polynomial permet de prendre en compte les effets non linéaires qui apparaissent localement dans le voisinage des intensités des pixels. Ces effets non-linéaires proviennent des variations lumineuses soudaines, des variations des paramètres de la caméra (auto-ajustements) et du bruit lors de l'acquisition. Cette méthode permet de plus de prendre en compte l'aspect additif de certaines variations et l'aspect multiplicatif d'autres variations. Lors de la minimisation, nous utilisons comme dans la méthode de la section précédente la norme $||.||_{\alpha,\beta}$ et l'algorithme IRLS.

La décomposition proposée est identique à celle de l'équation 5.19. Par contre, pour prendre en compte les effets non linéaires, la mimisation est formulée comme suit :

$$\min_{B,C} \; rank(\phi(B)C) + \lambda||S|| \;\; avec\, les\, contraintes \;\; \begin{cases} A = BC + S \\ det(CC') = 1 \end{cases}$$

$$\text{où} \quad \phi(.) = \; carte\, nonlinaire \tag{5.25}$$

$\phi()$ est une fonction non linéaire de $\mathbb{R}^p$ à $\mathbb{R}^q$ avec $p < q$. L'algorithme IRLS est utilisé pour miminiser de façon alternative l'équation 5.25 et obtenir ainsi les matrices $\phi(B)$ et C.

Pour la fonction $\phi()$, nous avons employé le cas polynomial multivarié dont un cas particulier est le suivant :

$$\phi([x_1 \;\; x_2 \;\; x_3]) = [1 \;\; x_1 \;\; x_2 \;\; x_3 \;\; x_1x_2 \;\; x_1x_3 \;\; x_2x_3 \;\; x_1x_2x_3] \tag{5.26}$$

La fonction $\phi()$ est caractérisée par un triplet $\{p, c, d\}$ où :

- p est le nombre de variables et correspond pratiquement au nombre d'eigenbackgrounds.
- c est le degré maximum sur une même variable.
- d est le degré maximum sur un monôme.

Le nombre de racines et le nombre de coefficients du polynôme sont des fonctions de $\{p, c, d\}$ [228]. Pratiquement, $p = 20$ et $c = d = 4$.

Les expérimentations (voir Section 5.3) montrent que cette approche est plus robuste que les méthodes PCP classiques dans la plupart des cas. Il reste donc comme dans le cas de la méthode RPCA-IRLS à développer une version incrémentale et temps réel.

5 Expérimentations et évaluations

Pour tester les trois contributions de ce chapitre, nous avons utilisé différentes bases de données tests comme Wallflower [521], ChangeDetection.net [204] et BMC 2012 [530]. Ces différentes bases de données vidéo ont permis de tester et d'évaluer les algorithmes qualitativement ou quantitativement dans différents contextes selon si les vérites terrain des séquences sont offertes ou pas, pour une partie ou pour toutes les séquences. Nous avons donc décidé, en premier temps, d'évaluer qualitativement les algorithmes. Cela a permis d'avoir une première idée sur le comportement des algorithmes proposés dans différentes situations critiques. Ensuite, pour les séquences fournies avec leurs vérités de terrain, nous avons évalué la méthode proposée quantitativement, en terme de faux positifs (FP) et de faux négatifs (FN) dans le domaine visible. Tous ces algorithmes ont été implémentés en utilisant Matlab. Les codes ont soit été fournis par les auteurs soit développés par Charles Guyon.

5.1 Evaluation des méthodes RPCA, suivi de sous-espaces et minimisation en rang faible sur BMC 2012

Nous avons évalué et comparé 19 algorithmes parmi ceux de l'ACP robustes, du suivi de sous-espaces et de la minimisation en rang faible comme suit :

- **ACP (1 algorithme) :** Nous avons choisi l'algorithme développé par Oliver et al. [387] comme référence en terme de performance.
- **ACP robustes via PCP (10 algorithmes) :** Pour cette catégorie, nous avons choisi les algorithmes les plus représentatifs qui sont les suivants : 1) PCP résolue par sept algorithmes différents (EALM [325], IALM [325], ADM (LRSD) [604], LADMAP [327], LSADM [199], LADM (LMaFit) [464], BLWS [328]), 2) SPCP résolue par NSA [30], 3) QPCP résolue par TFOCS [55] et 4) BPBP résolue par ALM [509].
- **Bayesian RPCA (2 algorithmes) :** Bayesian RPCA [156] et Varitional Bayesian RPCA [36].
- **Approximated RPCA (2 algorithmes) :** GoDec [643] et SemiSoft GoDec [643].
- **Suivi de sous-espaces (2 algorithmes) :** GRASTA [240] et pROST [230].
- **Minimisation en rang faible (2 algorithmes) :** DRMF [569] et PRMF [542].

Nous présentons les résultats visuels sur trois séquences de BMC : La séquence Rotary 122 (Fig. 5.2), la séquence "Wandering students" (Fig. 5.3) et la séquence "Traffic during windy day" (Fig. 5.4). Pour la séquence "Wandering students" (Figure 5.3), les variations lumineuses sont graduelles et on peut remarquer que les résultats restent visuellement corrects. De la figure 5.4, on peut constater que le modèle d'ACP robuste présente une bonne robustesse aux fonds dynamiques comme dans le cas du mouvement des arbres. Comme les résultats visuels sont assez proches, nous avons effectué une évaluation quantitative complète.

A partir des tableaux 5.18 et 5.20, on peut remarquer l'algorithme LSADM présente les meilleurs résultats pour les séquences 112, 212, 312 et 412. TFOCS semble émerger également en donnant les meilleurs résultats pour les séquences 512, 122, 222 et 522. De plus, TFOCS obtient la meilleure F-Mesure moyenne. A partir des tableaux 5.19 et 5.21 sur les séquences réelles, on peut voir que l'algorithme NSA est le plus robuste des algorithmes PCP pour les séquences 002, 003, 004 et 006. LSADM confirme sa robustesse en obtenant la seconde F-Mesure la plus élevée après l'algorithme BLWS. Les tableaux 5.26 et 5.27 regroupent les F-mesures moyennes pour les algorithmes les plus robustes.

Finalement, l'algorithme LSDAM donne la meilleure F-mesure sur les vidéos de synthèse. Pour les séquences réelles, le meilleur score est obtenu par l'algorithme TFOCS suivi par l'algorithme LSADM. Cependant, ces deux algorithmes ne sont ni incrémentaux ni temps réel. Une autre remarque est que la plupart des algorithmes RPCA apparaissent plus robustes que le modèle MOG adaptatif [466] tant sur les séquences de synthèse que sur les séquences réelles. Ce constat confirme donc le potentiel de ces méthodes pour la modélisation de fond et la détection d'objets mobiles.

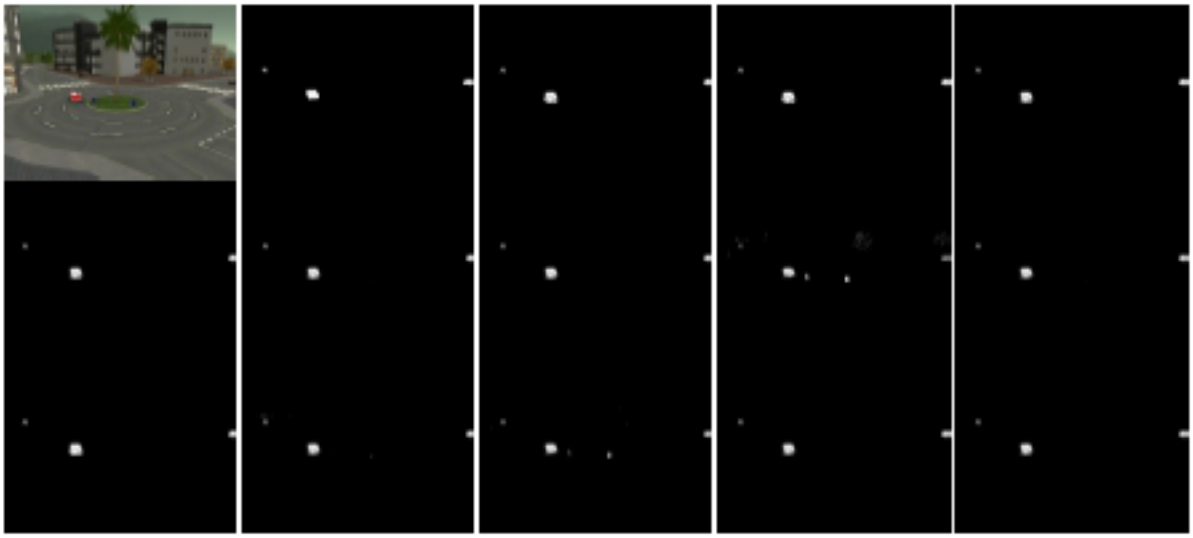


FIGURE 5.2 – Séquence Rotary 122 (Frame 1250) - Première ligne : Image originale, GT, EALM [325], IALM [325], ADM [604], Deuxième ligne : LADMAP [327], LSADM [199], LDAM [464], BLWS [328], NSA [30], TFOCS [55], Troisième ligne : BRPCA [156], VBRPCA [36], GoDec [643], SemiSoft GoDec [643].

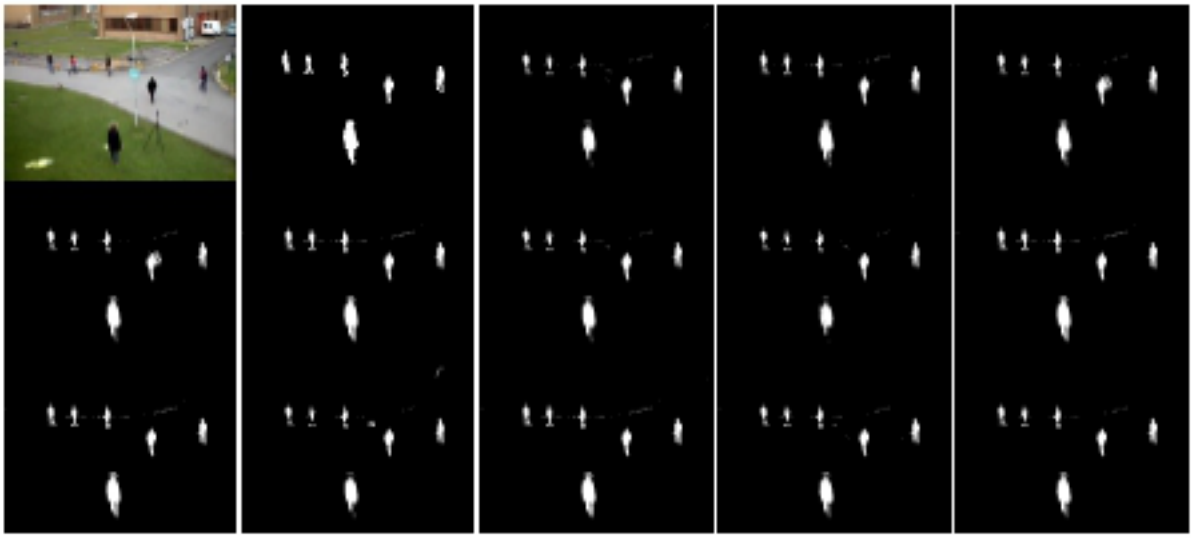


FIGURE 5.3 – Séquence Vidéo "Wandering student" (Frame 00000651)- Première ligne : Image originale, GT, EALM [325], IALM [325], ADM [604], Deuxième ligne : LADMAP [327], LSADM [199], LDAM [464], BLWS [328], NSA [30], TFOCS [55], Troisième ligne : BRPCA [156], VBRPCA [36], GoDec [643], SemiSoft GoDec [643].

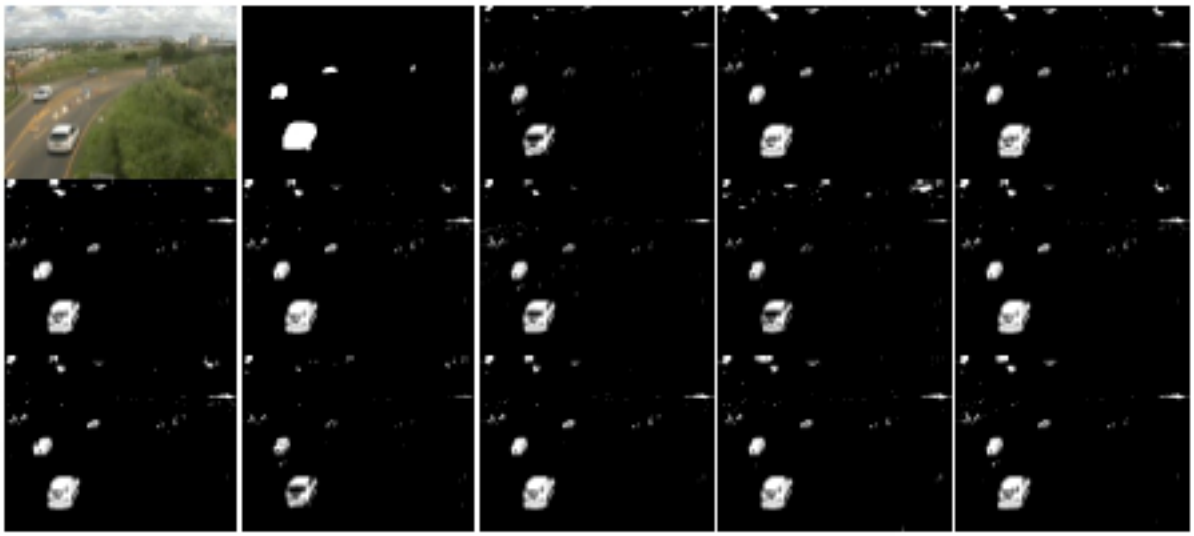


FIGURE 5.4 – Séquence Vidéo "Traffic during windy day" (Frame 00000140) - Première ligne : Image originale, GT, EALM [325], IALM [325], ADM [604], Deuxième ligne : LADMAP [327], LSADM [199], LDAM [464], BLWS [328], NSA [30], TFOCS [55], Troisième ligne : BRPCA [156], VBRPCA [36], GoDec [643], SemiSoft GoDec [643].

Méthodes	Algorithmes	Mesures	Street					Rotary					Moyenne
			112	212	312	412	512	122	222	322	422	522	
PCA	SVD	Recall	0.833	0.873	0.850	0.876	0.876	0.866	0.885	0.808	0.834	0.863	-
Oliver et al.[387]		Precision	0.960	0.944	0.966	0.945	0.898	0.895	0.910	0.946	0.775	0.869	-
		F-measure	0.892	0.907	0.805	0.809	0.787	0.880	0.897	0.871	0.803	0.866	0.851
RPCA-PCP	EALM	Recall	0.757	0.749	0.683	0.666	0.659	0.806	0.801	0.719	0.618	0.746	-
Candès et al. [92]	Lin et al. [325]	Precision	0.981	0.971	0.972	0.950	0.756	0.905	0.903	0.910	0.912	0.789	-
		F-measure	0.855	0.846	0.802	0.783	0.704	0.853	0.849	0.803	0.737	0.767	0.799
	IALM	Recall	0.924	0.839	0.891	0.888	0.827	0.893	0.900	0.891	0.855	0.855	-
	Lin et al. [325]	Precision	0.812	0.961	0.869	0.893	0.814	0.929	0.919	0.890	0.922	0.897	-
		F-measure	0.865	0.896	0.880	0.891	0.820	**0.911**	0.909	0.890	**0.887**	0.875	0.882
	ADM	Recall	0.849	0.841	0.863	0.887	0.830	0.877	0.878	0.861	0.808	0.868	-
	(LRSD)Yuan and Yang [604]	Precision	0.945	0.965	0.900	0.898	0.822	0.937	0.938	0.894	0.941	0.886	-
		F-measure	0.894	0.899	0.881	0.892	0.826	0.906	0.907	0.877	0.869	0.877	0.882
	LADMAP	Recall	0.849	0.841	0.874	0.888	0.831	0.877	0.880	0.863	0.805	0.868	-
	Lin et al. [327]	Precision	0.945	0.965	0.891	0.898	0.823	0.937	0.937	0.895	0.943	0.887	-
		F-measure	0.894	0.899	0.883	0.893	0.827	0.906	0.907	0.879	0.868	0.827	0.878
	LSADM	Recall	0.874	0.857	0.906	0.862	0.840	0.878	0.880	0.892	0.782	0.830	-
	Goldfarb et al. [199]	Precision	0.830	0.965	0.867	0.935	0.742	0.940	0.938	0.892	0.956	0.869	-
		F-measure	**0.912**	**0.908**	**0.886**	**0.897**	0.788	0.908	0.908	**0.892**	0.860	0.849	0.880
	LADM	Recall	0.829	0.803	0.777	0.720	0.740	0.875	0.862	0.750	0.747	0.792	-
	(LMaFit) Shen et al. [464]	Precision	0.979	0.984	0.963	0.966	0.724	0.944	0.953	0.902	0.910	0.832	-
		F-measure	0.898	0.884	0.860	0.825	0.732	0.908	0.906	0.819	0.821	0.812	0.846
	BLWS	Recall	0.700	0.786	0.697	0.785	0.664	0.663	0.633	0.603	0.620	0.793	-
	Lin and Wei [328]	Precision	0.981	0.847	0.971	0.848	0.966	0.921	0.560	0.538	0.775	0.711	-
		F-measure	0.817	0.816	0.812	0.815	0.787	0.771	0.594	0.569	0.689	0.750	0.742
RPCA-SPCP	NSA	Recall	0.857	0.856	0.872	0.850	0.838	0.867	0.871	0.892	0.786	0.837	-
Zhou et al. [647]	Aybat et al. [30]	Precision	0.969	0.966	0.898	0.946	0.747	0.949	0.945	0.891	0.948	0.861	-
		F-measure	0.910	0.907	0.885	0.896	0.790	0.906	0.906	0.891	0.860	0.849	0.880
RPCA-QPCP	TFOCS	Recall	0.910	0.843	0.867	0.903	0.834	0.898	0.892	0.892	0.831	0.877	-
Becker et al. [55]	Becker et al. [55]	Precision	0.830	0.965	0.899	0.889	0.824	0.924	0.932	0.887	0.940	0.879	-
		F-measure	0.868	0.900	0.882	0.896	**0.829**	**0.911**	**0.912**	0.889	0.882	**0.878**	**0.885**
RPCA-BPCP	ALM	Recall	0.757	0.806	0.749	0.801	0.683	0.719	0.666	0.618	0.659	0.746	-
Tang and Nehorai [509]	Tang and Nehorai [509]	Precision	0.981	0.905	0.971	0.903	0.972	0.910	0.950	0.912	0.756	0.789	-
		F-Measure	0.855	0.853	0.846	0.849	0.802	0.803	0.783	0.737	0.704	0.767	0.799
RPCA-IRLS	IRLS	Recall	-	-	-	-	-	-	-	-	-	-	-
Guyon et al. [226]	Guyon et al. [226]	Precision	-	-	-	-	-	-	-	-	-	-	-
		F-measure	-	-	-	-	-	-	-	-	-	-	-
Spatial RPCA-IRLS	IRLS	Recall	-	-	-	-	-	-	-	-	-	-	-
Guyon et al. [224]	Guyon et al. [224]	Precision	-	-	-	-	-	-	-	-	-	-	-
		F-measure	-	-	-	-	-	-	-	-	-	-	-
Spatial and Temporal RPCA-IRLS	IRLS	Recall	-	-	-	-	-	-	-	-	-	-	-
Guyon et al. [223]	Guyon et al. [223]	Precision	-	-	-	-	-	-	-	-	-	-	-
		F-measure	-	-	-	-	-	-	-	-	-	-	-
Polynomial RPCA	IRLS	Recall	0.944	0.903	0.851	0.896	0.864	0.879	0.855	0.824	0.801	0.856	-
Guyon et al. [228]	Guyon et al. [228]	Precision	0.841	0.916	0.968	0.921	0.888	0.892	0.895	0.879	0.723	0.850	-
		F-measure	0.890	0.910	0.906	0.908	0.876	0.886	0.875	0.879	0.760	0.853	0.874

TABLE 5.18 – Résultats sur les vidéos de synthèse pour la phase d'évaluation (BMC 2012 [530]) - Partie 1. Cette partie contient deux sous-parties. La première sous-partie du tableau concerne l'évaluation de méthodes RPCA existantes. La seconde sous-partie du tableau concerne l'évaluation des approches développées par Charles Guyon sous ma direction scientifique. Les expérimentations sont en cours pour les trois algorithmes de type IRLS.Dans les colonnes numériques, les scores les plus élevés sont indiqués en gras.

Méthode	Algorithmes	Mesure	Real Videos									Moyenne
			001	002	003	004	005	006	007	008	009	
PCA	SVD	Recall	0.818	0.672	0.890	0.824	0.799	0.793	0.580	0.753	0.861	-
Oliver et al. [387]		Precision	0.740	0.861	0.953	0.833	0.635	0.783	0.826	0.785	0.953	-
		F-measure	0.777	0.755	0.920	0.829	0.700	0.788	0.682	0.768	0.905	0.791
RPCA-PCP	EALM	Recall	0.723	0.765	0.878	0.839	0.660	0.800	0.739	0.670	0.748	-
Candès et al. [92]	Lin et al. [325]	Precision	0.554	0.855	0.958	0.844	0.543	0.729	0.899	0.728	0.965	-
		F-measure	0.628	**0.807**	0.916	0.842	0.596	0.763	0.811	0.698	0.843	0.697
	IALM	Recall	0.847	0.664	0.909	0.840	0.785	0.792	0.583	0.767	0.856	-
	Lin et al. [325]	Precision	0.735	0.873	0.947	0.827	0.632	0.793	0.833	0.781	0.957	-
		F-measure	0.787	0.754	0.928	0.833	0.701	0.792	0.686	0.774	0.904	0.794
	ADM	Recall	0.840	0.653	0.886	0.831	0.780	0.790	0.577	0.758	0.863	-
	(LRSD)Yuan and Yang [604]	Precision	0.751	0.876	0.946	0.820	0.633	0.782	0.828	0.781	0.952	-
		F-measure	**0.793**	0.749	0.915	0.825	0.699	0.786	0.680	0.769	**0.906**	0.791
	LADMAP	Recall	0.840	0.671	0.887	0.831	0.774	0.794	0.573	0.754	0.863	-
	Lin et al. [327]	Precision	0.751	0.865	0.945	0.818	0.634	0.782	0.840	0.791	0.952	-
		F-measure	**0.793**	0.756	0.915	0.824	0.697	0.788	0.681	0.772	**0.906**	0.792
	LSADM	Recall	0.842	0.685	0.933	0.870	0.793	0.805	0.599	0.770	0.850	-
	Goldfarb et al. [199]	Precision	0.661	0.873	0.952	0.879	0.624	0.805	0.842	0.783	0.959	-
		F-measure	0.741	0.768	**0.942**	0.875	0.699	**0.805**	0.700	0.776	0.901	0.800
	LADM	Recall	0.788	0.671	0.901	0.834	0.747	0.803	0.581	0.751	0.770	-
	(LMaFit) Shen et al. [464]	Precision	0.594	0.837	0.962	0.872	0.588	0.770	0.818	0.781	0.971	-
		F-measure	0.678	0.745	0.931	0.852	0.658	0.786	0.679	0.766	0.859	0.772
	BLWS	Recall	0.726	0.767	0.842	0.811	0.688	0.805	0.752	0.676	0.705	-
	Lin and Wei [328]	Precision	0.549	0.846	0.963	0.835	0.528	0.714	0.898	0.635	0.959	-
		F-measure	0.625	0.805	0.899	0.823	0.598	0.757	0.818	0.655	0.812	**0.839**
RPCA-SPCP	NSA	Recall	0.838	0.765	0.934	0.875	0.660	0.805	0.599	0.771	0.748	-
Zhou et al. [647]	Aybat et al. [30]	Precision	0.664	0.855	0.951	0.877	0.543	0.805	0.843	0.781	0.965	-
		F-measure	0.741	**0.807**	**0.942**	**0.876**	0.596	**0.805**	0.700	0.776	0.843	0.787
RPCA-QPCP	TFOCS	Recall	0.840	0.655	0.900	0.843	0.793	0.795	0.585	0.763	0.856	-
Becker et al. [55]	Becker et al. [55]	Precision	0.736	0.879	0.948	0.823	0.633	0.792	0.833	0.795	0.956	-
		F-measure	0.785	0.751	0.923	0.833	**0.704**	0.793	0.687	**0.779**	0.904	0.794
RPCA-BPCP	ALM	Recall	0.723	0.765	0.878	0.839	0.660	0.800	0.749	0.670	0.748	-
Tang and Nehorai [509]	Tang and Nehorai [509]	Precision	0.554	0.855	0.958	0.844	0.543	0.729	0.900	0.728	0.965	-
		F-measure	0.628	**0.807**	0.916	0.842	0.596	0.763	**0.817**	0.698	0.843	0.767
RPCA-IRLS	IRLS	Recall	-	-	-	-	-	-	-	-	-	-
Guyon et al. [226]	Guyon et al. [226]	Precision	-	-	-	-	-	-	-	-	-	-
		F-measure	-	-	-	-	-	-	-	-	-	-
Spatial RPCA-IRLS	IRLS	Recall	-	-	-	-	-	-	-	-	-	-
Guyon et al. [224]	Guyon et al. [224]	Precision	-	-	-	-	-	-	-	-	-	-
		F-measure	-	-	-	-	-	-	-	-	-	-
Spatial and Temporal RPCA-IRLS	IRLS	Recall	0.9139	0.8785	0.9658	0.9550	0.9102	0.9002	0.9116	0.8651	0.9309	-
Guyon et al. [223]	Guyon et al. [223]	Precision	0.7170	0.8656	0.8120	0.7187	0.5589	0.7727	0.8401	0.6710	0.8239	-
		F-measure	0.8036	0.8720	0.8822	0.8202	0.6925	0.8316	0.8744	0.7558	0.8741	0.822
Polynomial RPCA	IRLS	Recall	0.821	0.649	0.903	0.835	0.781	0.801	0.583	0.756	0.835	-
Guyon et al. [228]	Guyon et al. [228]	Precision	0.668	0.851	0.949	0.825	0.626	0.796	0.823	0.775	0.956	-
		F-measure	0.737	0.737	0.925	0.830	0.695	0.798	0.682	0.765	0.892	0.784

TABLE 5.19 – Résultats sur les vidéos réelles pour la phase d'évaluation (BMC 2012 [530]) - Partie 1. Cette partie contient deux sous-parties. La première sous-partie du tableau concerne l'évaluation de méthodes RPCA existantes. La seconde sous-partie du tableau concerne l'évaluation des approches développées par Charles Guyon sous ma direction scientifique. Les expérimentations sont en cours pour les trois algorithmes de type IRLS. Dans les colonnes numériques, les scores les plus élevés sont indiqués en gras.

Méthodes	Algorithmes	Mesures	Street					Rotary					Moyenne
			112	212	312	412	512	122	222	322	422	522	
Bayesian RPCA	Bayesian RPCA	Recall	0.809	0.874	0.776	0.666	0.658	0.806	0.661	0.679	0.624	0.746	-
	Ding et al. [156]	Precision	0.978	0.930	0.975	0.950	0.934	0.905	0.889	0.660	0.743	0.789	-
		F-measure	0.885	0.901	0.864	0.783	0.772	0.853	0.758	0.669	0.679	0.767	0.793
	Variational BRPCA	Recall	0.825	0.844	0.788	0.847	0.859	0.820	0.768	0.761	0.803	0.792	-
	Babacan et al. [36]	Precision	0.983	0.925	0.987	0.921	0.907	0.897	0.961	0.907	0.719	0.778	-
		F-measure	0.897	0.882	0.876	0.883	0.882	0.857	0.853	0.828	0.759	0.785	0.850
Approximated RPCA	GoDec	Recall	0.840	0.873	0.921	0.870	0.900	0.857	0.829	0.762	0.778	0.827	-
	Zhou and Tao [643]	Precision	0.966	0.941	0.874	0.944	0.852	0.897	0.865	0.957	0.686	0.831	-
		F-measure	0.899	0.906	0.897	0.906	0.875	0.877	0.846	0.848	0.729	0.829	0.861
	SemiSoft GoDec	Recall	0.841	0.875	0.850	0.868	0.866	0.822	0.879	0.792	0.813	0.866	-
	Zhou and Tao [643]	Precision	0.965	0.942	0.968	0.948	0.902	0.900	0.921	0.953	0.750	0.837	-
		F-measure	0.899	0.907	**0.905**	**0.906**	**0.884**	0.859	0.900	0.865	0.781	0.851	0.876
Subspace Tracking	GRASTA	Recall	0.700	0.787	0.695	0.787	0.669	0.680	0.637	0.619	0.623	0.791	-
	He et al. [240]	Precision	0.980	0.847	0.965	0.843	0.960	0.902	0.548	0.530	0.778	0.714	-
		F-measure	0.817	0.816	0.807	0.814	0.789	0.776	0.589	0.571	0.692	0.751	0.618
	pROST	Recall	0.944	0.878	0.853	0.889	0.785	0.819	0.838	0.789	0.738	0.863	-
	Hage and Kleinsteuber [230]	Precision	0.844	0.937	0.968	0.931	0.961	0.903	0.847	0.953	0.730	0.815	-
		F-measure	0.891	0.906	0.907	0.909	0.864	0.859	0.842	0.863	0.734	0.838	0.718
Low-Rank Minimization	DRMF	Recall	0.857	0.864	0.853	0.864	0.903	0.880	0.834	0.744	0.805	0.827	-
	Xiong et al. [569]	Precision	0.969	0.950	0.968	0.948	0.855	0.891	0.834	0.924	0.700	0.825	-
		F-measure	0.910	0.905	0.907	0.904	0.878	0.885	0.834	0.824	0.749	0.826	0.710
	PRMF	Recall	0.944	0.903	0.918	0.901	0.899	0.891	0.906	0.867	0.824	0.869	-
	Wang et al. [542]	Precision	0.819	0.919	0.879	0.922	0.862	0.887	0.849	0.888	0.707	0.821	-
		F-measure	0.877	0.911	0.898	0.911	0.880	0.889	0.877	0.878	0.761	0.845	0.727
Adaptive MOG		Recall	0.927	0.927	0.897	0.861	0.921	0.923	0.931	0.897	0.843	0.934	-
Shimada et al. [466]		Precision	0.866	0.868	0.580	0.526	0.619	0.886	0.890	0.626	0.535	0.840	-
		F-measure	0.896	0.896	0.705	0.653	0.740	0.904	0.910	0.738	0.655	**0.884**	0.798

TABLE 5.20 – Résultats sur les vidéos de synthèse pour la phase d'évaluation (BMC 2012 [530]) - Partie 2. Dans les colonnes numériques, les scores les plus élevés sont indiqués en gras.

Méthode	Algorithmes	Mesure	Real Videos									Moyenne
			001	002	003	004	005	006	007	008	009	
Bayesian RPCA	BRPCA	Recall	0.728	0.774	0.886	0.838	0.695	0.792	0.593	0.662	0.740	-
	Ding et al. [156]	Precision	0.553	0.857	0.949	0.824	0.536	0.732	0.839	0.732	0.966	-
		F-measure	0.629	**0.813**	0.917	0.831	0.605	0.761	0.695	0.695	0.839	0.753
	Variational BRPCA	Recall	0.834	0.690	0.935	0.875	0.781	0.817	0.606	0.755	0.847	-
	Babacan et al. [36]	Precision	0.621	0.862	0.949	0.874	0.614	0.809	0.843	0.781	0.950	-
		F-measure	0.712	0.766	**0.942**	0.874	0.687	**0.813**	0.705	0.768	0.896	0.795
Approximated RPCA	GoDec	Recall	0.833	0.701	0.910	0.858	0.771	0.820	0.615	0.747	0.850	-
	Zhou and Tao [643]	Precision	0.594	0.831	0.957	0.877	0.612	0.786	0.775	0.751	0.896	-
		F-measure	0.693	0.761	0.933	0.868	0.682	0.802	0.686	0.749	0.872	0.782
	SemiSoft GoDec	Recall	0.816	0.641	0.918	0.831	0.785	0.793	0.588	0.743	0.832	-
	Zhou and Tao [643]	Precision	0.697	0.855	0.959	0.843	0.638	0.781	0.792	0.778	0.965	-
		F-measure	0.752	0.733	0.938	0.837	**0.704**	0.787	0.675	0.760	0.894	0.787
Subspace Tracking	GRASTA	Recall	0.719	0.767	0.852	0.823	0.533	0.802	0.751	0.673	0.730	-
	He et al. [240]	Precision	0.542	0.845	0.963	0.796	0.516	0.711	0.900	0.696	0.950	-
		F-measure	0.618	0.804	0.904	0.809	0.524	0.754	0.819	0.684	0.826	0.674
	pROST	Recall	0.824	0.672	0.923	0.835	0.760	0.797	0.596	0.741	0.850	-
	Hage and Kleinsteuber [230]	Precision	0.632	0.844	0.958	0.884	0.631	0.787	0.796	0.770	0.874	-
		F-measure	0.715	0.749	0.940	0.859	0.689	0.792	0.682	0.755	0.862	0.704
Low-Rank Minimization	DRMF	Recall	0.828	0.719	0.934	0.874	0.772	0.823	0.617	0.762	0.842	-
	Xiong et al. [569]	Precision	0.600	0.856	0.949	0.877	0.618	0.801	0.811	0.766	0.929	-
		F-measure	0.696	0.782	0.941	0.875	0.686	0.812	0.701	0.764	0.883	0.714
	PRMF	Recall	0.830	0.720	0.937	0.874	0.790	0.805	0.596	0.746	0.839	-
		Precision	0.665	0.853	0.949	0.880	0.627	0.790	0.824	0.779	0.962	-
	Wang et al. [542]	F-measure	0.738	0.781	0.943	0.877	0.699	0.797	0.691	0.762	0.897	0.718
Adaptive MOG		Recall	0.949	0.680	0.959	0.929	0.854	0.880	0.791	0.823	0.928	-
Shimada et al. [466]		Precision	0.782	0.646	0.880	0.680	0.535	0.736	0.703	0.595	0.890	-
		F-measure	**0.857**	0.662	0.918	0.785	0.658	0.802	0.744	0.691	**0.909**	0.780

TABLE 5.21 – Résultats sur les vidéos réelles pour la phase d'évaluation (BMC 2012 [530]) - Partie 2. Dans les colonnes numériques, les scores les plus élevés sont indiqués en gras.

Methods	EALM [325]		IALM [325]		ADM (LRSD) [604]	
	SVDs	CPU	SVDs	CPU	SVDs	CPU
Sequence	550	00 :40 :15	38	00 :03 :47	510	00 :35 :20

TABLE 5.22 – Temps de calcul pour les algorithmes RPCA-PCP (Partie 1).

Methods	LADMAP [327]		LSADM [199]		LADM (LMaFit) [464]	
	SVDs	CPU	SVDs	CPU	SVDs	CPU
Sequence	16	00 :05 :20	43	00 :04 :03	35	00 :04 :55

TABLE 5.23 – Temps de calcul pour les algorithmes RPCA-PCP (Partie 2).

	RPCA-SPCP						Bayesian RPCA				Approximated RPCA			
Methods	ASALM [512]		NSA [30]		PSPG [31]		BRPCA [156]		VBRPCA [36]		GoDec [643]		Semi Soft GoDec [643]	
	SVDs	CPU	SVDs	CPU	SVDs	CPU	SVDs	CPU	SVDs	CPU	SVDs	CPU	SVDs	CPU
Sequence	94	00 :15 :17	19	00 :03 :07	23	00 :01 :05	-	00 :04 :01	-	00 :01 :07	-	00 :00 :50	-	00 :00 :55

TABLE 5.24 – Temps de calcul pour les algorithmes RPCA-SPCP, Bayesian RPCA et Approximated RPCA (Partie 3).

	Subspace Learning						Low Rank Minimization			
Methods	GRASTA [240]		pROST [230]		GOSUS [571]		DRMF [569]		PRMF [542]	
	SVDs	CPU	SVDs	CPU	SVDs	CPU	SVDs	CPU	SVDs	CPU
Sequence	-	00 :00 :30	-	00 :00 :19	-	00 :01 :50	-	00 :00 :17	-	00 :00 :42

TABLE 5.25 – Temps de calcul pour les algorithmes suivi de sous-espaces et de minimisation en rang faible (Partie 4).

	PCA [387]	RPCA-PCP [199]	RPCA-SPCP[647]
Synthetic videos	0.851	0.880	0.880
Real videos	0.791	**0.800**	0.787

TABLE 5.26 – F-Measure pour les vidéos de synthèse et réelles pour les méthodes les plus robustes (Partie 1).

	RPCA-QPCP[55]	RPCA-BPCP [509]	Bayesian RPCA [156]	Approximated RPCA [643]
Synthetic videos	**0.885**	0.799	0.850	0.876
Real videos	0.794	0.767	0.795	0.787

TABLE 5.27 – F-Measure pour les vidéos de synthèse et réelles pour les méthodes les plus robustes (Partie 2).

Les tableaux 5.22, 5.23, 5.24 et 5.25 montrent le nombre de SVDs et les temps de calcul de chacun des algorithmes. Le temps CPU est indiqué dans le format hh :mm :ss pour des séquences d'apprentissage de 200 images de taille 144×176 pour permettre la comparaison avec les publications antérieures qui présentent les performances dans ce format et sur la base de données I2R, notamment sur la séquence "Shopping Mall". L'algorithme EALM est le plus coûteux suivi par l'algorithme ADM (LRSD) et l'algorithme ASALM. Ces algorithmes dépassent les 15 minutes. Dans un deuxième groupe, les algorithmes IALM, LADMAP, LSADM, LADM (LMaFit), NSA, PSPG, BRPCA et VBRPCA donnent des temps de calcul entre 1 minute et 5 minutes. Les algorithmes de type GoDec et Semi-Soft GoDec semblent pouvoir atteindre des temps de calcul acceptables qui sont inférieurs à la minute. Les algorithmes de suivi de sous-espaces (GRASTA, pROST et GOSUS) présentent pour la plupart des temps de calcul inférieurs à la minute comme les algorithmes de minimisation en rang faible (DRMF et PRMF). Cependant, le temps réel à 25 fps n'est pas atteint. Il reste donc à faire des recherches pour développer des versions incrémentales et/ou temps-réel de ces algorithmes pour pouvoir les appliquer dans le domaine de la vidéo-surveillance.

5.2 Evaluation de la méthode RPCA-IRLS

Nous avons comparé la méthode proposée RPCA-IRLS [226] avec les méthodes suivantes : RPCA-BPCP [509], RPCA-PCP résolue via LADMAP [327], RPCA-PCP résolue via LSADM [199] et GRASTA [240]. Les expérimentations ont été effectuées sur les quatres bases de données que sont Wallflower [521], I2R [311], ChangeDetection.net [204] et BMC 2012 [530] comme suit :

— **Wallflower :** Les figures 5.5 et 5.6 montrent les résultats qualitatifs obtenus. Pour l'évaluation quantitative, nous avons utilisé les métriques basées sur les bonnes et fausses détections notées TP, TN, FP et FN. Ensuite, nous avons calculé les différentes métriques suivantes : le detection rate, la precision et la F-Measure. Un bon score est obtenu quand la F-Measure est proche de 1. Le tableau 5.28 montre les performances obtenues par chacun des algorithmes. La F-Measure ne peut pas être calculée pour la séquence "Moved Object" car il n'y a pas de détection positive dans la vérité-terrain. Les scores en gras correspondent au meilleur score pour la séquence de la ligne concernée.
— **I2R :** Pour cette base de données, les figures 5.7 et 5.8 montrent les résultats visuels sur six séquences et le tableau 5.29 montre les résultats quantitatifs correspondants.
— **ChangeDetection.net :** Pour chaque séquence, la vérité-terrain est fournie pour chaque images. Les résultats visuels sont montrés dans la figure 5.9. Le tableau 5.30 montre les résultats quantitatifs correspondants.
— **BMC 2012 :** Les tableaux 5.18 et 5.20 montrent les résultats quantitatifs obtenus respectivement sur les séquences de synthèses et les séquences réelles. Cette évaluation est en cours de réalisation.

FIGURE 5.5 – RPCA-IRLS : Résultats sur la base de données Wallflower [521]. De gauche à droite : Image originale, composante de rang faible (fond), composante parcimonieuse (objets mobiles).

La méthode RPCA-IRLS est plus robuste sur trois séquences de la base de données Wallflower, sur deux séquences de la base de données I2R, et de cinq séquences sur la base de données ChangeDetection.net. Globalement, la méthode développée fournit de meilleurs résultats sur dix séquences et elle est suivie par la méthode RPCA-PCP (LSADM) [199]. Sur le reste des séquences, les résultats sont acceptables. Pour les expérimentations, nous avons utilisé les paramétres indiqués par les auteurs pour leur méthode sur des décompositions effectuées sur 200 images consécutives.

Dataset	Sequence	Frame	RPCA-BPCP	RPCA-PCP (LADMAP)	RPCA-PCP (LSADM)	GRASTA	**RPCA-IRLS**
Wallflower [521]	Bootstrap	00299	**70.18**	69.40	67.02	55.37	58.78
	Foreground Aperture	00489	60.69	50.10	71.55	75.14	**75.19**
	Light Switch	01865	57.74	36.76	**69.33**	28.35	58.49
	Moved Objects	00985	-	-	-	-	-
	Time of Day	01850	71.43	54.41	**80.84**	79.80	80.43
	Waving Trees	00247	62.65	50.74	81.67	84.16	**84.70**
	Camouflage	00251	70.58	70.49	76.00	70.34	**82.27**

TABLE 5.28 – F-measure en pourcentage pour RPCA-BPCP [509],RPCA-PCP (LADMAP) [327], RPCA-PCP (LSADM) [199], GRASTA [240] et RPCA-IRLS [226]. RPCA-IRLS en gras est l'approche développée par Charles Guyon sous ma direction scientifique. Dans les colonnes numériques, les scores les plus élevés sont indiqués en gras.

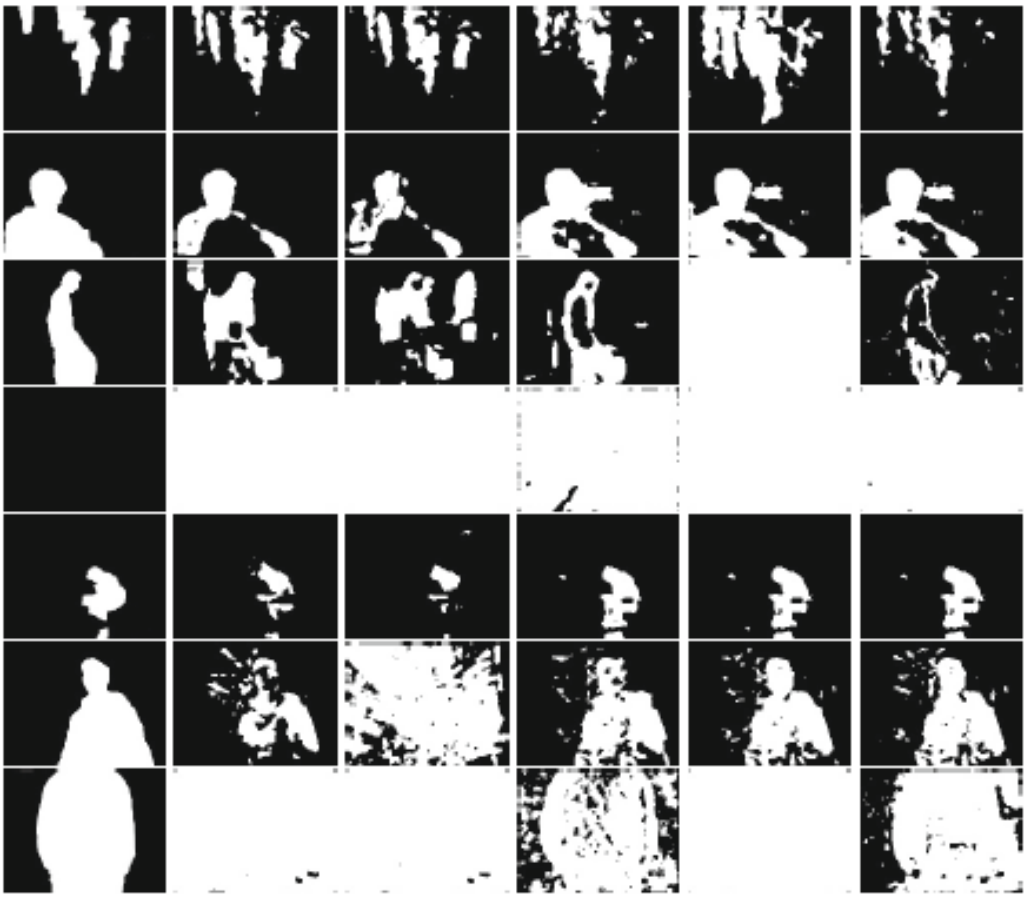


FIGURE 5.6 – RPCA-IRLS : Résultats sur la base de données Wallflower [521]. De gauche à droite : Vérité terrain, RPCA-BPCP [509], RPCA-PCP résolue via LADMAP [327], RPCA-PCP résolue via LSADM [199], GRASTA [240] et RPCA-IRLS [226].

FIGURE 5.7 – RPCA-IRLS : Résultats sur la base de données I2R. De gauche à droite : De gauche à droite : Image originale, composante de rang faible (fond), composante parcimonieuse (objets mobiles).

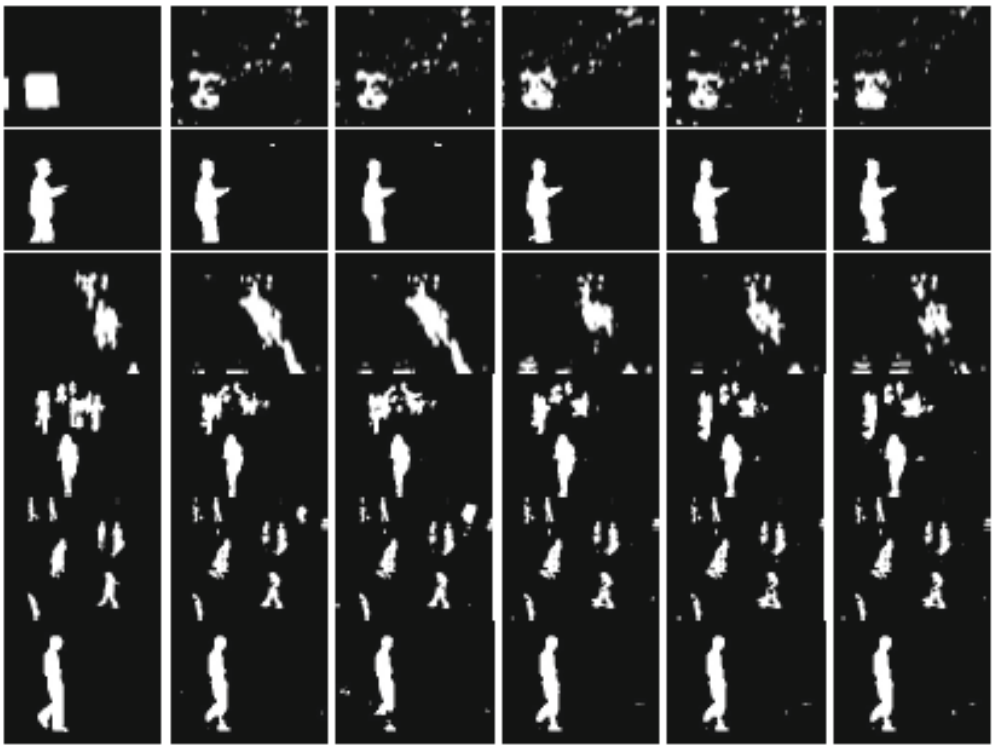


FIGURE 5.8 – RPCA-IRLS : Résultats sur la base de données Wallflower [521]. De gauche à droite : Vérité terrain, RPCA-BPCP [509], RPCA-PCP résolue via LADMAP [327], RPCA-PCP résolue via LSADM [199], GRASTA [240] et RPCA-IRLS [226].

Dataset	Sequence	Frame	RPCA-BPCP	RPCA-PCP (LADMAP)	RPCA-PCP (LSADM)	GRASTA	**RPCA-IRLS**
I2R	Campus	01650	59.93	59.99	68.14	61.84	**75.07**
	Curtain	22772	**91.08**	88.34	91.01	89.88	90.73
	Escalator	02424	65.28	63.41	59.81	**68.31**	63.68
	Hall	02926	73.58	69.75	77.57	78.25	**79.26**
	Shopping Mall	01862	80.15	77.88	**82.80**	79.54	81.94
	Water Surface	01499	90.37	83.57	**92.01**	90.97	91.02

TABLE 5.29 – F-measure en pourcentage RPCA-BPCP [509],RPCA-PCP (LADMAP) [327], RPCA-PCP (LSADM) [199], GRASTA [240] et RPCA-IRLS [226]. RPCA-IRLS en gras est l'approche développée par Charles Guyon sous ma direction scientifique. Dans les colonnes numériques, les scores les plus élevés sont indiqués en gras.

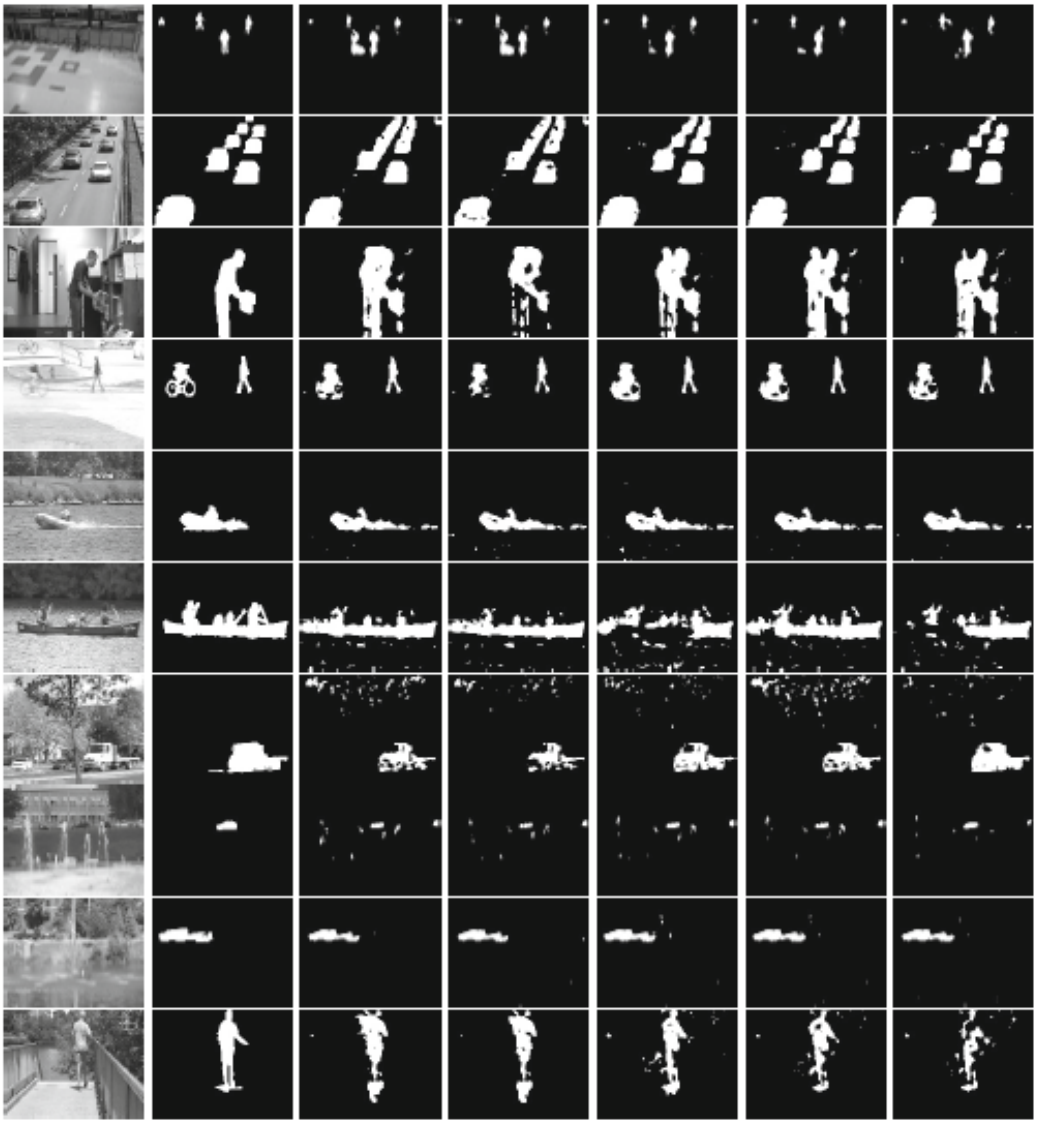


FIGURE 5.9 – RPCA-IRLS : Résultats sur la base de données ChangeDetection.net [204]. De gauche à droite : Vérité terrain, RPCA-BPCP [509], RPCA-PCP résolue via LADMAP [327], RPCA-PCP résolue via LSADM [199], GRASTA [240] et RPCA-IRLS [226].

Dataset	Sequence	Frame	RPCA-BPCP	RPCA-PCP (LADMAP)	RPCA-PCP (LSADM)	GRASTA	**RPCA-IRLS**
Baseline	PETS2006	00500	56.50	55.25	74.09	69.20	**78.12**
	Highway	00800	87.48	83.99	93.82	92.29	**94.44**
	Office	01886	76.01	58.42	75.67	78.36	**80.71**
	Pedestrians	00471	**82.74**	79.02	81.79	81.24	81.54
Dynamic Backgrounds	Boats	02000	84.01	82.29	83.36	**84.02**	83.39
	Canoe	00966	78.82	75.47	58.69	**79.73**	68.19
	Fall	01492	62.22	59.46	78.49	65.27	**88.39**
	Fountain01	00718	35.85	32.82	38.14	37.25	**54.46**
	Fountain02	00740	85.21	83.93	87.24	**87.25**	86.23
	Overpass	02529	72.01	72.55	**78.83**	78.65	70.33

TABLE 5.30 – RPCA-BPCP [509], RPCA-PCP (LADMAP) [327], RPCA-PCP (LSADM) [199], GRASTA [240] et RPCA-IRLS [226]. RPCA-IRLS en gras est l'approche développée par Charles Guyon sous ma direction scientifique. Dans les colonnes numériques, les scores les plus élevés sont indiqués en gras.

5.3 Evaluation de la méthode ACP polynomiale

Nous avons évalué la méthode ACP polynomiale proposée sur la base de données BMC 2012 [530]. Les tableaux 5.18 et 5.20 montrent les résultats quantitatifs obtenus respectivement sur les séquences de synthèses et les séquences réelles.

Sur les séquences de vidéos de synthèse, la méthode ACP polynomiale donne les meilleurs résultats sur les séquences 212, 312, 412 et 512 pour les vidéos de la catégorie "Street" en comparaison avec toutes les autres méthodes quelles soient de type RPCA, suivi de sous-espaces ou minimisation en rang faible. Pour la F-Measure, la méthode ACP polynomiale obtient un score meilleur que la version des MOG de Shimada et al. [466], la version RPCA de Candès et al. [92] et les méthodes de suivi de sous-espace et de minimization en rang faible. Cependant, quelques algorithmes de type RPCA résolus par LSADM ou TFOCS donnent de meilleurs résultats.

Sur les séquences de vidéos réelles, la méthode ACP polynomiale obtient un score meilleur que la version des MOG de Shimada et al. [466] et que les méthodes de suivi de sous-espace et de minimization en rang faible. Cependant, la plupart des autres méthodes RPCA obtiennent une meilleure F-Measure.

6 Conclusion

Dans ce chapitre, nous nous sommes focalisés sur l'apport des récentes avancées par ACP robustes dans la détection d'objets mobiles dans des séquences vidéo. Dans ce cadre, nous avons effectué les trois travaux de recherches suivants :

— **Un état de l'art :** pour une analyse systématique des méthodes par ACP robustes qui a permis de mettre en évidence le potentiel de ces méthodes en terme de robustesse en présence de fonds dynamiques et de variations lumineuses et leur limite actuelle en terme d'application temps-réel.

— **ACP robuste par "Iteratively Reweighted Least Squares" (IRLS)** : L'algorithme IRLS est plus robuste que les méthodes PCP classiques dans la plupart des cas. De plus, l'utilisation d'un terme de contrainte spatiale dans la minimisation renforce cette performance en présence de fonds dynamiques et de variations lumineuses. L'algorithme IRLS est donc une bonne alternative aux approches Lagrangiennes classiques. Il reste donc en à développer une version temps réel pour être utilisable dans une application comme la vidéo-surveillance.

— **ACP robuste polynomiale** : L'aspect polynomial permet de gérer les effets non linéaires qui apparaissent localement dans le voisinage des intensités des pixels. Ces effets non-linéaires proviennent des variations lumineuses soudaines, des variations des paramètres de la caméra (auto-ajustements) et du bruit lors de l'acquisition. Cette méthode polynomiale permet de plus de prendre en compte l'aspect additif de certaines variations et l'aspect multiplicatif d'autres variations.

Finalement, cette recherche a montré que ces récentes avancées en ACP robustes (RPCA), suivi de sous-espaces (Subspace Tracking) et minimisation en rang faible (Low Rank Minimization) sont donc fondamentales et offrent un potentiel important pour la modélisation de fond et la détection d'objets mobiles en vidéo-surveillance à l'image du mélange de Gaussiennes à ses débuts. Cependant, aucun algorithme n'émerge vraiment et ne permet de pallier simultanément les principaux challenges rencontrés dans les vidéos pour l'instant par manque d'investigations sur ces méthodes. De plus, des recherches doivent être menées pour rendre ces algorithmes à la fois incrémentaux et temps réels.

Chapitre 6

Conclusion Générale

Dans cette seconde partie, nous avons présenté les quatre grands axes des travaux que nous avons encadrés et effectués dans le cadre de la modélisation de fond et de la détection d'objets mobiles. Ce travail a permis de contribuer en partie à la structuration de ce domaine de recherche grâce à des états de l'art systématiques et d'ouvrir de nombreuses pistes de recherche pour poursuivre ce qui a été entrepris. Nous récapitulons dans cette conclusion notre contribution par les concepts flous, les approches de réduction de dimension discriminative et mixte, et les approches par ACP robustes ainsi que les perspectives que nous souhaitons mener dans ce domaine.

1 Synthèse des avancées proposées

1.1 Concepts flous

Tout d'abord, nous avons proposé une approche originale qui applique la théorie des ensembles flous de Type-2 à la modélisation de fond dynamique. Cette approche s'est montrée particulièrement adaptée à des fonds dynamiques (vibration de la caméra, mouvement d'arbres, écoulement d'eau dans le fond, ondulation de la surface d'eau). La méthode proposée est également robuste dans un environnement fermé (scène filmée en intérieur) face aux ombres et aux variations lumineuses. De plus, la méthode de classification floue des pixels en fond/objet qui utilisent la caractéristique couleur et une caractéristique texture (LBP) dont les résultats sont fusionnés par l'intégrale floue de Choquet permet d'obtenir méthode de détection plus robuste aux variations lumineuses. Finalement, la méthode floue de mise à jour adaptative basée sur le degré d'appartenance du pixel à la classe "fond"' et à la classe "objet mobile" renforce la robustesse de la méthode. Celle-ci peut être rendue plus adaptative en déterminant le coefficient d'apprentissage par une fonction plus évoluée des degrés d'appartenance.

Cette recherche a donc montré que les concepts flous permettent de prendre en compte l'imprécision et l'incertitude tant au niveau de la modélisation du fond qu'au niveau de la

classification fond/objet et à la mise à jour du fond. De plus, d'autres chercheurs ont repris ces travaux en les améliorant montrant ainsi la pertinence de ces approches floues.

1.2 Réduction de dimension discriminative et mixte

En premier, un état de l'art sur les approches de réduction de dimension a montré que seules les méthodes reconstructives avaient été utilisées jusqu'alors. Leur avantage principal est leur robustesse dans le cas de fonds comportant des variations lumineuses. Ensuite, nous avons présenté un modèle de fond par une approche discriminative de réduction de dimension adaptée à la classification nécessaire lors de la détection. Pour cela, nous avons utilisé l'approche discriminative incrémentale de réduction de dimension appelée "Incremental Maximum Margin Criterion". Finalement, nous avons développé un modèle de fond par une approche mixte de réduction de dimension combinant une approche reconstructive (ACP) avec une approche discriminative (LDA). Dans un premier temps, le fond est modélisé de façon robuste avec l'avantage d'une méthode non-supervisée et, dans un second temps, les pixels sont classés en "fond" ou "objet mobile" avec une méthode adaptée à la classification.

Ces travaux ont donc montré qu'une approche discriminative de réduction de dimension est plus performante pour l'étape de classification des pixels que les approches reconstructives. Cependant, une approche mixte parait plus adéquate car elle permet de combiner les avantages des approches reconstructives et discriminatives. Pour rendre plus applicable ces deux approches, il reste à développer des versions temps-réel de ces méthodes.

1.3 Décomposition en matrices de rang faible et parcimonieuse

Nous avons effectué un état de l'art le plus exhaustif possible pour une analyse systématique des approches basées sur la décomposition en une matrice de rang faible et une matrice parcimonieuse. Ensuite, nous avons proposé d'utiliser l'algorithme IRLS qui est plus robuste que les méthodes PCP classiques dans la plupart des cas. De plus, l'utilisation d'un terme de contrainte spatiale dans la minimisation renforce cette performance en présence de fonds dynamiques et de variations lumineuses. L'algorithme IRLS est donc une bonne alternative aux approches Lagrangiennes classiques. Il reste donc en à développer une version incrémentale et temps réel pour être utilisable dans une application comme la vidéo-surveillance.

Cette recherche a donc montré que les récentes avancées en ACP robustes (RPCA), suivi de sous-espaces (Subspace Tracking) et minimisation en rang faible (Low Rank Minimization) ont donc un potentiel important pour la modélisation de fond et la détection d'objets mobiles en vidéo-surveillance à l'image du mélange de Gaussiennes à ses débuts. Cependant, des recherches doivent être menées pour rendre ces algorithmes à la fois incrémentaux et temps réels.

2 A la poursuite de la méthode idéale

Outre les améliorations des méthodes précédentes, les perspectives de travail sont nombreuses car il reste encore un certain nombre de questions ouvertes qui concernent les étapes suivantes :

2.1 L'initialisation du fond

Contrairement à la modélisation de fond et la mise à jour du fond, relativement peu d'articles portent sur cette étape. Cependant, dans le cas de séquence courte, la qualité de la première image de fond conditionne la qualité des premières détections qui peuvent elle-même conditionner les suivantes selon le processus de mise à jour de l'image. S'attacher à cette étape présente donc un intérêt majeur et se présente ainsi comme une perspective de recherche.

2.2 La modélisation de fond

Bien qu'il y ait un grand nombre de modèles existants, il est encore nécessaire de développer des modèles plus robustes. Par exemple dans la continuité de nos travaux, l'idée est d'utiliser une décomposition tensorielle pour développer une approche spatio-temporelle robuste car la décomposition en matrices par ACP robustes et l'exploitation qui en est faite ne permet pas de prendre en compte les contraintes spatiales et temporelles des objets mobiles de façon optimale.

2.3 La sélection de caractéristiques

La couleur dans l'espace RGB est la caractéristique la plus employée car elle est directement fournie par les caméras. Bien qu'elle soit souvent discriminative pour détecter les objets, elle présente plusieurs limitations en présence de certains challenges comme les variations lumineuses, le camouflage et les ombres. Pour résoudre cela, les caractéristiques comme le contour ou la texture sont employées en plus de la couleur. Chacune d'entre elles ayant une robustesse spécifique à tel on tel challenge car invariantes à tel ou tel changement. A l'heure actuelle, très peu de travaux se sont intéressés vraiment à une étude approfondie de l'influence de l'utilisation de telle ou telle caractéristique et cela est donc à l'heure actuelle un problème ouvert. Il parait donc pertinent aussi de s'intéresser dans le cadre de futurs travaux à une méthode de sélection de caractéristiques permettant d'employer les caractéristiques les plus adaptées pour une zone de l'image où apparait tel ou tel challenge.

3 Conclusion

En conclusion, le constat qui peut être fait à la vue de ces travaux, et des résultats de ChangeDetection.net et BMC 2012, est que même si certaines méthodes sont globalement performantes et que d'énormes progrès ont été effectués pour être robuste à certains challenges, aucune méthode n'est performante dans tous les challenges.

La solution est donc sûrement dans une méthode multi-modèle qui, une fois le challenge identifié, sélectionnerait automatiquement le modèle le plus robuste dans une banque de modèles caractérisés par leur robustesse à tel ou tel challenge. Une autre voie consisterait aussi à choisir les caractéristiques les plus appropriées dans la situation identifiée comme indiqué dans le paragraphe précédent.

Au niveau des capteurs, il serait aussi intéressant d'imaginer un système multi-capteurs (caméra couleur, caméra IR et caméra RGB-D). L'idée serait soit de sélectionner tel ou tel capteur suivant les conditions détectées ou soit de fusionner les informations venant de ceux-ci. On peut aussi imaginer un système connecté GPS qui recevrait des informations sur les conditions climatiques pour adapter ou sélectionner le modèle le plus adéquat à la situation environnementale.

Finalement, une autre possibilité serait aussi de pouvoir entrer comme paramètre une information contextuelle comme par exemple la forme ou la couleur des objets rencontrés dans l'application concernée. Pratiquement, cela reviendrait à mettre des *a priori* de forme ou de couleur.

Bibliographie

[1] T. Aach and A. Condurache. Transformation of adaptive thresholds by significance invariance for change detection. *International Workshop on Statistical Signal Processing*, July 2005.

[2] T. Aach, L. Dumbgen, R. Mester, and D. Toth. Bayesian illumination-invariant motion detection. *IEEE International Conference on Image Processing, ICIP 2001*, 3 :640–643, October 2001.

[3] T. Aach and A. Kaup. Bayesian algorithms for adaptive change detection in image sequences using Markov Random Fields. *Signal Processing Image Communication,*, 7 :147–160, 1995.

[4] T. Aach, A. Kaup, and R. Mester. Change detection in image sequences using Gibbs random fields : a Bayesian approach. *IEEE Workshop Intelligent Signal Processing and Communications Systems*, October 1993.

[5] T. Aach, A. Kaup, and R. Mester. Statistical model-based change detection in moving video. *Signal Processing*, pages 165–180, 1993.

[6] T. Aach, D. Toth, and R. Mester. Motion estimation in varying illumination using a total least squares distance measure. *Picture Coding Symposium, PCS 2003*, pages 145–148, April 2003.

[7] R. Abbott and L. Williams. Multiple target tracking with lazy background subtraction and connected component analysis. *Machine Vision and Applications*, 20(2) :93–101, January 2009.

[8] A. Abdel-Hakim and M. El-Saban. FRPCA : fast robust principal component analysis. *International Conference on Pattern Recognition, ICPR 2012*, November 2012.

[9] J. Ahn. Fast adaptive robust subspace tracking for online background subtraction. *International Conference on Pattern Recognition, ICPR 2014*, 2014.

[10] C. Aicher. A variational bayes approach to robust principal component analysis. *REU 2013*, 2013.

[11] C. Aicher. A variational Bayes approach to robust principal component analysis. *REU 2013*, 2013.

[12] N. Al-Najdawi, H. Bez, J. Singhai, and E. Edirisinghe. A survey of cast shadow detection algorithms. *Pattern Recognition Letters*, 33(6) :752–764, April 2012.

[13] A.Leykin and M. Tuceryan. Detecting shopper groups in video sequences. *International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2007*, 2007.

[14] I. Ali, J. Mille, and L. Tougne. Space-time spectral model for object detection in dynamic textured background. *International Conference on Advanced Video and Signal Based Surveillance, AVSS 2010*, 33(13) :1710–1716, October 2012.

[15] I. Ali, J. Mille, and L. Tougne. Adding a rigid motion model to foreground detection : Application to moving object detection in rivers. *Pattern Analysis and Applications*, July 2013.

[16] M. Allili, N. Bouguila, and D. Ziou. A robust video foreground segmentation by using generalized Gaussian mixture modeling. *Canadian Conference on Computer and Robot Vision, CRV 2007*, pages 503–509, 2007.

[17] A. Amato, I. Huerta, M. Mozerov, F. Roca, and J. Gonzalez. Moving cast shadows detection methods for video surveillance applications. *Augmented Vision and Reality*, September 2012.

[18] A. Amato, M. Mozerov, F. Roca, and J. Gonzalez. Robust real-time background subtraction based on local neighborhood patterns. *EURASIP Journal on Advances in Signal Processing*, 2010, 2010.

[19] Z. An. Video background modeling based on optimization algorithms of robust pca. *Thesis*, February 2014.

[20] M. Anderson, G. Ballard, J. Demme, and K. Keutzer. Communication-avoiding QR decomposition for GPUs. *Technical Report, ECCS,*, 2010.

[21] W. Hu and·X. Li, X. Zhang, X. Shi, S. Maybank, and Z. Zhang. Incremental tensor subspace learning and its applications to foreground segmentation and tracking. *International Journal of Computer Vision*, 91(3) :303–327, 2011.

[22] B. Antic, D. Letic, V. Crnojevic, and D. Culibrk. Performance analysis of algorithms for video background subtraction. *National Conference of Digital Speech and Image Processing*, 2008.

[23] S. Apewokin, B. Valentine, L. Wills, S. Wills, and A. Gentile. Multimodal mean adaptive backgrounding for embedded real-time video surveillance. *Embedded Computer Vision Workshop, ECVW 2007*, June 2007.

[24] K. Appiah, A. Hunter, P. Dickinson, and H. Meng. Accelerated hardware video object segmentation : from foreground detection to connected components labelling. *Computer Vision and Image Understanding, CVIU 2010*, 2010.

[25] A. Aravkin, S. Becker, V. Cevher, and P. Olsen. A variational approach to stable principal component pursuit. *Preprint*, June 2014.

[26] M. Arif, S. Hafid, T.Brouard, and N. Vincent. AWFO : an aggregation operation for pattern recognition. *SPIE*, 1998.

[27] C. Arth, H. Bischof, and C. Leistner. Tricam-an embedded platform for remote traffic surveillance. *CVPR Workshop on Embedded Computer Vision, CVPRW 2006*, June 2010.

[28] R. Athilingam, K. Kumar, and G. Kavitha. Neuronal mapped hybrid background segmentation for video object tracking. *International Conference on Computing, Electronics and Electrical Technologies, ICCEET 2012*, pages 1061–1066, 2012.

[29] N. Avinash, M. Shashi Kumar, and S. Sagar. Automated video surveillance for retail store statistics generation. *International Conference on Signal and Image Processing, ICSIP 2012*, pages 585–596, 2012.

[30] N. Aybat, D. Goldfarb, and G. Iyengar. Fast first-order methods for stable principal component pursuit. *Preprint*, 2011.

[31] N. Aybat, D. Goldfarb, and G. Iyengar. Efficient algorithms for robust and stable principal component pursuit. *Preprint*, November 2012.

[32] N. Aybat and G. Iyengar. An alternating direction method with increasing penalty for stable principal component pursuit. *Computational Optimization and Applications*, 2014.

[33] M. Azab, H. Shedeed, and A. Hussein. A new technique for background modeling and subtraction for motion detection in real-time videos. *International Conference on Image Processing, ICIP 2010*, pages 3453–3456, September 2010.

[34] P. Azzari, L. Di Stefano, and A. Bevilacqua. An effective real-time mosaicing algorithm apt to detect motion through background subtraction using a PTZ camera. *IEEE Conference on Advanced Video and Signal Based Surveillance, AVSS 2005*, pages 511–516, September 2005.

[35] S. Babacan, M. Luessi, R. Molina, and A. Katsaggelos. Sparse additive matrix factorization for Robust PCA and its generalization. *ACML 2012*, 2012.

[36] S. Babacan, M. Luessi, R. Molina, and A. Katsaggelos. Sparse Bayesian methods for low-rank matrix estimation. *IEEE Transactions on Signal Processing*, 60(8) :3964–3977, 2012.

[37] F. El Baf and T. Bouwmans. Comparison of background subtraction methods for a multimedia learning space. *International Conference on Signal Processing and Multimedia, SIGMAP*, July 2007.

[38] F. El Baf, T. Bouwmans, and B. Vachon. Background Subtraction : State of the Art and Perspectives. *Journées PRIDES*, June 2006.

[39] F. El Baf, T. Bouwmans, and B. Vachon. Comparison of background subtraction methods for a multimedia application. *International Conference on systems, Signals and Image Processing, IWSSIP 2007*, pages 385–388, June 2007.

[40] F. El Baf, T. Bouwmans, and B. Vachon. Comparison of background subtraction methods for a multimedia learning space. *International Conference on Signal Processing and Multimedia, SIGMAP 2007*, July 2007.

[41] F. El Baf, T. Bouwmans, and B. Vachon. Detection d'objets en mouvement par l'integrale de Choquet. *Journees PRIDES*, July 2008.

[42] F. El Baf, T. Bouwmans, and B. Vachon. Foreground detection using the Choquet integral. *IEEE International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2008*, pages 187–190, May 2008.

[43] F. El Baf, T. Bouwmans, and B. Vachon. A fuzzy approach for background subtraction. *IEEE International Conference on Image Processing, ICIP 2008*, pages 2648–2651, October 2008.

[44] F. El Baf, T. Bouwmans, and B. Vachon. Fuzzy foreground detection for infrared videos. *IEEE International Conference on Computer Vision and Pattern Recognition, CVPR-Workshop OTCBVS 2008*, pages 1–6, June 2008.

[45] F. El Baf, T. Bouwmans, and B. Vachon. Fuzzy integral for moving object detection. *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2008*, pages 1729–1736, June 2008.

[46] F. El Baf, T. Bouwmans, and B. Vachon. Type-2 fuzzy mixture of Gaussians model : Application to background modeling. *International Symposium on Visual Computing, ISVC 2008*, pages 772–781, December 2008.

[47] F. El Baf, T. Bouwmans, and B. Vachon. Fuzzy statistical modeling of dynamic backgrounds for moving object detection in infrared videos. *IEEE International Conference on Computer Vision and Pattern Recognition, CVPR-Workshop OTCBVS 2009*, pages 60–65, June 2009.

[48] M. Balcilar and A. Sonmez. Region based fuzzy background subtraction using Choquet integral. *International Conference on Adaptive and Natural Computing Algorithms, ICANNGA 2013*, pages 287–296, April 2013.

[49] D. Baltieri, R. Cucchiara, and R. Vezzani. Fast background initialization with recursive hadamard transform. *International Conference on Advanced Video and Signal Based Surveillance, AVSS 2010*, September 2010.

[50] B. Bao, G. Liu, C. Xu, and S. Yan. Inductive robust principal component analysis. *IEEE Transactions on Image Processing*, pages 3794–3800, August 2012.

[51] O. Barnich and M. Van Droogenbroeck. ViBe : a powerful random technique to estimate the background in video sequences. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2009*, pages 945–948, April 2009.

[52] A. Bayona, J. San Miguel, and J. Martinez. Comparative evaluation of stationary foreground object detection algorithms based on background subtraction techniques. *AVSS 2009*, September 2009.

[53] A. Bayona, J. SanMiguel, and J. Martinez. Comparative evaluation of stationary foreground object detection algorithms based on background subtraction techniques. *International Conference on Advanced Video and Signal Based Surveillance, AVSS 2009*, pages 25–30, September 2009.

[54] M. Beal. Variational algorithms for approximate bayesian inference. *PhD Thesis, University of London*, 2003.

[55] S. Becker, E. Candes, and M. Grant. TFOCS : flexible first-order methods for rank minimization. *Low-rank Matrix Optimization Symposium, SIAM Conference on Optimization*, 2011.

[56] A. Bell and T. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(11) :1129–1159, 1995.

[57] C. Benedek and T. Sziranyi. Markovian framework for foreground-background-shadow segmentation of real world video scenes. *Asian Conference on Computer Vision, ACCV 2006*, January 2006.

[58] Y. Benezeth, P. Jodoin, B. Emile, H. Laurent, and C. Rosenberger. Review and evaluation of commonly-implemented background subtraction algorithms. *IEEE International Conference on Pattern Recognition, ICPR 2008*, December 2008.

[59] Y. Benezeth, P. Jodoin, B. Emile, H. Laurent, and C. Rosenberger. Comparative study of background subtraction algorithms. *Journal of Electronic Imaging*, 19, 2010.

[60] A. Bevilacqua, G. Capelli, L. Di Stefano, and A. Lanza. A novel approach to change detection based on a coarse-to-fine strategy. *IEEE International Conference on Image Processing, ICIP 2005*, pages 434–437, 2005.

[61] A. Bevilacqua, L. Di Stefano, and A. Lanza. Coarse-to-fine strategy for robust and efficient change detectors. *IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2005*, September 2005.

[62] S. Bhandarkar and X. Luo. Fast and robust background updating for real-time surveillance systems. *Workshop on Machine Vision for Intelligent Vehicles in conjunction with CVPR 2005*, pages 1–6, June 2005.

[63] H. Bhaskar, L. Mihaylova, and A. Achim. Video foreground detection based on symmetric alpha-stable mixture models. *IEEE Transactions on Circuits, Systems and Video Technology*, March 2010.

[64] X. Bian and H. Krim. Robust subspace recovery via bi-sparsity pursuit. *Preprint*, April 2014.

[65] P. Blauensteiner and M. Kampel. Visual surveillance of an airport's apron - an overview of the AVITRACK project. *Workshop of the Austrian Association for Pattern Recognition, AAPR 2004*, pages 213–220, 2004.

[66] B.Lee and M. Hedley. Background estimation for video surveillance. *Image and Vision Computing New Zealand, IVCNZ*, pages 315–320, 2002.

[67] D. Bloisi, A. Pennisi, and L. Iocchi. Background modeling in the maritime domain. *Machine Vision and Applications*, 2014.

[68] S. Boragno, B. Boghossian, J. Black, D. Makris, and S. Velastin. A DSP-based system for the detection of vehicles parked in prohibited areas. *AVSS 2007*, pages 1–6, 2007.

[69] T. Bouwmans. Subspace Learning for Background Modeling : A Survey. *Recent Patents on Computer Science, RPCS 2009*, 2(3) :223–234, November 2009.

[70] T. Bouwmans. Recent Advanced Statistical Background Modeling for Foreground Detection : A Systematic Survey. *Recent Patents on Computer Science, RPCS 2011*, 4(3) :147–176, September 2011.

[71] T. Bouwmans. Background Subtraction For Visual Surveillance : A Fuzzy Approach. *Chapter 5, Handbook on Soft Computing for Video Surveillance, Taylor and Francis Group, S.K. Pal, A. Petrosino, L. Maddalena*, pages 103–139, March 2012.

[72] T. Bouwmans. Recent Approaches in Background Modeling for Video Surveillance. *Handbook Background Modeling and Foreground Detection for Video Surveillance, Taylor and Francis Group, T. Bouwmans, B. Hoferlin, F. Porikli, A. Vacavant*, June 2014.

[73] T. Bouwmans. Traditional Approaches in Background Modeling for Video Surveillance. *Handbook Background Modeling and Foreground Detection for Video Surveillance, Taylor and Francis Group, T. Bouwmans, B. Hoferlin, F. Porikli, A. Vacavant*, June 2014.

[74] T. Bouwmans and F. El Baf. Modeling of Dynamic Backgrounds by Type-2 Fuzzy Gaussians Mixture Models. *MASAUM Journal of Basic and Applied Sciences*, 1(2) :265–277, November 2009.

[75] T. Bouwmans, F. El Baf, and B. Vachon. Background Modeling using Mixture of Gaussians for Foreground Detection - A Survey. *Recent Patents on Computer Science, RPCS 2008*, 1(3) :219–237, November 2008.

[76] T. Bouwmans, F. El Baf, and B. Vachon. Statistical Background Modeling for Foreground Detection : A Survey. *Part 2, Chapter 3, Handbook of Pattern Recognition and Computer Vision, World Scientific Publishing, Prof C.H. Chen*, 4 :181–199, January 2010.

[77] T. Bouwmans, F. El-Baf, and B. Vachon. Background modeling using mixture of Gaussians for foreground detection : A survey. *Recent Patents on Computer Science*, 1(3) :219–237, November 2008.

[78] T. Bouwmans, A. Sobral, and E. Zahzah. Foreground detection via decomposition in low-rank and spase matrices : A review for a comparative evaluation with large scale datasets. *En rédaction*, 2014.

[79] T. Bouwmans and E. Zahzah. Robust PCA via principal component pursuit : A review for a comparative evaluation in video surveillance. *Special Isssue on Background Models Challenge, Computer Vision and Image Understanding, CVIU 2014*, 122 :22–34, May 2014.

[80] R. Bowen. Implementation, integration, and optimization of a fuzzy foreground segmentation system. *Master Thesis, Rochester Institute of Technology Rochester, New York*, (1), February 2013.

[81] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distribute optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1) :1–123, 2011.

[82] S. Brutzer, B. Höferlin, and G. Heidemann. Evaluation of background subtraction techniques for video surveillance. *International Conference on Computer Vision and Pattern Recognition, CVPR 2011*, pages 1937–1944, June 2011.

[83] S. Bucak, B. Gunsel, and O. Gursoy. Incremental non-negative matrix factorization for dynamic background modeling. *International Workshop on Pattern Recognition in Information Systems*, June 2007.

[84] D. Butler, V. Bove, and S. Shridharan. Real time adaptive foreground/background segmentation. *EURASIP*, pages 2292–2304, 2005.

[85] R. Cabral, F. De la Torre, J. Costeira, and A. Bernardino. Unifying nuclear norm and bilinear factorization approaches for low-rank matrix decomposition. *International Conference on Computer Vision, ICCV 2013*, December 2013.

[86] J. Cai, E. Candes, and Z. Shen. A singular value thresholding algorithm for matrix completion. *International Journal of ACM*, May 2008.

[87] E. Calvo-Gallego, P. Brox, and S. Sanchez-Solano. A fuzzy system for background modeling in video sequences. *WILF 2013*, pages 184–192, 2013.

[88] J. Campbell, L. Mummert, and R. Sukthankar. Video monitoring of honey bee colonies at the hive entrance. *ICPR Workshop on Visual Observation and Analysis of Animal and Insect Behavior, VAIB 2008*, December 2008.

[89] M. Camplani, C. Blanco, L. Salgado, F. Jaureguizar, and N. Garcia. Advanced background modeling with RGB-D sensors through classifiers combination and inter-frame foreground prediction. *Machine Vision and Application, Special Issue on Background Modeling*, 2013.

[90] M. Camplani and L. Salgado. Background foreground segmentation with RGB-D Kinect data : an efficient combination of classifiers. *Journal on Visual Communication and Image Representation*, 2013.

[91] E. Candes. Compressive sampling. *International Congress of Mathematicians*, 1998.

[92] E. Candes, X. Li, Y. Ma, and J. Wright. Robust principal component analysis ? *International Journal of ACM*, 58(3), May 2011.

[93] L. Carminati and J. Benois-Pineau. Gaussian mixture classification for moving object detection in video surveillance environment. *International Conference on Image Processing, ICIP 2005*, pages 113–116, 2005.

[94] M. Casares and S. Velipasalar. Resource-efficient salient foreground detection for embedded smart cameras by tracking feedback. *International Conference on Advanced Video and Signal Based Surveillance, AVSS 2010*, September 2010.

[95] M. Casares, S. Velipasalar, and A. Pinto. Light-weight salient foreground detection for embedded smart cameras. *Computer Vision and Image Understanding*, 2010.

[96] V. Cevher, M. Duarte, C. Hedge, and R. Baraniuk. Sparse signal recovery using markov random fields. *NIPS 2008*, 2008.

[97] V. Cevher, D. Reddy, M. Duarte, A. Sankaranarayanan, R. Chellappa, and R. Baraniuk. Background subtraction for compressed sensing camera. *European Conference on Computer Vision, ECCV 2008*, October 2008.

[98] M. Chacon-Muguia and S. Gonzalez-Duarte. An adaptive neural-fuzzy approach for object detection in dynamic backgrounds for surveillance systems. *IEEE Transactions on Industrial Electronics*, 2011.

[99] M. Chacon-Murguia and S. Gonzalez-Duarte. An adaptive neural-fuzzy approach for object detection in dynamic backgrounds for surveillance systems. *IEEE Transactions on Industrial Electronic*, pages 3286–3298, August 2012.

[100] M. Chacon-Murguia, G. Ramirez-Alonso, and S. Gonzalez-Duarte. Improvement of a neural-fuzzy motion detection vision model for complex scenario conditions. *International Joint Conference on Neural Networks, IJCNN 2013*, August 2013.

[101] M. Chacon-Murguia and J. Urias-Zavala. A DTCNN approach on video analysis : Dynamic and static object segmentation. *Recent Advances on Hybrid Approaches for Designing Intelligent Systems, Studies in Computational Intelligence*, pages 315–336, 2014.

[102] A. Chan, V. Mahadevan, and N. Vasconcelos. Generalized Stauffer-Grimson background subtraction for dynamic scenes. *Machine Vision and Applications*, 21(5) :751–766, September 2011.

[103] T. Chang, T. Ghandi, and M. Trivedi. Vision modules for a multi sensory bridge monitoring approach. *ITSC 2004*, pages 971–976, 2004.

[104] R. Chartrand. Non convex splitting for regularized low-rank and sparse decomposition. *IEEE Transactions on Signal Processing*, 2012.

[105] B. Chen, W. Wang, and Q. Qin. Infrared target detection based on fuzzy ART neural network. *International Conference on Computational Intelligence and Natural Computing*, 2 :240–243, September 2010.

[106] B. Chen, W. Wang, and Q. Qin. Infrared target detection based on spatially related fuzzy ART neural network. *CSEE 2011*, pages 55–61, 2011.

[107] C. Chen and J. Aggarwal. An adaptive background model initialization algorithm with objects moving at different depths. *International Conference on Image Processing, ICIP 2008*, pages 2264–2267, 2008.

[108] C. Chen, J. Cai, W. Lin, and G. Shi. Surveillance video coding via low-rank and sparse decomposition. *ACM international conference on Multimedia*, pages 713–716, 2012.

[109] D. Chen and L. Zhang. An incremental linear discriminant analysis using fixed point method. *ISSN 2006*, 3971 :1334–1339, 2006.

[110] S. Chen, D. Donoho, and M. Saunders. Atomic decomposition by basis pursuit. *Journal on Scientific Computing*, 1998.

[111] Y. Chen, C. Chen, C. Huang, and Y. Hung. Efficient hierarchical method for background subtraction. *Journal of Pattern Recognition*, 40(10) :2706–2715, October 2007.

[112] Z. Chen, S. Babacan, R. Molina, and A. Katsaggelos. Variational Bayesian methods for multimedia problems. *IEEE Transaction on Multimedia*, 2014.

[113] Z. Chen, R. Molina, and A. Katsaggelos. A variational approach for sparse component estimation and low-rank matrix recovery. *Journal of Communication*, 8(9), September 2013.

[114] L. Cheng and M. Gong. Real time background subtraction from dynamics scenes. *International Conference on Computer Vision, ICCV 2009*, September 2009.

[115] L. Cheng, M. Gong, D. Schuurmans, and T. Caelli. Real-time discriminative background subtraction. *IEEE Transaction on Image Processing*, 20(5) :1401–1414, December 2011.

[116] S. Cheng, X. Luo, and S. Bhandarkar. A multiscale parametric background model for stationary foreground object detection. *International Workshop on Motion and Video Computing, WMCV 2007*, 2007.

[117] S. Cheung and C. Kamath. Robust background subtraction with foreground validation for urban traffic video. *Journal of Applied Signal Processing, EURASIP*, 2005.

[118] S. Chien, S. Ma, and L. Chen. Efficient moving object segmentation algorithm using background registration technique. *EEE Transaction on Circuits and Systems for Video Technology*, 12(7) :577–586, July 2012.

[119] P. Chiranjeevi and S. Sengupta. Detection of moving objects using fuzzy correlogram based background subtraction. *ICSIPA 2011*, 2011.

[120] P. Chiranjeevi and S. Sengupta. New fuzzy texture features for robust detection of moving objects. *IEEE Signal Processing Letters*, 19(10) :603–606, October 2012.

[121] P. Chiranjeevi and S. Sengupta. Robust detection of moving objects in video sequences through rough set theory framework. *Image and Vision Computing, IVC 2012*, 2012.

[122] S. Chiu, Y. Sheng, and C. Chang. Robust initial background extraction algorithm based on dynamic analysis. *International Conference on Information Science and Digital Content Technology, ICIDT 2012*, 3 :678–681, 2012.

[123] G. Choquet. Theory of capacities. *Ann. Inst. Fourier*, 5 :131–295, 1953.

[124] Y. Chu, X. Ye, J. Qian, Y. Zhang, and S. Zhang. Adaptive foreground and shadow segmentation using hidden conditional random fields. *Journal of Zhejiang University*, pages 586–592, 2007.

[125] T. Chua, K. Leman, and Y. Wang. Fuzzy rule-based system for dynamic texture and color based background subtraction. *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2012*, pages 1–7, June 2012.

[126] T. Chua, K. Leman, and Y. Wang. Fuzzy rule-based system for dynamic texture and color based background subtraction. *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2012*, pages 1–7, June 2012.

[127] Y. Chung, J. Wang, and S. Cheng. Progressive background image generation. *IPPR Conference on Computer Vision, Graphics and Image Processing, CVGIP 2002*, pages 858–865, 2002.

[128] A. Cichocki, A. Phan R. Zdunek, and S. Amari. Nonnegative matrix and tensor factorizations applications to exploratory multi-way data analysis and blind source separation. *Wiley*, September 2009.

[129] S. Cohen. Background estimation as a labeling problem. *International Conference on Computer Vision, ICCV 2005*, 2 :1034–1041, October 2005.

[130] R. Collins, A. Lipton, T. Kanade, H. Fujiyoshi, D. Duggin, Y. Tsin, D. Tolliver, N. Enomoto, and O. Hasegawa. A system for video surveillance and monitoring. *Technical Report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University*, May 2000.

[131] A. Colombari and A. Fusiello. Patch-based background initialization in heavily cluttered video. *IEEE Transactions on Image Processing*, 19(4) :926–933, April 2010.

[132] R. Mora Colque and G. Camara-Chavez. Progressive background image generation of surveillance traffic videos based on a temporal histogram ruled by a reward/penalty function. *SIBGRAPI 2011*, 2011.

[133] C. Conaire, N. O'Connor, E. Cooke, and A. Smeaton. Multispectral object segmentation and retrieval in surveillance video. *IEEE International Conference on Image Processing, ICIP 2006*, pages 2381–2384, 2006.

[134] M. Cristani, M. Bicego, and V. Murino. Integrated region- and pixel-based approach to background modeling. *IEEE Workshop on Motion and Video Computing, MOTION 2002*, 2002.

[135] M. Cristani, M. Bicego, and V. Murino. Multi-level background initialization using Hidden Markov Models. *ACM SIGMM Workshop on Video Surveillance, IWVS 2003*, pages 11–19, 2003.

[136] M. Cristani, M. Farenzena, D. Bloisi, and V. Murino. Background subtraction for automated multisensor surveillance : A comprehensive review. *EURASIP Journal on Advances in Signal Processing*, 2010 :24, 2010.

[137] M. Cristani and V. Murino. A spatial sampling mechanism for effective background subtraction. *International Conference on Computer Vision Theory and Applications, VISAPP 2007*, 2 :403–410, March 2007.

[138] M. Cristani and V. Murino. Background subtraction with adaptive spatio-temporal neighborhood analysis. *International Conference on Computer Vision Theory and Applications, VISAPP 2008*, January 2008.

[139] X. Cui and D. Metaxas Q. Liu, S. Zhang. Temporal spectral residual for fast salient motion detection. *Neurocomputing*, pages 24–32, 2012.

[140] Z. Cui, A. Li, and K. Jiang. Cooperative moving object segmentation using two cameras based on background subtraction and image registration. *Journal of Multimedia*, 9(3) :363–370, March 2014.

[141] D. Culbrik, O. Marques, D. Socek, H. Kalva, and B. Furht. A neural network approach to Bayesian background modeling for video object segmentation. *International Conference on Computer Vision Theory and Applications, VISAPP 2006*, February 2006.

[142] D. Culbrik, O. Marques, D. Socek, H. Kalva, and B. Furht. Neural network approach to background modeling for video object segmentation. *IEEE Transaction on Neural Networks*, 18(6) :1614–1627, 2007.

[143] D. Culibrk, V. Crnojevic, and B. Antic. Multiscale background modelling and segmentation. *International conference on Digital Signal Processing, DSP 2009*, pages 922–927, 2009.

[144] C. Dang, A. Moghadam, and H. Radha. RPCA-KFE : key frame extraction for consumer video based robust principal component analysis. *Preprint*, May 2014.

[145] S. Davarpanah, F. Khalid, , and M. Golchin. A block-based multi-scale background extraction algorithm. *Journal of Computer Science*, pages 1445–1451, 2010.

[146] S. Davarpanah, F. Khalid, N. Lili, S. Puteri, and M. Golchin. Using multi-scale filtering to initialize a background extraction model. *Journal of Computer Science*, pages 1077–1084, 2012.

[147] C. David, V. Gui, and F. Alexa. Foreground/background segmentation with learned dictionary. *International Conference on Circuits, Systems and Signals, CSS 2009*, pages 197–201, 2009.

[148] R. Davies, L. Mihaylova, N. Pavlidis, and I. Eckley. The effect of recovery algorithms on compressive sensing background subtraction. *Workshop Sensor Data Fusion : Trends, Solutions, and Applications*, 2013.

[149] J. Davis and M. Goadrich. The relationship between precision-recall and roc curves. *International Conference on Machine Learning, ICML 2006*, pages 233–240, 2006.

[150] Y. Deng, Q. Dai, R. Liu, and Z. Zhang. Low-rank structure learning via log-sum heuristic recovery. *Preprint*, 2012.

[151] B. Dey and M. Kundu. Robust background subtraction for network surveillance in H.264 streaming video. *IEEE Transactions on Circuits and Systems for Video Technology*, 2013.

[152] Y. Dhome, N. Tronson, A. Vacavant, T. Chateau, C. Gabard, Y. Goyat, and D. Gruyer. A benchmark for background subtraction algorithms in monocular vision : A comparative study. *International Conference on Image Processing Theory Tools and Applications, IPTA 2010*, pages 66–71, July 2010.

[153] M. Dikmen and T. Huang. Robust estimation of foreground in surveillance videos by sparse error estimation. *International Conference on Pattern Recognition, ICPR 2008*, December 2008.

[154] D. Ding, D. Zhou, X. He, and H. Zha. R1-PCA : rotational invariant l_1-norm principal component analysis for robust subspace factorization. *International Conference on Machine Learning, ICML 2006*, pages 281–288, 2006.

[155] J. Ding, M. Li, K. Huang, and T. Tan. Modeling complex scenes for accurate moving objects segmentation. *Asian Conference on Computer Vision, ACCV 2010*, pages 82–94, 2010.

[156] X. Ding, L. He, and L. Carin. Bayesian robust principal component analysis. *IEEE Transaction on Image Processing*, 2011.

[157] Y. Ding, W. Li, J. Fan, and H. Yang. A fuzzy background model for moving object detection. *IEEE International Conference on Computer-Aided Design and Computer Graphics, CAD/Graphics 2009*, pages 610–613, 2009.

[158] Y. Ding, W. Li, J. Fan, and H. Yang. Robust motion detection with color-texture feature fuzzy approach. *International Conference of COMPUTING in Engineering, Science and Informatics, ICC 2009*, April 2009.

[159] Y. Ding, W. Li, J. Fan, and H. Yang. A moving object detection algorithm based on Choquet integral. *Chinese Journal of Electronics*, 38(2) :263–268, February 2010.

[160] Y. Ding, W. Li, T. Fan, and H. Yang. Robust moving object detection under complex background. *Computer Science and Information Systems*, 7(1), February 2010.

[161] Y. Dong and G. DeSouza. Adaptive learning of multi-subspace for foreground detection under illumination changes. *Computer Vision and Image Understanding*, 2011.

[162] J. Dou and J. Li. Modeling the background and detecting moving objects based on SIFT flow. *Optik - International Journal for Light and Electron Optics*, 125(1) :435–440, January 2014.

[163] E. Durucan and T. Ebrahimi. Robust and illumination invariant change detection based on linear dependence. *European Signal Processing Conference, EUSIPCO 2000*, pages 1141–1144, September 2000.

[164] E. Durucan and T. Ebrahimi. Change detection and background extraction by linear algebra. *Special Issue of IEEE on Video Communications and Processing for Third Generation Surveillance*, October 2001.

[165] D. Dushnik, A. Schclar, and A. Averbuch. Video segmentation via diffusion bases. *Preprint*, 2013.

[166] A. Edelman, T. Arias, and S. Smith. The geometry of algorithms with orthogonality constraints. *IAM Journal on Matrix Analysis and Applications*, 20(2) :303–353, 1998.

[167] A. Elgammal and L. Davis. Non-parametric model for background subtraction. *European Conference on Computer Vision, ECCV 2000*, pages 751–767, June 2000.

[168] A. Elgammal, T. Parag, and A. Mittal. A framework for feature selection for background subtraction. *International Conference on Pattern Recognition, CVPR 2006*, June 2006.

[169] S. Elhabian, K. El-Sayed, and S. Ahmed. Moving object detection in spatial domain using background removal techniques - state-of-art. *Recent Patents on Computer Science*, 1(1) :32–54, January 2008.

[170] A. Elkabetz and Y. Yitzhak. Background modeling for moving object detection in long-distance imaging through turbulent medium. *Applied Optics*, January 2014.

[171] A. Elqursh and A. Elgammal. Online moving camera background subtraction. *European Conference on Computer Vision, ECCV 2012*, pages 228–241, 2012.

[172] D. Fan, M. Cao, and C. Lv. An updating method of self-adaptive background for moving objects detection in video. *ICALIP 2008*, pages 1497–1501, July 2008.

[173] W. Fan and N. Bouguila. Online variational learning of finite Dirichlet mixture models. *Evolving Systems*, January 2012.

[174] X. Fang, H. Xiong, B. Hu, and L. Wang. A moving object detection algorithm based on color information. *International Symposium on Instrumentation Science and Technology*, pages 384–387, 2006.

[175] X. Fang, W. Xiong, B. Hu, and L. Wang. A moving object detection algorithm based on color information. *International Symposium on Instrumentation Science and Technology*, 48 :384–387, 2006.

[176] D. Farcas and T. Bouwmans. Background modeling via a supervised subspace learning. *International Conference on Image, Video Processing and Computer Vision, IVPCV 2010*, pages 1–7, July 2010.

[177] D. Farcas, C. Marghes, and T. Bouwmans. Background Subtraction via Incremental Maximum Margin Criterion : A discriminative approach. *Machine Vision and Applications*, 2(3) :223–234, March 2012.

[178] A. Faro, D. Giordano, and C. Spampinato. Adaptive background modeling integrated with luminosity sensors and occlusion processing for reliable vehicle detection. *IEEE Transactions on Intelligent Transportation Systems*, 12(4) :1398–1412, December 2011.

[179] E. Fauske, L. Eliassen, and R. Bakken. A comparison of learning based background subtraction techniques implemented in cuda. *NAIS 2009*, pages 181–192, 2009.

[180] M. Fazel. Matrix rank minimization with applications. *Ph.D thesis,Stanford University*, March 2002.

[181] J. Feng, H. Xu, and S. Yan. Online robust pca via stochastic optimization. *NIPS 2013*, 2013.

[182] E. Fernandez-Sanchez, J. Diaz, and E. Ros. Background subtraction based on color and depth using active sensors. *Sensors*, 13 :8895–8915, 2013.

[183] E. Fernandez-Sanchez, L. Rubio, J. Diaz, and E. Ros. Background subtraction model based on color and depth cues. *Machine Vision and Application, Special Issue on Background Modeling*, 2013.

[184] A. Ferone and L. Maddalena. Neural background subtraction for Pan-Tilt-Zoom cameras. *IEEE Transactions on Systems, Man, and Cybernetics : Systems*, September 2013.

[185] M. Fettke, M. Naylor, K. Sammut, and F. He. Comparison of background models for video surveillance. *Digital Image Computing : Techniques and Applications Conference, DICTA 2002*, 2002.

[186] M. Fettke, K. Sammut, M. Naylor, and F. He. Evaluation of motion detection techniques for video surveillance. *IEEE International Conference on Information, Decision and Control*, 2002.

[187] M. Figueiredo, R. Nowak, and S. Wright. Gradient projection for sparse reconstruction : Application to compressed sensing and other inverse problems. *IEEE Journal on Selected Topics in Signal Processing*, 1 :586–597, 2007.

[188] S. Fu, G. Jiang, and M. Yu. Effective background subtraction method based on pixel change classification. *International Conference on Electrical and Control Engineering, ICECE 2010*, pages 4634–4637, 2010.

[189] N. El Gamal, F. Abou-Chadi, and H. El-Din Moustafa. A new technique for adding scribbles in video matting. *International Journal of Computer Science Issues, IJCSI 2014*, 11(1) :116, January 2014.

[190] S. Gandy and I. Yamada. Convex optimization techniques for the efficient recovery of a sparsely corrupted low-rank matrix. *Journal of Math-for-Industry*, 2 :147–156, 2010.

[191] B. Gao, T. Liu, Q. Cheng, and W. Ma. A linear approximation based method for noise-robust and illumination-invariant image change detection. *PCM 2004*, 2004.

[192] L. Gao, Y. Fan, N. Chen, Y. Li, and X. Li. Moving objects detection using adaptive region-based background model in dynamic scenes. *Foundations of Intelligent Systems, Advances in Intelligent and Soft Computing*, 2012.

[193] T. Gao, Z. Liu, W. Gao, and J. Zhang. A robust technique for background subtraction in traffic video. *International Conference on Neural Information Processing, ICONIP 2008*, pages 736–744, November 2008.

[194] Z. Gao, L. Cheong, and M. Shan. Block-sparse RPCA for consistent foreground detection. *European Conference on Computer Vision, ECCV 2012*, 2012.

[195] M. Genovese and E. Napoli. ASIC and FPGA implementation of the Gaussian mixture model algorithm for real-time segmentation of high definition video. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2013.

[196] M. Genovese, E. Napoli, D. De Caro, N. Petra, and A. Strollo. FPGA implementation of Gaussian mixture model algorithm for 47fps segmentation of 1080p video. *Journal of Electrical and Computer Engineering*, 2013.

[197] T. Gerhart. Convex optimization techniques and their application in hyperspectral video processing. *Thesis*, December 2013.

[198] J. Goes, T. Zhang, R. Arora, and G. Lerman. Robust stochastic principal component analysis. *AISTATS 2014*, 2014.

[199] D. Goldfarb, S. Ma, and K. Scheinberg. Fast alternating linearization methods for minimizing the sum of two convex function. *Preprint, Mathematical Programming Series A*, 2010.

[200] T. Goldstein and S. Osher. The split bregman method for l_1-regularized problems. *SIAM Journal of Image Science*, 2(2) :323–343, 2009.

[201] M. Gong and L. Cheng. Real time foreground segmentation on GPUs using local online learning and global graph cut optimization. *International Conference on Pattern Recognition, ICPR 2008*, December 2008.

[202] G. Gordon, T. Darrell, M. Harville, and J. Woodfill. Background estimation and removal based on range and color. *International Conference on Computer Vision and Pattern Recognition, CVPR 1999*, pages 459–464, June 1999.

[203] Y. Goyat, T. Chateau, and L. Trassoudaine. Vehicle trajectories evaluation by static video sensors. *IEEE Intelligent Transportation Systems Conference, ITSC 2006*, pages 864–869, September 2006.

[204] N. Goyette, P. Jodoin, F. Porikli, J. Konrad, and P. Ishwar. changedetection.net : A new change detection benchmark datase. *IEEE Workshop on Change Detection, CDW 2012 at CVPR 2012*, June 2012.

[205] M. Grabisch, S.A. Orlovski, and R.R. Yager. Fuzzy aggregation of numerical preferences. *The Handbooks of Fuzzy Sets Series*, 1998.

[206] M. Gregorio and M. Giordano. Change detection with weightless neural networks. *IEEE Change Detection Workshop, CDW 2014*, pages 315–336, 2014.

[207] M. Greiffenhagen, V. Ramesh, and H. Niemann. The systematics design and analysis cycle of a vision system : A case study in video surveillance. *CVPR*, 2 :704, 2001.

[208] E. Grossmann, A. Kale, and C. Jaynes. Towards interactive generation of ground truth in background subtraction from partially labeled examples. *IEEE Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, October 2005.

[209] E. Grossmann, A. Kale, C. Jaynes, and S. Cheung. Offline generation of high quality background subtraction data. *British Machine Vision Conference, BMVC 2005*, September 2005.

[210] G. Gu, B. He, and J. Yang. Inexact alternating direction based contraction methods for separable linearly constrained convex programming. *Journal of Optimization Theory and Applications*, December 2013.

[211] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. *Preprint*, March 2014.

[212] N. Guan, D. Tao, Z. Luo, and J. Shawe-Taylor. MahNMF : manhattan non-negative matrix factorization. *Journal of Machine Learning Research*, 2012.

[213] G. Guerra-Filho. Optical motion capture : Theory and implementation. *Journal of Theoretical and Applied Informatics (RITA), Brazilian Computing Society*, 12(2) :61–89, 2005.

[214] P. Guha, D. Palai, K. Venkatesh, and A. Mukerjee. A multiscale co-linearity statistic based approach to robust background modeling. *Asian Conference on Computer Vision, ACCV 2006*, pages 297–306, 2006.

[215] C. Guillot, M. Taron, P. Sayd, Q. Pham, C. Tilmant, and J. Lavest. Background subtraction for PTZ cameras performing a guard tour and application to cameras with very low frame rate. *ACCV 2010 Workshops*, pages 33–42, 2010.

[216] H. Guo, C. Qiu, and N. Vaswani. Practical ReProCS for separating sparse and low-dimensional signal sequences from their sum. *Preprint*, October 2013.

[217] J. Guo and C. Hsu. Cascaded background subtraction using block-based and pixel-based codebooks. *International Conference on Pattern Recognition, ICPR 2010*, August 2010.

[218] X. Guo and X. Cao. Speeding up low rank matrix recovery for foreground separation in surveillance videos. *International Conference on Multimedia and Expo, ICME 2014*, 2014.

[219] X. Guo, S. Li, and X. Cao. Motion matters : A novel framework for compressing surveillance videos. *ACM International Conference on Multimedia*, October 2013.

[220] D. Gutchess, M. Trajkovic, E. Cohen, D. Lyons, and A. Jain. A background model initialization for video surveillance. *International Conference on Computer Vision, ICCV 2001*, pages 733–740, 2001.

[221] C. Guyon, T. Bouwmans, and E. Zahzah. Foreground detection based on low-rank and block-sparse matrix decomposition. *IEEE International Conference on Image Processing, ICIP 2012*, September 2012.

[222] C. Guyon, T. Bouwmans, and E. Zahzah. Foreground detection by robust PCA solved via a linearized alternating direction method. *International Conference on Image Analysis and Recognition, ICIAR 2012*, pages 115–122, June 2012.

[223] C. Guyon, T. Bouwmans, and E. Zahzah. Foreground detection via robust low rank matrix decomposition including spatio-temporal constraint. *International Workshop on Background Model Challenges, ACCV 2012*, November 2012.

[224] C. Guyon, T. Bouwmans, and E. Zahzah. Foreground detection via robust low rank matrix factorization including spatial constraint with iterative reweighted regression. *International Conference on Pattern Recognition, ICPR 2012*, November 2012.

[225] C. Guyon, T. Bouwmans, and E. Zahzah. Moving object detection by robust PCA solved via a linearized symmetric alternating direction method. *International Symposium on Visual Computing, ISVC 2012*, July 2012.

[226] C. Guyon, T. Bouwmans, and E. Zahzah. Moving object detection via robust low rank matrix decomposition with IRLS scheme. *International Symposium on Visual Computing, ISVC 2012*, July 2012.

[227] C. Guyon, T. Bouwmans, and E. Zahzah. Robust Principal Component Analysis for Background Subtraction : Systematic Evaluation and Comparative Analysis. *Book 1, Chapter 12, Principal Component Analysis, INTECH*, pages 223–238, March 2012.

[228] C. Guyon, T. Bouwmans, and E. Zahzah. Background modeling and foreground detection via polynomial RPCA. *En rédaction*, 2014.

[229] D. Ha, J. Lee, and Y. Kim. Neural-edge-based vehicle detection and traffic parameter extraction. *Image and Vision Computing*, 22 :899–907, May 2004.

[230] C. Hage and M. Kleinsteuber. Robust PCA and subspace tracking from incomplete observations using l_0-surrogates. *Preprint*, 2012.

[231] T. Haines and T. Xiang. Background subtraction with Dirichlet processes. *European Conference on Computer Vision, ECCV 2012*, October 2012.

[232] D. Hall, J. Nascimento, P. Ribeiro, E. Andrade, P. Moreno, S. Pesnel, T. List, R. Emonet, R. Fisher, J. Santos-Victor, and J. Crowley. Comparison of target detection algorithms using adaptive background models. *IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, PETS 2005*, pages 113–120, October 2005.

[233] B. Han and L. Davis. Density-based multi-feature background subtraction with support vector machine. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI 2012*, 34(5) :1017–1023, May 2012.

[234] B. Han and R. Jain. Real-time subspace-based background modeling using multichannel data. *International Symposium on Visual Computing, ISVC 2007*, pages 162–172, November 2007.

[235] B. Han and X. Lin. Update the GMMs via adaptive kalman filtering. *Proceedings of SPIE*, pages 1506–1515, 2005.

[236] J. Han and Z. Pan. Robust moving objects detection in dynamic scenes based on mean shift. *International Conference on Cyberworlds, ICC 2008*, pages 271–275, 2008.

[237] J. Han, M. Zhang, and D. Zhang. Background modeling fusing local and global cues for temporally irregular dynamic textures. *Advanced Science Letters*, 7 :58–63, 2012.

[238] D. Hardoon, S. Szedmak, and S. Shawe. Canonical correlation analysis : An overview with application to learning methods. *Neural Computation*, 16 :2639–2664, 2006.

[239] M .Harville, G. Gordon, and J. Woodfill. Foreground segmentation using adaptive mixture models in color and depth. *International Workshop on Detection and Recognition of Events in Video*, July 2001.

[240] J. He, L. Balzano, and J. Luiz. Online robust subspace tracking from partial information. *IT 2011*, September 2011.

[241] J. He, M. Gao, L. Zhang, and H. Wu. Sparse signal recovery from fixed low-rank subspace via compressive measurement. *Algorithms 2013*, 6(4) :871–882, 2013.

[242] J. He, D. Zhang, L. Balzano, and T. Tao. Iterative Grassmannian optimization for robust image alignment. *Image and Vision Computing*, June 2013.

[243] R. He, B. Hu, W. Zheng, and X. Kong. Robust principal component analysis based on maximum correntropy criterion. *IEEE Transactions on Image Processing*, pages 1485–1494, 2011.

[244] M. Hedayati, W. Zaki, and A. Hussain. Real-time background subtraction for video surveillance : From research to reality. *International Colloquium on Signal Processing and Its Applications, CSPA 2010*, pages 1–6, May 2010.

[245] M. Heikkila and M. Pietikainen. A texture-based method for modeling the background and detecting moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI 2006*, 28(4) :657–62, 2006.

[246] S. Herrero and J. Bescos. Background subtraction techniques : Systematic evaluation and comparative analysis. *ACVIS 2009*, pages 33–42, 2009.

[247] J. Hiraiwa, E. Vargas, and S. Toral. An FPGA based embedded vision system for real-time motion segmentation. *International Conference on Systems, Signals and Image Processing, IWSSIP 2010*, pages 360–363, 2010.

[248] M. Hofmann, P.Tiefenbacher, and G. Rigoll. Background segmentation with feedback : The pixel-based adaptive segmenter. *IEEE Workshop on Change Detection, CVPR 2012*, June 2012.

[249] P. Holtzhausen, V. Crnojevic, and B. Herbst. An illumination invariant framework for real-time foreground detection. *Journal of Real Time Image Processing*, November 2012.

[250] Z. Hou and C. Han. A background reconstruction algorithm based on pixel intensity classification. *Journal of Software*, 9(16) :1568–1576, 2005.

[251] H. Hsiao and J. Leou. Background initialization and foreground segmentation for bootstrapping video sequences. *EURASIP Journal on Image and Video Processing*, pages 1–19, 2013.

[252] D. Hsu, S. Kakade, and T. Zhang. Robust matrix decomposition with sparse corruptions. *IEEE Transactions on Information Theory*, 57(11) :7221–7234, 2011.

[253] J. Hu and T. Su. Robust background subtraction with shadow and highlight removal for indoor surveillance. *International Journal on Advanced Signal Processing*, pages 1–14, 2007.

[254] J. Huang, X. Huang, and D. Metaxas. Learning with dynamic group sparsity. *International Conference on Computer Vision, ICCV 2009*, October 2009.

[255] J. Huang, T. Zhang, and D. Metaxas. Learning with structured sparsity. *International Conference on Machine Learning, ICML 2009*, 2009.

[256] I. Huerta, A. Amato, X. Roca, and J. Gonzalez. Exploiting multiple cues in motion segmentation based on background subtraction. *Neurocomputing, Special issue : Behaviours in video*, pages 183–196, January 2013.

[257] I. Huerta, D. Rowe, J. Gonzalez, and J. Villanueva. Efficient incorporation of motionless foreground objects for adaptive background segmentation. *Articulated Motion and Deformable Objects, AMDO 2006*, pages 424–433, 2006.

[258] M. Itoh, M. Kazui, and H. Fujii. Robust object detection based on radial reach correlation and adaptive background estimation for real-time video surveillance systems. *Real Time image Processing, SPIE 2008*, 2008.

[259] A. Mc Ivor. Background subtraction techniques. *International Conference on Image and Vision Computing New Zealand, IVCNZ 2000*, November 2010.

[260] V. Jain, B. Kimia, and J. Mundy. Background modelling based on subpixel edges. *ICIP 2007*, 6 :321–324, September 2007.

[261] O. Javed, K. Shafique, and M. Shah. A hierarchical approach to robust background subtraction using color and gradient information. *IEEE Workshop on Motion and Video Computing, WMVC 2002*, December 2002.

[262] S. Javed, S. Oh, and S. Jung. Foreground object detection via background modeling using histograms of oriented gradient. *International Conference On Human Computer Interaction, HCI-2013*, January 2013.

[263] T. Jebara and A. Pentland. Maximum conditional likelihood via bound maximization and the CEM algorithm. *NIPS 1999*, 1999.

[264] H. Jiang, H. Ardö, and V. Öwall. A hardware architecture for real-time video segmentation utilizing memory reduction techniques. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(2) :226–236, February 2009.

[265] H. Jiang, W. Deng, and Z. Shen. Surveillance video processing using compressive sensing. *Inverse Problems and Imaging*, 6(2) :201–214, 2012.

[266] P. Jimenez, S. Bascon, R. Pita, and H. Moreno. Background pixel classification for motion detection in video image sequences. *International Work Conference on Artificial and Natural Neural Network, IWANN 2003*, pages 718–725, 2003.

[267] P. KaewTraKulPong and R. Bowden. An improved adaptive background mixture model for real-time tracking with shadow detection. *AVBS 2001*, September 2001.

[268] A. Kamkar-Parsi, R. Laganier, and M. Bouchard. Multi-criteria model for robust foreground extraction. *VSSN 2005*, pages 67–70, November 2007.

[269] S. Kanna and N. Murthy. Background subtraction for dynamic texture scenes using local FCH features with adaptive updating procedure. *Indian Streams Research Journal*, 3 :1–13, December 2013.

[270] S. Kanprachar and S. Tangkawanit. Performance of RGB and HSV color systems in object detection applications under different illumination intensities. *International Multi Conference of Engineers and Computer Scientists*, 2 :1943–1948, March 2007.

[271] K. Karmann and A. Von Brand. Moving object recognition using an adaptive background memory. *Time-Varying Image Processing and Moving Object Recognition, Elsevier*, 1990.

[272] H. Kashani, S. Seyedin, and H. Yazdi. A novel approach in video scene background estimation. *International Journal of Computer Theory and Engineering*, 2(2) :274–282, April 2010.

[273] H. Kashani, S. Seyedin, and H. Yazdi. Background estimation in kernel space. *International Journal of Pattern Recognition and Artificial Intelligence*, 25(1) :1–35, 2011.

[274] I. Kavasidis, S. Palazzo, and C. Spampinato. An innovative web-based collaborative platform for video annotation. *Multimedia Tools and Applications*, pages 1–20, 2013.

[275] S. Kawabata, S. Hiura, and K. Sato. Real-time detection of anomalous objects in dynamic scene. *International Conference on Pattern Recognition, ICPR 2006*, 3 :1171–1174, August 2006.

[276] Y. Kawanishi, I. Mitsugami, M. Mukunoki, and M. Minoh. Background image generation keeping lighting condition of outdoor scenes. *International Conference on Security Camera Network, Privacy Protection and Community Safety, SPC 2009*, October 2009.

[277] P. Khorrami, J. Wang, and T. Huang. Multiple animal species detection using robust principal component analysis and large displacement optical flow. *Workshop on Visual Observation and Analysis of Animal and Insect Behavior, VAIB 2012*, 2012.

[278] E. Kim, M. Lee, C. Choi, N. Kwak, and S. Oh. Efficient l_1-norm-based low-rank matrix approximations for large-scale problems using alternating rectified gradient method. *International Conference on Multimedia and Expo, ICME 2014*, 2014.

[279] H. Kim, R. Sakamoto, I. Kitahara, T. Toriyama, and K. Kogure. Robust foreground extraction technique using Gaussian family model and multiple thresholds. *Asian Conference on Computer Vision, ACCV 2007*, pages 758–768, November 2007.

[280] K. Kim, T. Chalidabhongse, D. Harwood, and L. Davis. Background modeling and subtraction by codebook construction. *IEEE International Conference on Image Processing, ICIP 2004*, 2004.

[281] K. Kim, T. Chalidabhongse, D. Harwood, and L. Davis. Real time foreground background segmentation using codebook model. *Real time Imaging*, 11(3) :167–256, 2005.

[282] K. Kim, T. Chalidabhongse, D. Harwood, and L. Davis. PDR : performance evaluation method for foreground-background segmentation algorithms. *EURASIP Journal on Applied Signal Processing*, 2006.

[283] K. Kim, D. Harwood, and L. Davis. Background updating for visual surveillance. *International Symposium, ISVC 2005*, pages 337–346, December 2005.

[284] T. Kim, S. Lee, and J. Paik. Evolutionary algorithm-based background generation for robust object detection. *ICIC 2006*, pages 542–552, 2006.

[285] T. Kim, S. Wong, B. Stenger, J. Kittler, and R. Cipolla. Incremental linear discriminant analysis using sufficient spanning set approximations. *CVPR*, pages 1–8, June 2007.

[286] W. Kim and C. Kim. Background subtraction for dynamic texture scenes using fuzzy color histograms. *IEEE Signal Processing Letters*, 3(19) :127–130, March 2012.

[287] T. Kimura, M. Ohashi, K. Crailsheim, T. Schmickl, R. Odaka, and H. Ikeno. Tracking of multiple honey bees on a flat surface. *International Conference on Emerging Trends in Engineering and Technology, ICETET 2012*, pages 36–39, November 2012.

[288] B. Klare and S. Sarkar. Background subtraction in varying illuminations using an ensemble based on an enlarged feature set. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2009*, June 2009.

[289] T. Ko, S. Soatto, and D. Estrin. Background subtraction on distributions. *European Conference on Computer Vision, ECCV 2008*, pages 222–230, October 2008.

[290] T. Ko, S. Soatto, and D. Estrin. Warping background subtraction. *IEEE International Conference on Computer Vision and Pattern Recognition, CVPR 2010*, June 2010.

[291] M. Krishna, V. Aradhya, M. Ravishankar, and D. Babu. LoPP : Locality preserving projections for moving object detection. *International Conference on Computer, Communication, Control and Information Technology, C3IT 2012*, 4 :624–628, February 2012.

[292] M. Krishna, V. Aradhya, M. Ravishankar, and D. Babu. Ten-LoPP : tensor locality preserving projections approach for moving object detection and tracking. *nternational Conference on Computing and Information Technology, IC2IT 2013*, 4 :291–300, 2013.

[293] F. Kristensen, P. Nilsson, and V. Wall. Background segmentation beyond RGB. *ACCV 2006*, pages 602–612, 2006.

[294] G. J. Kulr and T. A. Folger. Fuzzy sets, Uncertainty, and Information. *Prentice-Hall International*, 1988.

[295] A. Kumar and V. Sindhwani. Near-separable non-negative matrix factorization with l_1 and Bregman loss functions. *Preprint*, 2013.

[296] P. Kumar and K. Sengupta. Foreground background segmentation using temporal and spatial Markov processes. *Department of Electrical and Computer Engineering, National University of Singapore*, November 2000.

[297] P. Kumar, A. Singhal, S. Mehta, and A. Mittal. Real-time moving object detection algorithm on high-resolution videos using GPUs. *ournal of Real-Time Image Processing*, January 2013.

[298] V. Kumar, B. Neha, C. Subhasis, and S. Guna. Fast compensation of illumination changes for background subtraction. *IEEE Applied Imagery Pattern Recognition Workshop : Sensing for Control and Augmentation, AIPR 2013*, pages 1–7, October 2013.

[299] N. Kwak. Principal component analysis based on l_1-norm maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1672–1680, 2008.

[300] C. Lallier, E. Reynaud, L. Robinault, and L. Tougne. A testing framework for background subtraction algorithms comparison in intrusion detection. *IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2011*, pages 314–319, September 2011.

[301] K. Lange, D. Hunter, and I. Yang. Optimization transfer using surrogate objective functions. *Journal of Computational and Graphical Statistics*, 9 :1–59, 2000.

[302] B. Langmann, S. Ghobadi, K. Hartmann, and O. Loffeld. Multi-model background subtraction using Gaussian mixture models. *Symposium on Photogrammetry Computer Vision and Image Analysis, PCV 2010*, pages 61–66, 2010.

[303] B. Langmann, K. Hartmann, and O. Loffeld. Depth assisted background subtraction for color capable tof-cameras. *International Conference on Image and Video Processing and Computer Vision, IVPCV 2010*, pages 75–82, July 2012.

[304] N. Lazarevic-McManus, J. Renno, D. Makris, and G. Jones. An object-based comparative methodology for motion detection based on the F-Measure. *Computer Vision and Image Understanding*, 111(1) :74–85, 2008.

[305] D. Lee. Improved adaptive mixture learning for robust video background modeling. *APR Workshop on Machine Vision for Application, MVA 2002*, pages 443–446, December 2002.

[306] S. Lee, N. Kim, I. Paek, M. Hayes, and J. Paik. Moving object detection using unstable camera for consumer surveillance systems. *International Conference on Consumer Electronics, ICCE 2013*, pages 145–146, January 2013.

[307] Y. Lee, J. Jung, and I. Kweon. Hierarchical on-line boosting based background subtraction. *Workshop on Frontiers of Computer Vision, FCV 2011*, pages 1–5, February 2011.

[308] W. Leow, Y. Cheng, L. Zhang, T. Sim, and L. Foo. Background recovery by fixed-rank robust principal component analysis. *CAIP 2013*, 2013.

[309] H. Li, T. Jiang, and K. Zhang. Efficient and robust feature extraction by maximum margin criterion. *Advances in Neural Information Processing Systems*, 16, 2004.

[310] J. Li, J. Wang, and W. Shen. Moving object detection in framework of compressive sampling. *Journal of Systems Engineering and Electronics*, 21(5) :740–745, October 2010.

[311] L. Li and W. Huang. Statistical modeling of complex background for foreground object detection. *IEEE Transaction on Image Processing*, 13(11) :1459–1472, November 2004.

[312] L. Li, P. Wang, Q. Hu, and S. Cai. A sparse outliers iterative removal algorithm to model the background in the video sequences. *Preprint*, 2014.

[313] R. Li, Y. Chen, and X. Zhang. Fast robust eigen-background updating for foreground detection. *ICIP 2006*, pages 1833–1836, 2006.

[314] W. Li, Y. Ding, and Y. Wang. Color-texture feature based background subtraction of video. *International Workshop on Information Technology, IWIT 2009*, September 2009.

[315] X. Li, W. Hu, Z. Zhang, and X. Zhang. Robust foreground segmentation based on two effective background models. *MIR 2008*, pages 223–228, October 2008.

[316] X. Li and X. Jing. FPGA based mixture Gaussian background modeling and motion detection. *International Conference on Natural Computation, ICNC 2011*, 4 :2078–2081, 2011.

[317] Y. Li. On incremental and robust subspace learning. *Pattern Recognition*, 37(7) :1509–1518, 2004.

[318] Y. Li, J. Yan, Y. Zhou, and J. Yang. Optimum subspace learning and error correction for tensors. *European Conference on Computer Vision, ECCV 2010*, 2010.

[319] D. Liang, S. Kaneko, M. Hashimoto, K. Iwata, X. Zhao, and Y. Satoh. Co-occurrence-based adaptive background model for robust object detection. *International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2013*, September 2013.

[320] Y. Liang, Z. Wang, X. Xu, and X. Cao. Background pixel classification for motion segmentation using mean shift algorithm. *ICLMC 2007*, pages 1693–1698, 2007.

[321] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikäinen, and S. Li. Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes. *International Conference on Computer Vision and Pattern Recognition, CVPR 2010*, June 2010.

[322] X. Lijun. Moving object segmentation based on background subtraction and fuzzy inference. *International Conference on Mechatronic Science, Electric Engineering and Computer, MEC 2011*, pages 434–437, August 2011.

[323] H. Lin, T. Liu, and J. Chuang. A probabilistic SVM approach for background scene initialization. *International Conference on Image Processing, ICIP 2002*, 3 :893–896, September 2002.

[324] L. Lin, Y. Xu, and X. Liang. Complex background subtraction by pursuing dynamic spatio-temporal manifolds. *Transactions on Transactions on Image Processing*, 2013.

[325] Z. Lin, M. Chen, L. Wu, and Y. Ma. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. *UIUC Technical Report*, November 2009.

[326] Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, and Y. Ma. Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix. *UIUC Technical Report*, August 2009.

[327] Z. Lin, R. Liu, and Z. Su. Linearized alternating direction method with adaptive penalty for low-rank representation. *NIPS 2011*, December 2011.

[328] Z. Lin and S. Wei. A block Lanczos with warm start technique for accelerating nuclear norm minimization algorithms. *Preprint*, 2010.

[329] J. Lindstrom, F. Lindgren, K. Ltrstrom, J. Holst, and U. Holst. Background and foreground modeling using an online EM algorithm. *IEEE International Workshop on Visual Surveillance VS 2006 in conjunction with ECCV 2006*, pages 9–16, May 2006.

[330] L. Liu and N. Sang. Metrics for objective evaluation of background subtraction algorithms. *International Conference on Image and Graphics, ICIG 2011*, August 2011.

[331] R. Liu, Z. Lin, S. Wei, and Z. Su. Solving principal component pursuit in linear time via l_1 filtering. *International Journal on Computer Vision, IJCV 2011*, 2011.

[332] X. Liu, Z. Wen, and Y. Zhang. Limited memory block Krylov subspace optimization for computing dominant singular value decomposition. *Preprint*, 2012.

[333] X. Liu, Z. Wen, and Y. Zhang. An efficient Gauss-Newton algorithm for symmetric low-rank product matrix approximations. *Technical Report*, June 2014.

[334] Y. Liu, L. Jiao, and F. Shang. A fast tri-factorization method for low-rank matrix recovery and completion. *Pattern Recognition, PR 2013*, 46 :163–173, January 2012.

[335] Y. Liu, L. Jiao, and F. Shang. An efficient matrix factorization based low-rank representation for subspace clustering. *Pattern Recognition, PR 2013*, 46 :284–292, January 2013.

[336] Y. Liu, H. Yao, W. Gao, X. Chen, and D. Zhao. Nonparametric background generation. *International Conference on Pattern Recognition, ICPR 2006*, 4 :916–919, 2006.

[337] Z. Liu, W. Chen, K. Huang, and T. Tan. Probabilistic framework based on KDE-GMM hybrid model for moving object segmentation in dynamic scenes. *International Workshop on Visual Surveillance, ECCV 2008*, October 2008.

[338] Z. Liu, K. Huang, and T. Tan. Foreground object detection using top-down information based on em framework. *IEEE Transactions on Image Processing*, 21(9) :4204–4217, September 2012.

[339] S. Liwick, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic. Euler principal component analysis. *International Journal on Computer Vision, IJCV 2013*, 101(3) :498–518, February 2013.

[340] W. Long and Y. Yang. Stationary background generation : An alternative to the difference of two images. *Pattern Recognition*, 12(23) :1351–1359, 1990.

[341] X. Lu, T. Izumi, L. Teng, T. Hori, and L. Wang. A novel background subtraction method for moving vehicle detection. *IEEJ Transactions on Fundamentals and Materials*, 132(10) :857–863, 2012.

[342] R. Luo, L. Li, and I. Gu. Efficient adaptive background subtraction based on multi-resolution background modelling and updating. *Pacific-RIM Conference on Multimedia, PCM 2007*, December 2007.

[343] R. Luque, E. Dominguez, E. Palomo, and J. Munoz. A neural network approach for video object segmentation in traffic surveillance. *International Conference on Image Analysis and Recognition, ICIAR 2008*, pages 151–158, 2008.

[344] R. Luque, F. Dominguez, E E. Palomo, and J. Muñoz. An ART-type network approach for video object detection. *European symposium on artificial neural networks*, pages 423–428, 2010.

[345] R. Luque, D. Lopez-Rodriguez, E. Dominguez, and E. Palomo. A dipolar competitive neural network for video segmentation. *Ibero-American Conference on Artificial Intelligence, IBERAMIA 2008*, pages 103–112, 2008.

[346] R. Luque, D. Lopez-Rodriguez, E. Merida-Casermeiro, and E. Palomo. Video object segmentation with multivalued neural networks. *IEEE International Conference on Hybrid Intelligent Systems, HIS 2008*, pages 613–618, 2008.

[347] F. Ma and N. Sang. Background subtraction based on multi-channel SILTP. *Asian Conference on Computer Vision, ACCV 2012*, November 2012.

[348] S. Ma. Algorithms for sparse and low-rank optimization : Convergence, complexity and applications. *Thesis*, Jun. 2011.

[349] L. Mackey, A. Talwalkar, and M. Jordan. Divide-and-conquer matrix factorization. *Neural Information Processing Systems, Neural Information Processing Systems, NIPS 2011*, December 2011.

[350] L. Maddalena and A. Petrosino. A self organizing approach to background subtraction for visual surveillance applications. *IEEE Transactions on Image Processing*, 17(7) :1729–1736, 2008.

[351] L. Maddalena and A. Petrosino. Multivalued background/foreground separation for moving object detection. *International Workshop on Fuzzy Logic and Applications, WILF 2009*, 5571 :263–270, June 2009.

[352] L. Maddalena and A. Petrosino. Self organizing and fuzzy modelling for parked vehicles detection. *Advanced Concepts for Intelligent Vision Systems, ACIVS 2009*, pages 422–433, 2009.

[353] L. Maddalena and A. Petrosino. A fuzzy spatial coherence-based approach to background/foreground separation for moving object detection. *Neural Computing and Applications, NCA 2010*, pages 1–8, 2010.

[354] L. Maddalena and A. Petrosino. The SOBS algorithm : What are the limits ? *IEEE Workshop on Change Detection, CVPR 2012*, June 2012.

[355] L. Maddalena and A. Petrosino. The 3dSOBS+ algorithm for moving object detection. *Computer Vision and Image Understanding, CVIU 2014*, 122(65-73), May 2014.

[356] D. Magee. Tracking multiple vehicles using foreground, background and motion models. *Image and Vision Computing*, 22 :143–155, 2004.

[357] H. Mansour and A. Vetro. Video background subtraction using semi-supervised robust matrix completion. *International Conference on Acoustics, Speech and Signal Processing, ICASSP 2014*, 2014.

[358] C. Marghes and T. Bouwmans. Background modeling via incremental maximum margin criterion. *International Workshop on Subspace Methods, ACCV 2010 Workshop Subspace 2010*, November 2010.

[359] C. Marghes, T. Bouwmans, and R. Vasiu. Background modeling and foreground detection via a reconstructive and discriminative subspace learning approach. *International Conference on Image Processing, Computer Vision, and Pattern Recognition, IPCV 2012*, July 2012.

[360] S. Mashak, B. Hosseini, and S. Abu-Bakar. Background subtraction for object detection under varying environments. *International Journal of Computer Information Systems and Industrial Management Applications*, 4 :506–513, 2012.

[361] G. Mateos and G. Giannakis. Sparsity control for robust principal component analysis. *International Conference on Signals, Systems, and Computers*, November 2010.

[362] C. McFarlane and N. Schofield. Segmentation and tracking of piglets in images. *British Machine Vision and Applications*, pages 187–193, 1995.

[363] J. Mendel. Type-2 fuzzy sets : Some questions and answers. *IEEE Connections, Newsletter of the IEEE Neural Networks Society*, pages 10–13, 2003.

[364] J. M. Mendel and R. I. John. Type-2 Fuzzy Sets Made Simple. *IEEE Trasactions on Fuzzy Systems*, 10(2), Avril 2002.

[365] D. Meng and F. De la Torre. Robust matrix factorization with unknown noise. *International Conference on Computer Vision, ICCV 2013*, December 2013.

[366] S. Messelodi, C. Modena, N. Segata, and M. Zanin. A Kalman filter based background updating algorithm robust to sharp illumination changes. *ICIAP 2005*, 3617 :163–170, September 2005.

[367] R. Mester, T. Aach, and L. Duembgen. Illumination-invariant change detection using a statistical colinearity criterion. *DAGM 2001*, pages 170–177, September 2001.

[368] J. San Miguel and J. Martinez. On the evaluation of background subtraction algorithms without ground-truth. *IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2010*, September 2010.

[369] M. Mizumoto and K. Tanaka. Some properties of fuzzy sets of Type-2. *Information and Control*, 31 :312–340, 1976.

[370] S. Molina-Giraldo, J. Carvajal-Gonzalez, A. Alvarez-Meza, and G. Castellanos-Dominguez. Video segmentation based on multi-kernel learning and feature relevance analysis for object classification. *International Conference on Pattern Recognition Applications and Methods, ICPRAM 2013*, February 2013.

[371] M. Molinier, T. Hame, and H. Ahola. 3d connected components analysis for traffic monitoring in image sequences acquired from a helicopter. *Scandinavian Conference, SCIA 2005*, page 141, June 2005.

[372] A. Moro, E. Mumolo, M. Nolich, K. Terabayashi, and K. Umeda. Improved foreground-background segmentation using Dempster-Shafer fusion. *nternational Symposium on Image and Signal Processing and Analysis, ISPA 2013*, pages 72–77, September 2013.

[373] Y. Mu, J. Dong, X. Yuan, and S. Yan. Accelerated low-rank visual recovery by random projection. *International Conference on Computer Vision, CVPR 2011*, 2011.

[374] D. Mukherjee and J. Wu. Real-time video segmentation using student's t mixture model. *International Conference on Ambient Systems, Networks and Technologies, ANT 2012*, pages 153–160, 2012.

[375] S. Nadimi and B. Bhanu. Physics-based cooperative sensor fusion for moving object detection. *IEEE Workshop on Learning in Computer Vision and Pattern Recognition, CVPR 2004*, June 2004.

[376] N. Naraghi. A comparative study of background estimation algorithms. *Master Thesis, Eastern Mediterranean University, North Cyprus*, September 2009.

[377] Y. Narukawa and T. Murofushi. Decision modelling using the Choquet integral. *Modeling Decisions for artificial Intelligence*, 3131 :183–193, 2004.

[378] J. Nascimento and J. Marques. New performance evaluation metrics for object detection algorithms. *International Workshop on Performance Evaluation of Tracking and Surveillance, PETS 2004*, May 2004.

[379] J. Nascimento and J. Marques. Novel metrics for performance evaluation of object detection algorithms. *ISR Workshop on Systems, Decision and Control Robotic Monitoring and Surveillance*, June 2005.

[380] J. Nascimento and J. Marques. Performance evaluation of object detection algorithms for video surveillance. *IEEE Transaction on Multimedia*, pages 761–774, August 2006.

[381] D. Needell and J. Tropp. CoSaMP : iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, June 2008.

[382] A. Nghiem, F. Bremond, M. Thonnat, and V. Valentin. ETISEO, performance evaluation for video surveillance systems. *IEEE Conference on Advanced Video and Signal Based Surveillance, AVSS 2007*, pages 476–481, September 2007.

[383] A. Nghiem, F.Bremond, and M. Thonnat. Controlling background subtraction algorithms for robust object detection. *ICDP 2009*, pages 1–6, December 2009.

[384] B. Nikolov, N. Kostov, and N. Velikov. Background subtraction methods for motion detection in video surveillance systems. *Journal of Marine Technology and Environment*, 3(2) :107–116, December 2010.

[385] S. Noh and M. Jeon. A new framework for background subtraction using multiple cues. *Asian Conference on Computer Vision, ACCV 2012*, November 2012.

[386] T. Oh. A novel low-rank constraint method with the sparsity model for moving object analysis. *Master Thesis, KAIST 2012*, 2012.

[387] N. Oliver, B. Rosario, and A. Pentland. A Bayesian computer vision system for modeling human interactions. *International Conference on Vision Systems, ICVS 1999*, January 1999.

[388] F. Orabona, A. Argyriou, and N. Srebro. Prisma : Proximal iterative smoothing algorithm. *Optimization and Control*, 2012.

[389] O. Oreifej, X. Li, and M. Shah. Simultaneous video stabilization and moving object detection in turbulence. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI 2012*, 2012.

[390] G. Pajares, J. Ruz, and J. Manuel de la Cruz. Performance analysis of homomorphic systems for image change detection. *IBPRIA 2005*, pages 563–570, 2005.

[391] E. Palomo, E. Dominguez, R. Luque, and J. Munoz. Image hierarchical segmentation based on a ghsom. *International Conference on Neural Information Processing, ICONIP 2009*, pages 743–750, 2009.

[392] S. Panahi, S. Sheikhi, S. Hadadan, and N. Gheissari. Evaluation of background subtraction methods. *Digital Image Computing : Techniques and Applications, DICTA 2008*, pages 357–364, 2008.

[393] G. Park. Background initialization by spatiotemporal similarity. *Journal of Broadcast Engineering*, 12(3) :289–292, 2007.

[394] T. Parker and P. Schniter. Bilinear generalized approximate message passing (big-amp) for matrix completion. *Asilomar Conference on Signals, Systems, and Computers*, November 2012.

[395] D. Parks and S. Fels. Evaluation of background subtraction algorithms with post-processing. *IEEE International Conference on Advanced Video and Signal-based Surveillance, AVSS 2008*, 2008.

[396] M. Paul, W. Lin, C. Lau, and B. Lee. Video coding with dynamic background. *EURASIP Journal on Advances in Signal Processing*, 2013 :11, 2013.

[397] N. Pejcic, N. Reljin, S. Daniel, D. Pokrajac, and A. Lazarevic. Detection of moving objects using incremental connectivity outlier factor algorithm. *ACM Conference*, 2009.

[398] D. Penciuc, F. El Baf, and T. Bouwmans. Comparison of background subtraction methods for an interactive learning space. *NETTIES 2006*, September 2006.

[399] D. Peng, C. Lin, W. Sheu, and T. Tsai. Architecture design for a low-cost and low-complexity foreground object segmentation with multi-model background maintenance algorithm. *ICIP 2009*, pages 3241–3244, 2009.

[400] T. Pham and A. Smeulders. Efficient projection pursuit density estimation for background subtraction. *International Workshop on Visual Surveillance, VS 2006*, May 2006.

[401] V. Pham, P. Vo, H. Vu Thanh, and B. Le Hoai. GPU implementation of extended Gaussian mixture model for background subtraction. *IEEE International Conference on Computing and Telecommunication Technologies, RIVF 2010*, November 2010.

[402] M. Piccardi. Background subtraction techniques : a review. *IEEE International Conference on Systems, Man and Cybernetics*, October 2004.

[403] M. Piccardi and T. Jan. Mean shift background image modeling. *IEEE International Conference on Image Processing, ICIP 2004*, October 2004.

[404] A. Pnevmatikakis and L. Polymenakos. 2d person tracking using Kalman filtering and adaptive background learning in a feedback loop. *CLEAR Workshop 2006*, 2006.

[405] D. Pokrajac and L. Latecki. Spatiotemporal blocks-based moving objects identification and tracking. *IEEE Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS 2003),*, pages 70–77, October 2003.

[406] G. Pope, M. Baumann, C. Studery, and G. Durisi. Real-time principal component pursuit. *Asilomar Conference on Signals, Systems, Computation*, November 2011.

[407] F. Porikli. Human body tracking by adaptive background models and mean-shift analysis. *IEEE Int Workshop on Performance Evaluation of Tracking and Surveillance, PETS 2003*, March 2003.

[408] F. Porikli. Detection of temporarily static regions by processing video at different frame rates. *IEEE International Conference on Advanced Video and Signal based Surveillance, AVSS 2007*, 2007.

[409] F. Porikli, Y. Ivanov, and T. Haga. Robust abandoned object detection using dual foregrounds. *EURASIP Journal on Advances in Signal Processing*, 2008 :11, 2008.

[410] F. Porikli and O. Tuzel. Bayesian background modeling for foreground detection. *ACM International Workshop on Video Surveillance and Sensor Networks, VSSN 2005*, pages 55–58, November 2005.

[411] F. Porikli and C. Wren. Change detection by frequency decomposition : Wave-back. *International Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2005*, April 2005.

[412] F. Porikli and Z. Yin. Temporally static region detection in multi-camera systems. *PETS 2007*, October 2007.

[413] A. Prati, I. Mikic, M. Trivedi, and R. Cucchiara. Detecting moving shadows : Algorithms and evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(4) :918–923, July 2003.

[414] B. Qin, J. Wang, J. Gao, T. Pang, and F. Su. A traffic video background extraction algorithm based on image content sensitivity. *CSI 2010*, pages 603–610, 2010.

[415] C. Qiu and N. Vaswani. Real-time robust principal components pursuit. *International Conference on Communication Control and Computing*, 2010.

[416] C. Qiu and N. Vaswani. ReProCS : a missing link between recursive robust PCA and recursive sparse recovery in large but correlated noise. *Preprint*, 2011.

[417] C. Qiu and N. Vaswani. Support predicted modified-CS for recursive robust principal components' pursuit. *IEEE International Symposium on Information Theory, ISIT 2011*, 2011.

[418] C. Qiu and N. Vaswani. Automated recursive projected CS (ReProCS) for real-time video layering. *International Conference on Computer Vision and Pattern Recognition, CVPR 2012,*, 2012.

[419] C. Qiu and N. Vaswani. Recursive sparse recovery in large but structured noise - part 1. *Preprint*, November 2012.

[420] C. Qiu and N. Vaswani. Recursive sparse recovery in large but structured noise - part 2. *Preprint*, 2012.

[421] C. Quivy and I. Kumazawa. Background images generation based on the nelder-mead simplex algorithm using the eigenbackground model. *International Conference onon Image Analysis and Recognition, ICIAR 2011*, pages 21–29, June 2011.

[422] R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms : A systematic survey. *IEEE Transactions on Image Processing*, 14(3) :294–307, March 2005.

[423] N. Rai, S. Chourasia, and A. Sethi. An efficient neural network based background subtraction method. *International Conference on Bio-Inspired Computing : Theories and Applications, BIC-TA 2012*, pages 453–460, 2012.

[424] S. Ramadan. Using time series analysis to visualize and evaluate background subtraction results in computer vision applications. *Master Thesis, University of Maryland*, 2006.

[425] I. Ramirez and G. Sapiro. An MDL framework for sparse coding and dictionary learning. *Preprint*, 2012.

[426] V. Reddy, C. Sanderson, and B. Lovell. Robust foreground object segmentation via adaptive region-based background modelling. *International Conference on Pattern Recognition, ICPR 2010*, August 2010.

[427] V. Reddy, C. Sanderson, A. Sanin, and B. Lovell. MRF-based background initialisation for improved foreground detection in cluttered surveillance videos. *Asian Conference on Computer Vision, ACCV 2010*, November 2010.

[428] N. Reljin, S. Daniel, D. Pokrajac, N. Pejcic, T. Vance, A. Lazarevic, and L. Latecki. Small moving targets detection using outlier detection algorithms. *SPIE 2010*, 2010.

[429] Q. Ren. Type-2 takagi-sugeno-kang fuzzy logic system and uncertainty in machining. *Ph.D. Thesis, Univ. Montréal, Canada*, 2012.

[430] R. Revathi and M. Hemalatha. Moving and immovable object in video processing using intuitionistic fuzzy logic. *International Conference on Computing, Communications and Networking Technologies, ICCCNT 2013*, pages 1–6, July 2013.

[431] H. Ribeiro and A. Gonzaga. Hand image segmentation in video sequence by GMM : a comparative analysis. *XIX Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI 2006*, pages 357–364, 2006.

[432] C. Robert and G. Casellah. Monte carlo statistical methods. *2nd edition New York : Springer*, 2004.

[433] A. Sanchez Rodriguez, J. Gonzalez Castolo, and O. Deniz Suarez. Timeviewer : a tool for visualizing the problems of the background subtraction. *Pacific-Rim Symposium, PSIVT 2013*, pages 372–384, October 2013.

[434] P. Rodriguez and B. Wohlberg. Fast principal component pursuit via alternating minimization. *IEEE International Conference on Image Processing, ICIP 2013*, September 2013.

[435] P. Rodriguez and B. Wohlberg. Video background modeling under impulse noise. *IEEE International Conference on Image Processing, ICIP 2014*, October 2014.

[436] R. Rodriguez-Gomez, E. Fernandez-Sanchez, J. Diaz, and E. Ros. Codebook hardware implementation on FPGA for background subtraction. *Journal of Real-Time Image Processing*, 2012.

[437] P. Rosin and E. Ioannidis. Evaluation of global image thresholding for change detection. *Pattern Recognition Letters*, 24 :2345–2356, October 2003.

[438] R. Rosipal and N. Kramer. Overview and recent advances in partial least squares. *SLSFS*, 3940 :34–35, 2005.

[439] J. Rossel-Ortega, G. Andrieu, A. Rodas-Jorda, and V. Atienza-Vanacloig. A combined self-configuring method for object tracking in colour video. *International Conference on Computer Vision, ICPR 2010*, August 2010.

[440] H. Rueda, L. Polania, and K. Barner. Robust tracking and anomaly detection in video surveillance sequences. *SPIE Airborne Intelligence, Surveillance, Reconnaissance, ISR 2012, Systems and Applications*, May 2012.

[441] J. Rymel, J. Renno, D. Greenhill, J. Orwell, and G. Jones. Adaptive eigen-backgrounds for object detection. *International Conference on Image Processing, ICIP 2004*, pages 1847–1850, October 2004.

[442] A. Sanin, C. Sanderson, and B. Lovell. Shadow detection : A survey and comparative evaluation of recent methods. *Pattern Recognition*, 45(4) :1684–1689, April 2012.

[443] Y. Satoh. Object detection based on radial reach filter under the change of background. *SPIE International Symposium on Optomechatronic Technologies*, pages 1–10, 2005.

[444] Y. Satoh, S. Kaneko, and S. Igarashi. Robust object detection and segmentation by peripheral increment sign correlation image. *Systems and Computers in Japan*, 35(9) :70–80, 2004.

[445] Y. Satoh, S. Kaneko, Y. Niwa, and K. Yamamoto. Robust object detection using a radial reach filter (RRF). *Wiley*, 35(10) :63–73, 2004.

[446] Y. Satoh and K. Sakaue. Robust background subtraction based on bi-polar radial reach correlation. *IEEE International Conference on Computers, Communications, Control and Power Engineering, TENCON 2005*, 2005.

[447] Y. Satoh, H. Tanahashi, Y. Niwa, S. Kaneko, and K. Yamamoto. Robust object detection and segmentation based on radial reach correlation. *IAPR Workshop on Machine Vision Applications, MVA 2002*, pages 512–517, 2002.

[448] Y. Satoh, H. Tanahashi, C. Wang, and S. Kaneko. Robust event detection by radial reach filter (RRF). *ICPR 2002*, 2002.

[449] K. Schindler and H. Wang. Smooth foreground-background segmentation for video processing. *Asian Conference on Computer Vision, ACCV 2006*, pages 581–590, January 2006.

[450] J. Schutte and A. Groenwold. A study of global optimization using particle swarms. *Jounal of Global Optimization*, 31(1) :93–108, 2005.

[451] F. Seidel, C. Hage, and M. Kleinsteuber. pROST - a smoothed Lp-norm robust online subspace tracking method for realtime background subtraction in video. *Machine Vision and Applications, Special Issue on Background Modeling*, 2013.

[452] N. Setiawan, S. Hong, J. Kim, and C. Lee. Gaussian mixture model in improved IHLS color space for human silhouette extraction. *International Conference on Artificial Reality and Telexistence, ICAT 2006*, pages 732–741, 2006.

[453] M. Shah, J. Deng, and B. Woodford. Illumination invariant background model using mixture of Gaussians and SURF features. *International Workshop on Background Models Challenge, ACCV 2012*, January 2006.

[454] M. Shah, J. Deng, and B. Woodford. Localized adaptive learning of mixture of Gaussians models for background extraction. *International Conference of Image and Vision Computing New Zealand, ICVNZ 2010*, November 2010.

[455] M. Shah, J. Deng, and B. Woodford. Enhanced codebook model for real-time background subtraction. *International Conference on Neural Information Processing, ICONIP 2011*, November 2011.

[456] M. Shakeri, H. Deldari, H. Foroughi, A. Saberi, and A. Naseri. A novel fuzzy background subtraction method based on cellular automata for urban traffic applications. *International Conference on Signal Processing, ICSP 2008*, pages 899–902, October 2008.

[457] M. Shakeri, H. Deldari, H. Foroughi, A. Saberi, and A. Naseri. A novel fuzzy background subtraction method based on cellular automata for urban traffic applications. *International Conference on Signal Processing, ICSP 2008*, pages 899–902, October 2008.

[458] M. Shakeri and H. Zhang. Real-time bird detection based on background subtraction. *World Congress on Intelligent Control and Automation, WCICA 2012*, pages 4507–4510, July 2012.

[459] W. Shao, Q. Ge, Z. Gan, H. Deng, and H. Li. A generalized robust minimization framework for low-rank matrix recovery. *Mathematical Problems in Engineering*, 2014.

[460] Y. Sheikh, O. Javed, and T. Kanade. Background subtraction for freely moving cameras. *IEEE International Conference on Computer Vision, ICCV 2009*, pages 1219–1225, September 2009.

[461] Y. Sheikh and M. Shah. Bayesian modeling of dynamic scenes for object detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27 :1778–1792, 2005.

[462] Y. Sheikh and M. Shah. Bayesian object detection in dynamic scenes. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2005*, June 2005.

[463] Y. Shen, W. Hu, Mi. Yang, J. Liu, C. Chou, and B. Wei. Efficient background subtraction for tracking in embedded camera networks. *International Conference on Information Processing in Sensor Networks, IPSN 2012*, April 2012.

[464] Y. Shen, Z. Wen, and Y. Zhang. Augmented Lagrangian alternating direction method for matrix separation based on low-rank factorization. *Preprint*, January 2011.

[465] H. Sheng, W. Suzhen, and W. Xin. l_1-regularized outlier isolation and regression. *Preprint*, June 2014.

[466] A. Shimada, D. Arita, and R. Taniguchi. Dynamic control of adpative mixture of Gaussians background model. *IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2006*, 2006.

[467] A. Shimada, H. Nagahara, and R. Taniguchi. Object detection based on spatio-temporal light field sensing. *IPSJ Transactions on Computer Vision and Applications*, 5 :129–133, 2013.

[468] A. Shimada and R. Taniguchi. Object detection based on Gaussian mixture predictive background model under varying illumination. *International Workshop on Computer Vision, MIRU 2008*, July 2008.

[469] A. Shimada and R. Taniguchi. Hybrid background model using spatial-temporal lbp. *IEEE International Conference on Advanced Video and Signal based Surveillance, AVSS 2009*, September 2009.

[470] A. Shimada and R. Taniguchi. Object detection based on fast and low-memory hybrid background model. *IEEJ Transactions on Electronics, Information and Systems*, 129-C(5) :846–852, May 2009.

[471] A. Shimada, S. Yoshinaga, and R. Taniguchi. Adaptive background modeling for paused object regions. *International Workshop on Visual Surveillance, VS 2010*, November 2010.

[472] A. Shimada, S. Yoshinaga, and R. Taniguchi. Maintenance of blind background model for robust object detection. *IPSJ Transactions on Computer Vision and Applications*, 3 :148–159, 2011.

[473] G. Shobha and N. Satish Kumar. Adaptive background modeling and foreground detection in video sequence using artificial neural network. *International Conference on Intelligent Computational Systems, ICICS 2012*, January 2012.

[474] X. Shu. Advanced imaging via multiplexed sensing and compressive sensing. *PhD Thesis,University of Illinois at Urbana-Champaign,*, 2013.

[475] X. Shu, F. Porikli, and N. Ahuja. Robust orthonormal subspace learning : Efficient recovery of corrupted low-rank matrices. *International Conference on Computer Vision and Pattern Recognition, CVPR 2014*, June 2014.

[476] M. Sigari. Fuzzy background modeling/subtraction and its application in vehicle detection. *World Congress on Engineering and Computer Science, WCECS 2008*, October 2008.

[477] M. Sigari, N. Mozayani, and H. Pourreza. Fuzzy running average and fuzzy background subtraction : Concepts and application. *International Journal of Computer Science and Network Security*, 8(2) :138–143, 2008.

[478] J. Silveira, C. Jung, and S. Musse. Background subtraction and shadow detection in grayscale video sequences. *Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI 2005*, pages 189–196, 2005.

[479] D. Silvestre. Video surveillance using a time-of-flight camera. *Master Thesis, Informatics and Mathematical Modelling, University of Denmark*, 2007.

[480] M. Singh, V. Parameswaran, and V. Ramesh. Order consistent change detection via fast statistical significance testing. *IEEE Computer Vision and Pattern Recognition Conference, CVPR 2008*, June 2008.

[481] M. Sivabalakrishnan and D.Manjula. Adaptive background subtraction in dynamic environments using fuzzy logic. *International Journal of Image Processing*, 4(1), 2010.

[482] D. Skocaj and A. Leonardis. Weighted and robust incremental method for subspace learning. *International Conference on Computer Vision, ICCV 2003*, pages 1494–1501, 2003.

[483] D. Skocaj and A. Leonardis. Canonical correlation analysis for appearance-based orientation and self-estimation and self-localization. *CogVis Meeting*, January 2004.

[484] D. Skocaj, A. Leonardis, M. Uray, and H. Bischof. Why to combine reconstructive and discriminative information for incremental subspace learning. *Computer Vision Winter Workshop, Czech Society for Cybernetics and Informatics*, February 2006.

[485] M. Smids. Background subtraction for urban traffic monitoring using webcams. *Master Thesis, Univ. Amsterdam*, December 2006.

[486] A. Sobral and A. Vacavant. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *Computer Vision and Image Understanding*, 2014.

[487] C. Spampinato, Y. Burger, G. Nadarajan, and R. Fisher. Detecting, tracking and counting fish in low quality unconstrained underwater videos. *VISAPP 2008*, pages 514–519, 2008.

[488] C. Spampinato, S. Palazzo, and I. Kavasidis. A texton-based kernel density estimation approach for background modeling under extreme conditions. *Computer Vision and Image Understanding, CVIU 2014*, 122 :74–83, 2014.

[489] P. Sprechmann, A. Bronstein, and G. Sapiro. Learning robust low-rank representations. *Optimization and Control*, 2012.

[490] C. Stauffer and E. Grimson. Adaptive background mixture models for real-time tracking. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 1999*, pages 246–252, 1999.

[491] E. Stringa. Morphological change detection algorithms for surveillance applications. *British Machine Vision Conference, BMVC 2000*, September 2000.

[492] T. Su, Y. Chen, and S. Lai. Over-segmentation based background modeling and foreground detection with shadow removal by using hierarchical MRFs. *Asian Conference on Computer Vision, ACCV 2010*, November 2010.

[493] B. Subudhi, S. Ghosh, and A. Ghosh. Change detection for moving object segmentation with robust background construction under wronskian framework. *Machine Vision and Applications, MVA 2013*, January 2013.

[494] Y. Sugaya and K. Kanata. Extracting moving objects from a moving camera video sequence. *Symposium on Sensing via Image Information, SSII 2004*, pages 279–284, June 2004.

[495] M. Sugeno. Theory of fuzzy integrals and its application. *Tokyo Institute of Technology*, 1974.

[496] M. Sugeno. Fuzzy measures and fuzzy integrals : A survey. *Fuzzy Automata and Decision Process*, pages 89–102, 1977.

[497] M. Sugeno and S. Kwon. A new approach to time series modeling with fuzzy measures and the Choquet integral. *IEEE International Conference on Fuzzy Systems*, pages 799–804, Mars 1995.

[498] Q. Sun, S. Xiang, and J. Ye. Robust principal component analysis via capped norms. *International Conference on Knowledge Discovery and Data Mining, KDD 2013*, pages 311–319, 2013.

[499] Y. Sun, X. Tao, Y. Li, and J. Lu. Robust two-dimensional principal component analysis via alternating optimization. *International Conference on Image Processing, ICIP 2013*, September 2013.

[500] H. Tahani and J. Keller. Information fusion in computer vision using the fuzzy integral. *Transaction on Systems, Man, and Cybernetics*, 20(3) :733–741, 1990.

[501] H. Tan, B. Cheng, J. Feng, G. Feng, W. Wang, and Y. Zhang. Low-n-rank tensor recovery based on multi-linear augmented Lagrange multiplier method. *Neurocomputing*, January 2013.

[502] H. Tan, B. Cheng, J. Feng, G. Feng, and Y. Zhang. Tensor recovery via multi-linear augmented Lagrange multiplier method. *International Conference on Image and Graphics, ICIG 2011*, pages 141–146, August 2011.

[503] H. Tan, B. Cheng, J. Feng, L. Liu, and W. Wang. Mixture augmented lagrange multiplier method for tensor recovery and its applications. *Special Issue on Green Intelligent Transport System*, 2013.

[504] T. Tanaka, A. Shimada, D. Arita, and R. Taniguchi. Object detection under varying illumination based on adaptive background modeling considering spatial locality. *International Workshop on Computer Vision, MIRU 2008*, July 2008.

[505] T. Tanaka, A. Shimada, D. Arita, and R. Taniguchi. Object segmentation under varying illumination based on combinational background modeling. *Joint Workshop on Machine Perception and Robotics, MPR 2008*, 2008.

[506] T. Tanaka, A. Shimada, D. Arita, and R. Taniguchi. Object detection under varying illumination based on adaptive background modeling considering spatial locality. *PSVIT 2009*, pages 645–656, January 2009.

[507] A. Taneja, L. Ballan, and M. Pollefeys. Modeling dynamic scenes recorded with freely moving cameras. *Asian Conference on Computer Vision, ACCV 2010*, 2010.

[508] F. Tang and H. Tao. Fast linear discriminant analysis using binary bases. *International Conference on Pattern Recognition,ICPR 2006*, 2, 2006.

[509] G. Tang and A. Nehorai. Robust principal component analysis based on low-rank and block-sparse matrix decomposition. *CISS 2011*, 2011.

[510] G. Tang and A. Nehorai. Robust principal component analysis based on low-rank and block-sparse matrix decomposition. *CISS 2011*, 2011.

[511] P. Tang, L. Gao, and Z. Liu. Salient moving object detection using stochastic approach filtering. *Fourth International Conference on Image and Graphics*, pages 530–535, 2007.

[512] M. Tao and X. Yuan. Recovering low-rank and sparse components of matrices from incomplete and noisy observations. *SIAM Journal on Optimization*, 21(1) :57–81, 2011.

[513] A. Tavakkoli, M. Nicolescu, and G. Bebis. Novelty detection approach for foreground region detection in videos with quasi-stationary backgrounds. *International Symposium on Visual Computing, ISVC 2006*, pages 40–49, November 2006.

[514] L. Taycher, J. Fisher, and T. Darrell. Incorporating object tracking feedback into background maintenance framework. *IEEE Workshop on Motion and Video Computing, WMVC 2005*, 2 :120–125, 2005.

[515] L. Teixeira, J. Cardoso, and L. Corte-Real. Object segmentation using background modelling and cascaded change detection. *Journal of Multimedia*, 2(5) :55–65, 2007.

[516] H. Tezuka and T. Nishitani. A precise and stable foreground segmentation using fine-to-coarse approach in transform domain. *International Conference on Image Processing, ICIP 2008*, pages 2732–2735, October 2008.

[517] F. De La Torre and M. Black. Robust principal component analysis for computer vision. *International Conference on Computer Vision, ICCV 2001*, July 2001.

[518] F. De La Torre and M. Black. A framework for robust subspace learning. *International Journal on Computer Vision*, pages 117–142, 2003.

[519] D. Toth, T. Aach, and V. Metzler. Bayesian spatio-temporal motion detection under varying illumination. *European Signal Processing Conference, EUSIPCO 2000*, pages 2081–2084, 2000.

[520] D. Toth, T. Aach, and V. Metzler. Illumination-invariant change detection. *IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 3–7, April 2000.

[521] K. Toyama, J. Krumm, B. Brumiit, and B. Meyers. Wallflower : Principles and practice of background maintenance. *International Conference on Computer Vision*, pages 255–261, September 1999.

[522] L. Tran, C. Navasca, and J. Luo. Video detection anomaly via low-rank and sparse decompositions. *IEEE New York Image Processing Workshop, WNYIPW 2012*, pages 17–20, November 2012.

[523] D. Tsai and C. Lai. Independent component analysis-based background subtraction for indoor surveillance. *IEEE Transactions on Image Processing, IP 2009*, 8(1) :158–167, January 2009.

[524] T. Tsai, D. Peng, C. Lin, and W. Sheu. A low cost foreground object detection architecture design with multi-model background maintenance algorithm. *VLSI 2008*, 2008.

[525] M. Himmelsbach U. Knauer, F. Winkler, F. Zautke, K. Bienefeld, and B. Meffert. Application of an adaptive background model for monitoring honeybees. *VIIP 2005*, 2005.

[526] M. Unger, M. Asbach, and P. Hosten. Enhanced background subtraction using global motion compensation and mosaicing. *International Conference on Image Processing, ICIP 2008*, pages 2708–2711, October 2008.

[527] L. Unzueta, M. Nieto, A. Cortes, J. Barandiaran, O. Otaegui, and P. Sanchez. Adaptive multi-cue background subtraction for robust vehicle counting and classification. *IEEE Transaction on Intelligent Transportation Systems*, 13(2) :527–540, June 2012.

[528] M. Uray, D. Skocaj, P. Roth, H. Bischof, and A. Leonardis. Incremental LDA learning by combining reconstructive and discriminative approaches. *BMVC 2007*, pages 272–281, 2007.

[529] A. Utasi and L. Czuni. Reducing the foreground aperture problem in mixture of Gaussians based motion detection. *EURASIP Conference on Speech and Image Processing, Multimedia Communications and Services, EC-SIPMCS 2007*, 2, 2007.

[530] A. Vacavant, T. Chateau, A. Wilhelm, and L. Lequievre. A benchmark dataset for foreground/background extraction. *International Workshop on Background Models Challenge, ACCV 2012*, November 2012.

[531] F. Wang and C. Zhang. Feature extraction by maximizing the average neighborhood margin. *CVPR 2007*, pages 1–8, 2007.

[532] H. Wang and A. Banerjee. Online alternating direction method. *Preprint*, 2013.

[533] H. Wang, A. Banerjee, and Z. Luo. Parallel direction method of multipliers. *Preprint*, June 2014.

[534] H. Wang and D. Suter. A re-evaluation of Mixture-of-Gaussian background modeling. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2005*, pages 1017–1020, March 2005.

[535] H. Wang and D. Suter. A re-evaluation of mixture of Gaussians background modeling. *IEEE Internationm Conference on Acoustics, Speech, and Signal Processing, ICASSP 2005*, pages 1017–1020, March 2005.

[536] H. Wang and D. Suter. A novel robust statistical method for background initialization and visual surveillance. *Asian Conference on Computer Vision, ACCV 2006*, pages 328–337, January 2006.

[537] J. Wang, G. Bebis, and R. Miller. Robust video-based surveillance by integrating target detection with tracking. *IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum in conjunction with CVPR 2006*, June 2006.

[538] J. Wang, M. Wan, X. Hu, and S. Yan. Image denoising with a unified schattern-p norm and l_q norm regularization. *Journal of Optimization Theory and Applications*, April 2014.

[539] L. Wang and C. Pan. Fast and effective background subtraction based on ELBP. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010*, March 2010.

[540] L. Wang and C. Pan. Effective multi-resolution background subtraction. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2011,*, May 2011.

[541] L. Wang, L. Wang, Q. Zhuo, H. Xiao, and W. Wang. Adaptive eigenbackground for dynamic background modeling. *Intelligent Computing in Signal Processing and Pattern Recognition, Lecture Notes in Control and Information Sciences*, 345 :670–675, 2006.

[542] N. Wang, T. Yao, J. Wang, and D. Yeung. A probabilistic approach to robust matrix factorization. *European Conference on Computer Vision, ECCV 2012*, 2012.

[543] N. Wang and D. Yeung. Bayesian robust matrix factorization for image and video processing. *International Conference on Computer Vision, ICCV 2013*, December 2013.

[544] T. Wang, J. Liang, X. Wang, and S. Wang. Background modeling using local binary patterns of motion vector. *EEE Visual Communications and Image Processing, VCIP 2012*, pages 1–5, November 2012.

[545] W. Wang, D. Chen, W. Gao, and J. Yang. Modeling background from compressed video. *International Conference on Computer Vision, ICCV 2005*, 2005.

[546] W. Wang and R. Wu. Fusion of luma and chroma GMMs for HMM-based object detection. *First Pacific Rim Symposium on Advances in Image and Video Technology, PSIVT 2006*, pages 573–581, December 2006.

[547] X. Wang, F. Liu, and Z. Ye. Background modeling in compressed sensing scheme. *ESEP 2011*, pages 4776–4783, December 2011.

[548] X. Wang, Z. Zhang, Y. Ma, X. Bai, W. Liu, and Z. Tu. Robust subspace discovery via relaxed rank minimization. *Neural Computation*, 2013.

[549] Y. Wang. Real-time moving vehicle detection with cast shadow removal in video based on conditional random field. *IEEE Transaction on Circuits Systems Video Technologies*, 19 :437–441, 2009.

[550] Y. Wang and H. Chen. The design of background subtraction on reconfigurable hardware. *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 182–185, 2012.

[551] Y. Wang, Y. Liu, and L. Wu. Study on background modeling method based on robust principal component analysis. *Annual Conference on Electrical and Control Engineering, ICECE 2011*, pages 6787–6790, September 2011.

[552] Y. Wang, K. Loe, T. Tan, and J. Wu. A dynamic Hidden Markov Random field model for foreground and shadow segmentation. *Workshops on Application of Computer Vision, WACV 2005*, 1 :474–480, January 2005.

[553] Y. Wang, K. Loe, and J. Wu. A dynamic conditional random field model for foreground and shadow segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(2) :279–289, February 2006.

[554] G. Warnell, S. Bhattacharya, R. Chellappa, and T. Basar. Adaptive-rate compressive sensing via side information. *Preprint*, January 2014.

[555] G. Warnell, D. Reddy, and R. Chellappa. Adaptive rate compressive sensing for background subtraction. *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, March 2012.

[556] A. Waters, A. Sankaranarayanan, and R. Baraniuk. SpaRCS : recovering low-rank and sparse matrices from compressive measurements. *Neural Information Processing Systems, NIPS 2011*, December 2011.

[557] C. Wei, Y. Huang, Y. Wang, and M. Shih. Background recovery in railroad crossing videos via incremental low-rank matrix decomposition. *Asian Conference on Pattern Recognition, ACPR 2013*, November 2013.

[558] J. Wen, Y. Xu, J. Tang, Y. Zhan, Z. Lai, and X. Guo. Joint video frame set division and low-rank decomposition for background subtraction. *IEEE Transactions on Circuits and Systems For Video Technology*, 2014.

[559] B. White and M. Shah. Automatically tuning background subtraction parameters using particle swarm optimization. *IEEE International Conference on Multimedia and Expo, ICME 2007*, pages 1826–1829, 2007.

[560] B. Wohlberg, R. Chartrand, and J. Theiler. Local principal component analysis for nonlinear datasets. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, March 2012.

[561] M. Wojcikowski, R. Zaglewsk, and B. Pankiewicz. FPGA-based real-time implementation of detection algorithm for automatic traffic surveillance sensor network. *Journal of Signal Processing Systems*, December 2010.

[562] C. Wren and A. Azarbayejani. Pfinder : Real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7) :780–785, July 1997.

[563] C. Wren and F. Porikli. Waviz : Spectral similarity for object detection. *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, PETS 2005*, January 2005.

[564] J. Wright, Y. Peng, Y. Ma, A. Ganesh, and S. Rao. Robust principal component analysis : Exact recovery of corrupted low-rank matrices by convex optimization. *Neural Information Processing Systems, NIPS 2009*, December 2009.

[565] J. Wu and Y. Wu. Detecting moving objects based on fuzzy non-symmetric membership function in color space. *International Journal of Engineering and Industries*, 2(1), March 2011.

[566] X. Wu, Y. Wang, and J. Li. Video background segmentation using adaptive background models. *ICIAP 2009*, pages 623–632, September 2009.

[567] M. Xiao, C. Han, and X. Kang. A background reconstruction for dynamic scenes. *International Conference on Information Fusion, ICIF 2006*, pages 1–6, July 2006.

[568] B. Xie, V. Ramesh, and T. Boult. Sudden illumination change detection using order consistency. *Image and Vision Computing*, 22(2) :117–125, February 2004.

[569] L. Xiong, X. Chen, and J. Schneider. Direct robust matrix factorization for anomaly detection. *International Conference on Data Mining, ICDM 2011*, 2011.

[570] H. Xu, C. Caramanis, and S. Sanghavi. Robust PCA via outlier pursuit. *NIPS 2010*, 2010.

[571] J. Xu, V. Ithapu, L. Mukherjee, J. Rehg, and V. Singhy. GOSUS : grassmannian online subspace updates with structured-sparsity. *International Conference on Computer Vision, ICCV 2013*, December 2013.

[572] M. Xu and T. Ellis. Illumination-invariant motion detection using color mixture models. *British Machine Vision Conference, BMVC 2001*, pages 163–172, September 2001.

[573] M. Xu and J. Lu. K-cluster-valued compressive sensing for imaging. *EURASIP Journal on Advances in Signal Processing*, 2011.

[574] Z. Xu, I. Gu, and P. Shi. Recursive error-compensated dynamic eigenbackground learning and adaptive background subtraction in video. *Optical Engineering*, 47(5), May 2008.

[575] G. Xue, L. Song, J. Sun, and M. Wu. Foreground estimation based on robust linear regression model. *International Conference on Image Processing, ICIP 2011*, pages 3330–3333, September 2011.

[576] G. Xue, L. Song, J. Sun, and M. Wu. Hybrid center-symmetric local pattern for dynamic background subtraction. *International Conference in Multimedia and Exposition, ICME 2011*, July 2011.

[577] G. Xue, J. Sun, and L. Song. Dynamic background subtraction based on spatial extended center-symmetric local binary pattern. *International Conference in Multimedia and Exposition, ICME 2010*, July 2010.

[578] Y. Xue, X. Gu, and X. Cao. Motion saliency detection using low-rank and sparse decomposition. *International Conference on Acoustics, Speech, and Signal Processing,ICASSP 2012*, March 2012.

[579] M. Yamazaki, G. Xu, and Y. Chen. Detection of moving objects by independent component analysis. *Asian Conference on Computer Vision, ACCV 2006*, pages 467–478, 2006.

[580] J. Yan, B. Zhang, S. Yan, Q. Yang, H. Li, Z. Chen, W. Xi, W. Fan, W. Ma, and Q. Cheng. IMMC : incremental maximum margin criterion. *KDD 2004*, pages 725–730, August 2004.

[581] H. Yang, Y. Nam, W. Cho, and Y. Choi. Adaptive background modeling for effective ghost removal and robust left object detection. *International Conference on Information Technology Convergence and Services, ITCS 2010*, pages 1–6, August 2010.

[582] H. Yang, Y. Tan, J. Tian, and J. Liu. Accurate dynamic scene model for moving object detection. *International Conference on Image Processing, ICIP 2007*, pages 157–160, 2007.

[583] J. Yang, X. Sun, X. Ye, and K. Li. Background extraction from video sequences via motion-assisted matrix completion. *IEEE International Conference on Image Processing, ICIP 2014*, October 2014.

[584] J. Yang and X. Yuan. Linearized augmented Lagrangian and alternating direction methods for nuclear norm minimization. *Preprint*, 2011.

[585] M. Yang. Background modeling from surveillance video using rank minimization. *Artificial Intelligence and Computational Intelligence, AICI 2012*, pages 769–774, 2012.

[586] M. Yang and Z. An. Video background modeling using low-rank matrix recovery. *Journal of Nanjing University of Posts and Telecommunications*, April 2013.

[587] M. Yang and Y. Wang. Fast alternating direction method of multipliers for robust PCA. *Journal of Nanjing University*, 34(2) :83–88, April 2014.

[588] S. Yang and C. Hsu. Background modeling from GMM likelihood combined with spatial and color coherency. *ICIP 2006*, pages 2801–2804, 2006.

[589] T. Yang, S. Li, Q. Pan, and J. Li. Real-time and accurate segmentation of moving objects in dynamic scene. *ACM International Workshop on Video Surveillance and Sensor Networks, VSSN 2004*, October 2004.

[590] W. Yang, L. Dou, and J. Zhan. A multi-histogram clustering approach toward markov random field for foreground segmentation. *International Journal of Image and Graphics*, October 2010.

[591] W. Yang, J. Wang, M. Ren, and J. Yang. Feature extraction base on local maximum margin criterion. *ICPR 2008*, pages 1–4, December 2008.

[592] Y. Yang, Y. Feng, and J. Suykens. Nonconvex relaxation approach to robust matrix completion. *Preprint*, 2014.

[593] H. Yao and I. Ahmad. Adaptive foreground segmentation using fuzzy approach. *International Conference on Digital Information Management, ICDIM 2009*, pages 104–109, November 2009.

[594] B. Yeo, W. Lim, H. Lim, and W. Wong. Extended fuzzy background modeling for moving vehicle detection using infrared vision. *IEICE Electronics Express*, pages 340–345, 2011.

[595] B. Yeo, W. Lim, H. Lim, and W. Wong. Extended fuzzy background modeling for moving vehicle detection using infrared vision. *IEICE Electronics Express*, pages 340–345, 2011.

[596] H. Yin, H. Yang, H. Su, and C. Zhang. Dynamic background subtraction based on appearance and motion pattern. *International Conference on Multimedia and Expo Workshops, ICMEW 2013*, pages 1–6, July 2013.

[597] K. Yokoi. Illumination-robust change detection using texture based featuress. *Machine Vision Applications, MVA 2007*, pages 487–491, May 2007.

[598] K. Yokoi. Probabilistic BPRRC : robust change detection against illumination changes and background movements. *Machine Vision Applications, MVA 2009*, pages 148–151, May 2009.

[599] S. Yoshinaga, A. Shimada, H. Nagahara, and R. Taniguchi. Object detection using local difference patterns. *Asian Conference on Computer Vision, ACCV 2010*, pages 216–227, November 2010.

[600] S. Yoshinaga, A. Shimada, H. Nagahara, and R. Taniguchi. Background model based on statistical local difference pattern. *International Workshop on Background Models Challenge, ACCV 2012*, November 2012.

[601] L. Yu and Y. Kong. An algorithm of background reconstruction based on morphology and PIC. *International Workshop on Information Security and Application*, November 2009.

[602] G. Yuan, Y. Gao, D. Xu, and M. Jiang. A new background subtraction method using texture and color information. *ICIC 2011*, pages 541–548, 2011.

[603] X. Yuan. Nuclear-norm-free variational models for background extraction from surveillance video. *Cross-straits Optimization Workshop, COW 2013*, March 2013.

[604] X. Yuan and J. Yang. Sparse and low-rank matrix decomposition via alternating direction methods. *Optimization Online*, November 2009.

[605] J. Yuk and K. Wong. An efficient pattern-less background modeling based on scale invariant local states. *IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2011*, 2011.

[606] S. Yusuf, D. Brown, A. Mackinnon, and R. Papanicolaou. Pixel-level, intensity-based nonlinear autoregressive classifier (NARX) with chromatic exogenous input for efficient image background subtraction. *International Joint Conference on Neural Networks, IJCNN 2013*, August 2013.

[607] Y. Zhang Z. Li, W. Liu. A background modeling method based on adaptive fuzzy estimation. *Journal of South China University of Technology*, 41(9) :77–81, September 2013.

[608] Y. Zhang Z. Li, W. Liu. A background modeling method based on adaptive fuzzy estimation. *Journal of South China University of Technology*, 41(9) :77–81, September 2013.

[609] L. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information Sciences*, 8 :199–249, 1975.

[610] D. Zamalieva and A. Yilmaz. Background subtraction for the moving camera : A geometric approach. *Computer Vision and Image Understanding*, 2014.

[611] D. Zamalieva, A. Yilmaz, and J. Davis. Exploiting temporal geometry for moving camera background subtraction. *ICPR 2014*, August 2014.

[612] D. Zamalieva, A. Yilmaz, and J. Davis. A multi-transformational model for background subtraction with moving cameras. *ECCV 2014*, September 2014.

[613] Q. Zang and R. Klette. Evaluation of an adaptive composite Gaussian model in video surveillance. *CITR Technical Report 114*, August 2002.

[614] Q. Zang and R. Klette. Robust background subtraction and maintenance. *International Conference on Pattern Recognition, ICPR 2004*, pages 90–93, 2004.

[615] H. Zen and S. Lai. Adaptive foreground object extraction for real-time video surveillance with lighting variations. *ICASSP 2007*, 1 :1201–1204, 2007.

[616] J. Zeng, L. Xie, and Z. Liu. Type-2 Fuzzy Gaussian Mixture Models. *Pattern Recognition*, June 2008.

[617] J. Zhan and N. Vaswani. Robust PCA with partial subspace knowledge. *Preprint*, 2014.

[618] E. Zhang, F. Chen, and W. Zhang. A novel particle filter based background subtraction method. *Computational Intelligence and Security, CIS 2006*, November 2006.

[619] H. Zhang and L. Liu. Recovering low-rank and sparse components of matrices for object detection. *Electronics Letters*, 49(2), January 2013.

[620] H. Zhang and D. Xu. Fusing color and texture features for background model. *International Conference on Fuzzy Systems and Knowledge Discovery, FSKD*, 4223(7) :887–893, September 2006.

[621] J. Zhang, Y. Tian, Y. Yang, and C. Zhu. Robust foreground segmentation using subspace based background model. *APCIP 2009*, 2 :214–217, July 2009.

[622] J. Zhang and Y. Zhuang. Adaptive weight selection for incremental eigen-background modeling. *ICME 2007*, July 2007.

[623] S. Zhang, H. Yao, and S. Liu. Dynamic background and subtraction using spatio-temporal local binary patterns. *IEEE International Conference on Image Processing, ICIP 2008*, pages 1556–1559, October 2008.

[624] S. Zhang, H. Yao, and S. Liu. Dynamic background subtraction based on local dependency histogram. *International Journal of Pattern Recognition and Artificial Intelligence, IJPRAI 2009*, 2009.

[625] W. Zhang, J. Wu, and X. Fang. Moving cast shadow detection. *INTECH*, page 546, June 2007.

[626] Y. Zhang, Z. Liang, Z. Hou, H. Wang, and M. Tan. An adaptive mixture Gaussian background model with online background reconstruction and adjustable foreground mergence time for motion segmentation. *ICIT 2005*, pages 23–27, December 2005.

[627] Z. Zhang, C. Wang, B. Xiao, S. Liu, and W. Zhou. Multi-scale fusion of texture and color for background modeling. *International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2012*, September 2012.

[628] C. Zhao, X. Wang, and W. Cham. Background subtraction via robust dictionary learning. *EURASIP Journal on Image and Video Processing, IVP 2011*, January 2011.

[629] D. Zhao, D. Liu, and Y. Yang. An improved PIC algorithm of background reconstruction for detecting moving object. *International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2008*, 4 :24–28, 2008.

[630] L. Zhao, X. Zhang, Y. Tian, R. Wang, and T. Huang. A background proportion adaptive lagrange multiplier selection method for surveillance video on high HEVC. *International Conference on Multimedia and Expo, ICME 2013*, July 2013.

[631] L. Zhao, X. Zhang, Y. Tian, R. Wang, and T. Huang. A background proportion adaptive Lagrange multiplier selection method for surveillance video on high HEVC. *International Conference on Multimedia and Expo, ICME 2013*, July 2013.

[632] M. Zhao, J. Bu, and C. Chen. Robust background subtraction in HSV color space. *SPIE Multimedia Systems and Applications*, 4861 :325–332, July 2002.

[633] Q. Zhao, D. Meng, Z. Xu, W. Zuo, and L. Zhang. Robust principal component analysis with complex noise. *ICML 2014*, 2014.

[634] Y. Zhao, H. Gong, L. Lin, and Y. Jia. Spatio-temporal patches for night background modeling by subspace learning. *ICPR 2008*, 2008.

[635] J. Zheng, Y. Wang, N. Nihan, and E. Hallenbeck. Extracting roadway background image : A mode based approach. *Journal of Transportation Research Report*, 1944 :82–88, 2006.

[636] Y. Zheng, G. Liu, S. Sugimoto, S. Yan, and M. Okutomi. Practical low-rank matrix approximation under robust $l1$-norm. *CVPR 2012*, 2012.

[637] B. Zhong, S. Liu, and H. Yao. Local spatial co-occurrence for background subtraction via adaptive binned kernel estimation. *Asian Conference on Computer Vision, ACCV 2009*, September 2009.

[638] B. Zhong, S. Liu, H. Yao, and B. Zhang. Multi-resolution background subtraction for dynamic scenes. *International Conference on Image Processing, ICIP 2009*, pages 3193–3196, November 2009.

[639] B. Zhong, H. Yao, S. Shan, X. Chen, and W. Gao. Hierarchical background subtraction using local pixel clustering. *IEEE International Conference on Pattern Recognition, ICPR 2008*, December 2008.

[640] B. Zhong, H. Yao, and X Yuan. Local histogram of figure/ground segmentations for dynamic background subtraction. *EURASIP Journal on Advances in Signal Processing*, 2010 :14, 2010.

[641] Q. Zhong, L. Dai, Y. Song, and R. Wang. A hierarchical motion detection algorithm with the fusion of the two types of motion information. *Pattern Recognition*, 18(5) :552–557, 2005.

[642] D. Zhou and H. Zhang. Modified GMM background modeling and optical flow for detection of moving objects. *IEEE Internationm Conference on Systems, Man and Cybernetics, SMC 2005*, pages 2224–2229, October 2005.

[643] T. Zhou and D. Tao. GoDec : randomized low-rank and sparse matrix decomposition in noisy case. *International Conference on Machine Learning, ICML 2011*, 2011.

[644] T. Zhou and D. Tao. Greedy bilateral sketch, completion and smoothing for large-scale matrix completion, robust PCA and low-rank approximation. *AISTATS 2013*, 2013.

[645] W. Zhou, Y. Liu, W. Zhang, L. Zhuang, and N. Yu. Dynamic background subtraction using spatial-color binary patterns. *International Conference on Image and Graphics, ICIG 2011*, August 2011.

[646] X. Zhou, C. Yang, and W. Yu. Moving object detection by detecting contiguous outliers in the low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 :597–610, 2013.

[647] Z. Zhou, X. Li, J. Wright, E. Candes, and Y. Ma. Stable principal component pursuit. *IEEE ISIT Proceedings*, pages 1518–1522, June 2010.

[648] W. Zhu, S. Shu, and L. Cheng. Proximity point algorithm for low-rank matrix recovery from sparse noise corrupted data. *Applied Mathematics and Mechanics*, 35(2) :259–268, February 2014.

[649] Z. Zivkovic. Improved adaptive Gaussian mixture model for background subtraction. *International Conference Pattern Recognition, ICPR 2004*, pages 28–31, 2004.

[650] H. Zou, T. Hastie, and T. Tibshirani. Sparse principal component analysis. *Journal of Computation and Graphical Statistics*, 15(2) :265–286, 2006.

Index

, 157, 163, 238

Résumé

Ce manuscrit présente mes activités scientifiques à l'Université de la Rochelle depuis ma nomination. Ces activités de recherche se sont concentrées sur les méthodes de soustraction de fond pour la détection d'objets mobiles dans des séquences vidéos acquises par des caméras fixes dans des conditions non-contrôlées. Bien que cette étape peut ne pas être nécessaire dans certaines applications, elle se révèle souvent très utile et très utilisée pour distinguer et séparer le fond et les objets mobiles dans des images vidéo d'applications comme la vidéo-surveillance, la capture de mouvement optique et d'autres applications du domaine multimédia. Durant la dernière décennie, on recense un nombre important de publications présentant des méthodes spécifiques pour des applications spécifiques. Devant ce contexte, j'ai engagé un état de l'art systématique du domaine tant sur certaines catégories de modèles que des sous-catégories plus ou moins étudiées à l'époque. J'ai ensuite dégagé des pistes nouvelles afin de développer des méthodes robustes aux challenges importants du domaine comme les fonds dynamiques et les variations lumineuses. Pour cela, j'ai engagé et encadré des travaux de recherche dans le cadre de différentes thèses pour proposer l'utilisation d'autres outils mathématiques comme les concepts flous, les méthodes de réduction de dimension discriminatives et mixtes et les méthodes d'ACP robuste par décomposition de matrices de rang faible et parcimonieuse.

Mots-clés : Modélisation de fond, Détection d'objets mobiles, Concepts flous, Réduction de dimension, ACP robustes