



HAL
open science

Reconstruction et analyse de trajectoires 2D d'objets mobiles par modélisation markovienne et par la théorie de l'évidence à partir de séquences d'images monoculaires - Application à l'évaluation du danger aux passages à niveau

Houssam Salmane

► **To cite this version:**

Houssam Salmane. Reconstruction et analyse de trajectoires 2D d'objets mobiles par modélisation markovienne et par la théorie de l'évidence à partir de séquences d'images monoculaires - Application à l'évaluation du danger aux passages à niveau. Traitement des images [eess.IV]. Université de Technologie de Belfort-Montbéliard, 2013. Français. NNT: . tel-00953503

HAL Id: tel-00953503

<https://theses.hal.science/tel-00953503>

Submitted on 4 Mar 2014

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.

SPIM

Thèse de Doctorat



école doctorale sciences pour l'ingénieur et microtechniques

UNIVERSITÉ DE TECHNOLOGIE BELFORT-MONTBÉLIARD

Reconstruction et analyse de trajectoires 2D d'objets mobiles par modélisation markovienne et par la théorie de l'évidence à partir de séquences d'images monoculaires - Application à l'évaluation du danger aux passages à niveau

Thèse préparée à IRTES-SET et LEOST-IFSTTAR, Projet ANR PANSAFER

■ Houssam SALMANE

SPIM

Thèse de Doctorat



école doctorale sciences pour l'ingénieur et microtechniques
UNIVERSITÉ DE TECHNOLOGIE BELFORT-MONTBÉLIARD

N° 2 | 0 | 5

THÈSE présentée par

Houssam SALMANE

pour obtenir le

Grade de Docteur de

l'Université de Technologie de Belfort-Montbéliard

Spécialité : **Informatique**

Reconstruction et analyse de trajectoires 2D d'objets mobiles par modélisation markovienne et par la théorie de l'évidence à partir de séquences d'images monoculaires - Application à l'évaluation du danger aux passages à niveau

Thèse préparée à IRTES-SET et LEOST-IFSTTAR, Projet ANR PANSAFER

Soutenue le 09 Juillet 2013 devant le Jury :

Michel DEVY	Rapporteur	LAAS-CNRS (Toulouse)
Olivier COLOT	Rapporteur	LAGIS (Lille)
Jacques JACOT	Président	LPM (Lausanne)
Alain CROUZIL	Examineur	irit (Toulouse)
Olivier STRAUSS	Examineur	Lirmm (Montpellier)
Dominique BERTRAND	invité	CERTU (Lyon)
Yassine RUICHEK	Directeur de thèse	UTBM (Belfort)
Louahdi KHOUDOUR	Co-Directeur de thèse	CETE (Toulouse)

Remerciements

Je tiens dans un premier temps à remercier le directeur de cette thèse, Monsieur Yassine Ruichek, Professeur à l'UTBM, pour m'avoir confié ce travail de recherches et pour m'avoir guidé, ainsi que pour son aide et ses précieux conseils au cours de ces années. Je remercie également Monsieur Louahdi Khoudour, Chargé de Recherches à IFSTTAR (CETE actuellement) et co-encadrant de ce travail de thèse, pour sa sympathie, sa disponibilité, ses idées et conseils, ainsi que pour son aide précieuse de tous les jours. Travailler avec Monsieur Yassine et Monsieur Louahdi est un réel plaisir, d'une part pour les discussions animées tant sur le plan scientifique que sur le plan personnel et d'autre part pour le respect pour la vie et le rythme de chacun qui permet d'avoir un environnement de travail plaisant et convivial. Merci

Je tiens à remercier vivement les différents membres de mon jury. Je remercie Jacques Jacot, professeur au LPM en suisse, de m'avoir fait l'honneur de présider mon jury. Je remercie Michel Devy, directeur de recherche CNRS à LAAS Toulouse, et Olivier Colot, Professeur des universités au laboratoire LAGIS Lille, d'avoir pris le temps de lire ce manuscrit et d'y avoir apporté des remarques pertinentes (ainsi que de nombreuses questions durant la soutenance) sur les travaux présentés. Messieurs Alberto Beswick et Jacques Meyer m'ont fait l'honneur de participer au Jury de soutenance ; je les en remercie profondément. Messieurs Alain CROUZIL, Maître de conférences à IRIT Toulouse, et Dominique Bertrand, Chargé d'études à CERTU Lyon, m'ont fait l'honneur de participer au Jury de soutenance ; je les en remercie profondément. Je remercie aussi tout particulièrement Olivier Strauss, Maître de conférences HDR au Lirmm Montpellier et mon ancien prof. de Master, pour sa participation à ce jury ainsi qu'aux différentes discussions toujours intéressantes que nous avons eu.

Je tiens également à remercier tous les membres du projet PANsafer qui ont permis de rendre ces années très agréables. Je n'oublie pas : Jean-Luc Bruyelle, Amaury Flancquart et Marc Heddebaut. Ils ont toujours fait preuve de gentillesse et de compétence. Je tiens aussi à remercier toutes les personnes rencontrées dans les conférences, les doctorants et les membres de l'équipe PENA (Frederic, Cyril, Cindy), qui ont permis de rendre ces années de thèse très enrichissantes. Mes remerciements vont également à M.Nizar Fakhfakh, pour la gentillesse, pour tous les conseils et les programmes qu'il a bien m'envoyer.

Je souhaiterais adresser des remerciements tout particuliers à M. Mazen ghandour, mon ancien professeur à l'université Libanaise, pour ces encouragements pendant toutes ces années.

Je terminerais ces remerciements par toutes les personnes et amis qui m'ont accompagné durant l'élaboration et surtout la rédaction de cette thèse. Enfin, bien sûr, ces remerciements ne seraient pas complets sans mentionner mes parents, sans eux, toute cette belle aventure n'aurais pu naître ni aboutir, je vous embrasse très fort.

Résumé : Les travaux présentés dans ce mémoire s'inscrivent dans le cadre du projet PANsafer (Vers un Passage A Niveau plus sûr), lauréat de l'appel ANR-VTT 2008. Ce projet est labellisé par les deux pôles de compétitivité i-Trans et Véhicule du Futur. Le travail de la thèse est mené conjointement par le laboratoire IRTES-SET de l'UTBM et le laboratoire LEOST de l'IFSTTAR.

L'objectif de cette thèse est de développer un système de perception permettant l'interprétation de scénarios dans l'environnement d'un passage à niveau. Il s'agit d'évaluer des situations potentiellement dangereuses par l'analyse spatio-temporelle des objets présents autour du passage à niveau.

Pour atteindre cet objectif, le travail est décomposé en trois étapes principales. La première étape est consacrée à la mise en place d'une architecture spatiale des capteurs vidéo permettant de couvrir de manière optimale l'environnement du passage à niveau. Cette étape est mise en oeuvre dans le cadre du développement d'un simulateur d'aide à la sécurité aux passages à niveau en utilisant un système de perception multi-vues. Dans ce cadre, nous avons proposé une méthode d'optimisation permettant de déterminer automatiquement la position et l'orientation des caméras par rapport à l'environnement à percevoir.

La deuxième étape consiste à développer une méthode robuste de suivi d'objets en mouvement à partir d'une séquence d'images. Dans un premier temps, nous avons proposé une technique permettant la détection et la séparation des objets. Le processus de suivi est ensuite mis en oeuvre par le calcul et la rectification du flot optique grâce respectivement à un modèle gaussien et un modèle de filtre de Kalman.

La dernière étape est destinée à l'analyse des trajectoires 2D reconstruites par l'étape précédente pour l'interprétation de scénarios. Cette analyse commence par une modélisation markovienne des trajectoires 2D. Un système de décision à base de théorie de l'évidence est ensuite proposé pour l'évaluation de scénarios, après avoir modélisé les sources de danger.

L'approche proposée a été testée et évaluée avec des données issues de campagnes expérimentales effectuées sur site réel d'un passage à niveau mis à disposition par RFF.

Mots clés : passage à niveau, système de perception, situations dangereuses, sécurité, détection d'objets, suivi d'objets

Table des matières

1	Introduction	1
1.1	Estimation du danger autour d'un passage à niveau : contexte	1
1.1.1	Etat actuel de la situation en Europe	5
1.1.2	Etat actuel de la situation à l'international	7
1.1.3	Campagnes pour l'amélioration de la sécurité aux PN	8
1.1.3.1	ILCAD	8
1.1.3.2	Le Plan Bussereau	8
1.1.3.3	Projet européen Selcat (Safer European Level Crossing Appraisal and Technology) (2006-2008)	10
1.1.3.4	Projet PANsafer (2009-2012)	10
1.2	Problématique générale	11
1.3	Problématique de la thèse	12
1.4	Contribution	13
1.5	Plan de la thèse	13
2	Developpement d'un algorithme d'optimisation de positions de cameras dans l'espace	17
2.1	Introduction	17
2.2	Plateforme de réalité virtuelle pour l'étude des PN	19
2.2.1	Contexte et démarche	19
2.2.1.1	Contexte	19
2.2.1.2	Intérêt de la réalité virtuelle	20
2.2.1.3	Démarche générale	21
2.2.2	Description de l'application	21
2.2.2.1	Modélisation 3D de l'univers virtuel	22
2.2.2.2	Positionnement des capteurs vidéo	23
2.2.2.3	Edition de scenarii	23
2.2.2.4	Simulation de scenarii critiques	24
2.3	Problématique	24
2.3.1	Modélisation du champ de vue	24
2.3.2	Intersection de pyramides : Géométrie descriptive	26
2.3.3	Plans de projection	28
2.4	Processus d'optimisation	31
2.4.1	Optimisation relative au plan horizontal	31
2.4.2	Optimisation relative au plan vertical d'orientation	35
2.4.3	Optimisation relative au plan vertical de projection	36
2.5	Résultats expérimentaux	37
2.6	Conclusion	44

3	Suivi d'objets en mouvement à partir d'une séquence d'images	47
3.1	Introduction	47
3.2	État de l'art sur le suivi d'objets en mouvement	49
3.2.1	Détection des objets	49
3.2.1.1	Modèle à valeur médiane	49
3.2.1.2	Codebook	49
3.2.1.3	Distributions gaussiennes	50
3.2.1.4	Mélange de gaussiennes	50
3.2.1.5	Modèle de Markov	50
3.2.1.6	Analyse en composantes indépendantes (ACI)	51
3.2.2	Séparation des objets	51
3.2.3	Suivi des objets	52
3.2.4	Méthode proposée	53
3.3	Détection et séparation des objets	53
3.3.1	Introduction	53
3.3.2	Processus de cadrage pour la détection des objets fixes et mobiles	54
3.3.3	Méthode de séparation des objets	55
3.4	Suivi dense des objets	58
3.4.1	Introduction	58
3.4.2	Suivi d'objets à partir d'une séquence d'images monoculaires	58
3.4.2.1	Suivi des points d'intérêt de Harris	58
3.4.2.2	Propagation gaussienne du flot optique des points de Harris	60
3.4.2.2.1	Première proposition	60
3.4.2.2.2	Deuxième proposition	61
3.4.2.3	Rectification du flot optique par un modèle de filtre de Kalman itératif	61
3.4.2.3.1	Contrainte de similarité	61
3.4.2.3.2	Modèle de filtre de Kalman proposé	63
3.4.2.3.3	Adaptation des paramètres	65
3.4.2.4	Modèle global de suivi	66
3.4.2.5	Algorithme d'optimisation	67
3.5	Résultats expérimentaux	69
3.5.1	Expérimentation 1 : Suivi de piétons	69
3.5.1.1	Évaluation des résultats de suivi obtenus	71
3.5.2	Expérimentation 2 : Evaluation avec des courbes gaussiennes	75
3.5.2.1	1ère distribution proposée (en fonction du flot optique moyen des points de Harris)	76
3.5.2.2	2ème distribution proposée (en fonction du flot optique de point de Harris le plus proche)	77
3.5.3	Expérimentation 3 : Suivi d'un objet dans une séquence d'images simulées	78
3.6	Conclusion	83

4	Evaluation du danger aux passages à niveau	85
4.1	Introduction : Problématique de la sécurité aux passages à niveau . .	85
4.2	Prédiction des trajectoires idéales d'objets mobiles	87
4.2.1	Introduction	87
4.2.2	Segmentation des objets à base du flot optique	87
4.2.3	Estimation des trajectoires idéales	88
4.2.3.1	Modèle de Markov Caché : définition des paramètres	88
4.2.3.2	Prédiction des trajectoires idéales	89
4.3	Estimation du danger de situations potentiellement dangereuses . .	95
4.3.1	Introduction	95
4.3.2	Théorie de l'évidence : définition des paramètres	95
4.3.2.1	Cadre de discernement	95
4.3.2.2	Jeux de masses	95
4.3.2.3	Fonction de crédibilité	96
4.3.2.4	Combinaison de sources de danger (Loi de combinaison de Dempster)	96
4.3.3	Modèle d'évaluation des trajectoires idéales	98
4.3.4	Définition des sources potentielles de danger	98
4.3.5	Condition d'indépendance des sources de danger	103
4.4	Résultats expérimentaux	103
4.4.1	Scénario 1 : Détection d'obstacles	104
4.4.2	Scénario 2 : Remontée de file	105
4.4.3	Scénario 3 : Passage en chicane, sans arrêt sur platelage . . .	109
4.4.4	Scénario 4 : Passage en chicane, avec arrêt sur platelage . . .	112
4.4.5	Scénario 5 : chute d'un piéton	115
4.5	Conclusion	121
5	Conclusion générale	125
A	Annexe	129
A.1	Algorithme de Levenberg-Marquardt	129
B	Annexe	131
B.1	Calcul de la fonction $d_c^{t_f}$ liée à la masse de croyance $m_p^A(D)$ de la source de danger relatif à la position du centre de la région considérée (voir chapitre 4) :	131
C	Liste des publications	133
C.1	Conférences internationales	133
C.2	Conférences nationales	134
C.3	Journal	134
	Bibliographie	135

Table des figures

1.1	Causes principales des accidents routiers.	2
1.2	Evolution du nombre de passages à niveau depuis la création de la SNCF.	3
1.3	Répartition des accidents mortels par types d’usagers routiers [Vic 2012].	4
1.4	Répartition du nombre moyen de collisions moyen par type de PN entre 1990-2010.	5
1.5	Evolution du nombre d’accidents et de victimes.	6
1.6	Les victimes de chemin de fer en Europe par catégorie (2008-2010) [Agency 2012].	6
1.7	Nombre de tués aux passages à niveau dans les pays de l’union européenne (2008-2010) [Agency 2012].	7
1.8	liste des 42 pays participant à la communauté ILCAD [UIC 2012].	9
1.9	Les pays impliqués dans le projet Selcat.	11
1.10	Schéma matériel du système de prise de vue.	14
2.1	Etapes pour l’utilisation du simulateur.	22
2.2	Contraintes d’optimisation (champ commun / total).	25
2.3	Représentation pyramidale du champ de vue d’une caméra.	27
2.4	Intersection entre deux faces latérales de pyramides.	29
2.5	Représentation du champ de vue commun de deux caméras (forme d’un polyèdre convexe).	30
2.6	Plan horizontal.	30
2.7	Plan vertical d’orientation.	32
2.8	Plan vertical de projection.	32
2.9	Etapes du processus d’optimisation.	34
2.10	Optimisation relative au plan horizontal.	35
2.11	Optimisation relative au plan vertical d’orientation.	36
2.12	Optimisation relative au plan horizontal et au plan vertical de projection.	37
2.13	Dimensions de la scène à observer.	39
2.14	Position initiale des caméras par rapport à l’environnement à surveiller.	39
2.15	Phase d’optimisation par rapport au plan horizontal : centrage et optimisation de la surface horizontale de la base de vue d’une caméra.	40
2.16	Centrage de la base de vue.	41
2.17	Evolution des surfaces durant l’optimisation relative au plan horizontal.	41
2.18	Phase d’optimisation par rapport au plan vertical d’orientation : orientation du champ de vue d’une caméra.	42
2.19	Phase d’optimisation par rapport au plan horizontal et au plan vertical de projection : optimisation des surfaces horizontales et verticales (totales et communes) des champs de vue des deux caméras.	43

2.20	Evolution des surfaces durant le processus d'optimisation relative au plan d'orientation.	43
2.21	Evolution des surfaces durant le processus d'optimisation relative au plan horizontal et au plan vertical de projection.	44
3.1	Détection des pixels affectés par du mouvement.	54
3.2	Extraction de cibles entre deux images successives.	55
3.3	Résultats de l'algorithme de séparation d'objets.	57
3.4	Séparation multi-objets.	57
3.5	Méthodologie de suivi d'objets.	59
3.6	Suivi des points de Harris.	60
3.7	Zone de recherche gaussienne (1 ^{ère} proposition).	62
3.8	Zone de recherche gaussienne (2 ^{ème} proposition).	63
3.9	Filtre de Kalman adaptatif.	65
3.10	Synoptique du suivi des objets.	67
3.11	Résultats de suivi avec un objet en mouvement : (a) méthode de Lucas-Kanade (b) méthode proposée.	70
3.12	Evolution du nombre de pixels suivis par notre méthode et celle de Lucas-Kanade.	71
3.13	Résultats de suivi multi-objets par la méthode proposée.	72
3.14	répartition de l'erreur d'intensité.	73
3.15	Représentation de l'erreur d'intensité par la méthode proposée.	73
3.16	Représentation de l'erreur d'intensité par la méthode de Lucas-Kanade.	74
3.17	Courbes médianes obtenues par les différentes méthodes.	75
3.18	Distribution initiale du flot optique.	76
3.19	Distribution du flot optique après application de la méthode proposée (avec la 1 ^{ère} distribution gaussienne).	77
3.20	Distribution du flot optique après application de la méthode proposée (avec la deuxième distribution gaussienne).	79
3.21	Distribution du flot optique (avec la 2 ^{ème} distribution gaussienne), en utilisant la moyenne et la variance du flot optique des points de Harris.	80
3.22	Séquence d'images simulées.	81
3.23	Résultats du suivi de la grille.	82
3.24	Erreur moyenne de position des pixels suivis.	82
3.25	distribution de l'erreur de position des pixels.	83
4.1	Synoptique de l'approche d'évaluation du danger.	88
4.2	Segmentation d'un objet en mouvement par clustering du flot optique.	89
4.3	Paramètres liés à l'environnement du passage à niveau nécessaires à l'estimation d'une trajectoire idéale.	90
4.4	Exemple de modèle de Markov pour l'estimation d'une trajectoire.	91
4.5	Illustration de l'estimation de la trajectoire idéale à partir d'une trajectoire idéale courante.	92

4.6	Exemples de trajectoires idéales prédites.	94
4.7	Représentation graphique de la fusion de deux sources de danger distinctes selon la loi de combinaison de Dempster.	97
4.8	Analyse de la trajectoire idéale prédite.	99
4.9	Installation du système de perception sur site (Mouzon-France). . . .	104
4.10	Détection d'obstacles - DV : danger lié au véhicule; DP : danger lié au piéton. Sens de lecture de la séquence : haut-bas, gauche-droite. .	107
4.11	Détection d'obstacles - Evolution du danger lié aux différents objets au cours du temps.	108
4.12	Remontée de file 1 - DV _i : danger liée au véhicule V _i ; DP : danger lié au piéton. Sens de lecture de la séquence : haut-bas, gauche-droite.	111
4.13	Remontée de file 1 - Evolution du danger lié aux différents objets. .	112
4.14	Deuxième scénario de remontée de file. Sens de lecture de la sé- quence : haut-bas, gauche-droite.	114
4.15	Remontée de file 2 - Evolution du danger lié aux différents objets. .	115
4.16	Passage en chicane sans arrêt sur platelage. Sens de lecture de la séquence : haut-bas, gauche-droite.	117
4.17	Passage en chicane sans arrêt sur platelage - Evolution du danger lié au véhicule.	118
4.18	Passage en chicane avec arrêt sur platelage. Sens de lecture de la séquence : haut-bas, gauche-droite.	120
4.19	Passage en chicane avec arrêt sur platelage - Evolution du danger lié au véhicule.	121
4.20	Chute d'un piéton sur platelage. Sens de lecture de la séquence : haut-bas, gauche-droite.	123

Introduction

Sommaire

1.1	Estimation du danger autour d'un passage à niveau : contexte	1
1.1.1	Etat actuel de la situation en Europe	5
1.1.2	Etat actuel de la situation à l'international	7
1.1.3	Campagnes pour l'amélioration de la sécurité aux PN	8
1.2	Problématique générale	11
1.3	Problématique de la thèse	12
1.4	Contribution	13
1.5	Plan de la thèse	13

1.1 Estimation du danger autour d'un passage à niveau : contexte

Le transport ferroviaire est probablement l'un des modes les plus efficaces et les plus écologiques. Pourtant les passages à niveau, qui représentent le croisement entre une voie ferrée avec une voie routière, constituent un environnement très dangereux pour les usagers. La solution la plus sûre à ce problème serait de remplacer tous les passages à niveau par des ponts ou des souterrains. Cependant, cette solution est très coûteuse (5 à 10 millions d'Euros par an en Europe) et elle n'est pas toujours possible à réaliser en particulier dans les zones très urbanisées.

Normalement, le franchissement des passages à niveau peut être réalisé en toute sécurité en cas de respect de la signalisation routière définie à partir de la "Convention sur la signalisation routière" signée à Vienne en 1968 et appliquée dans le monde entier [Cousin 1981]. Malheureusement, les statistiques montrent que les usagers de la route ou les piétons ne respectent pas la signalisation routière, ce qui occasionne près de 98% des accidents aux passages à niveau [Min 2011]. En réalité, des milliers de presque-accidents ou incidents s'y produisent chaque jour dans le monde entier. Le risque d'accident augmente aussi en cas d'implication d'autocars ou de véhicules lourds et peut entraîner des conséquences catastrophiques [Bea 2012]. Selon l'agence ferroviaire européenne (ERA) [Agency 2012], 619 accidents ont tué 359 personnes dans l'Union européenne en 2010 (alors qu'en 2009, 831 accidents avaient fait 405 morts). Ce chiffre représente 28% de l'ensemble des décès par accidents dans les chemins de fer et 1,2% des morts sur la route. En plus des coûts médicaux élevés,

les accidents mortels aux passages à niveau laissent des effets très négatifs à long terme sur la conscience personnelle des passagers de la voiture ou du train, sur le conducteur du train, sur les témoins et les personnes blessées.

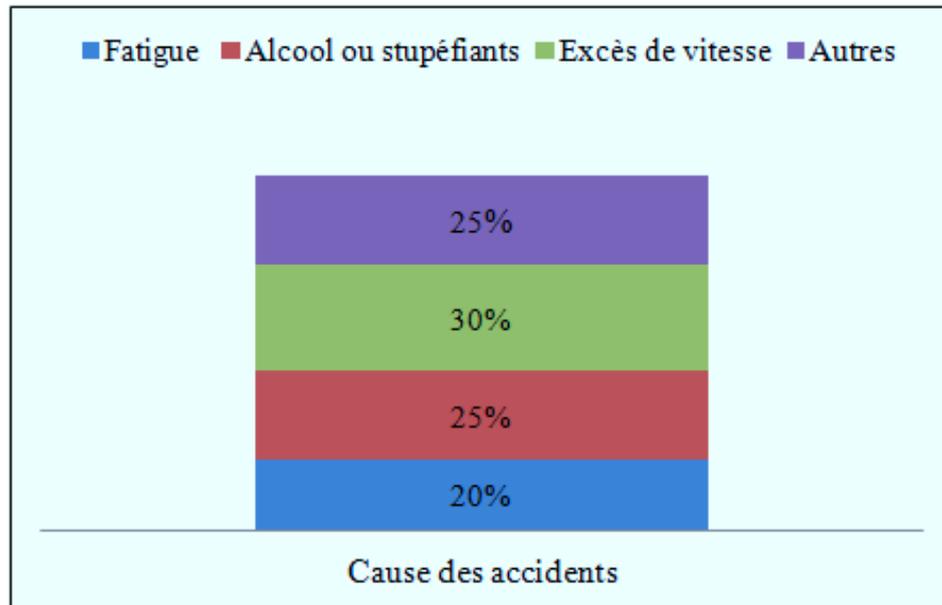


FIGURE 1.1 – Causes principales des accidents routiers.

Plusieurs raisons expliquent la prise de risque au franchissement des passages à niveau : la pression du temps, l'inattention, des raisons physiologiques (par exemple la fatigue, l'alcool). Selon les chiffres de la commission européenne [TREN 2003], 20% des accidents de la route sont dus à la fatigue et 25% à une conduite sous l'influence de l'alcool ou de stupéfiants. L'excès de vitesse constitue le plus fort facteur de risque de collision aux passages à niveau. Selon les chiffres de la commission européenne, 30% des accidents routiers mortels survenus dans l'UE sont dus à un excès de vitesse (Fig. 1.1). Malgré la diversité des solutions techniques proposées [Silmon 2010][Tey 2009], telles que les systèmes d'alerte, les systèmes de navigation par satellite et l'émission de signaux d'alerte transmis aux véhicules à l'approche des trains aux passages à niveau (systèmes en cours d'essai en France et en Australie), le contrôle ou bien la prédiction du comportement humain reste le problème le plus difficile à résoudre [Tey 2012] d'où l'importance de l'organisation de journées internationales pour la sécurité aux passages à niveau. Le but de ces journées est la sensibilisation des usagers quant au respect de la signalisation routière et au danger que peut représenter le franchissement illicite d'un passage à niveau (PN).

Malgré les efforts menés par les gouvernements pour sécuriser ces endroits et sensibiliser les usagers, le nombre d'accidents aux passages à niveau reste toujours trop élevé, même s'il baisse progressivement.

Malgré les efforts importants menés par le gouvernement français, la S.N.C.F.

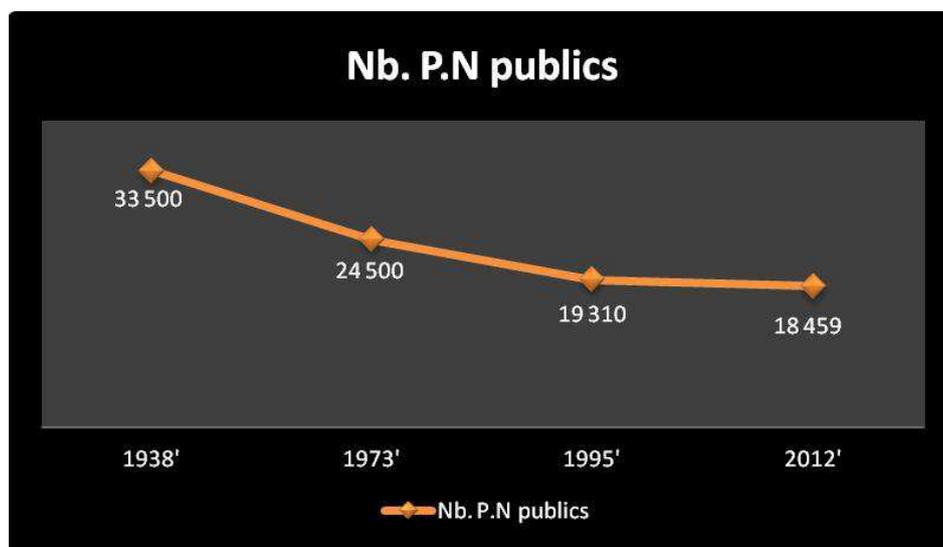


FIGURE 1.2 – Evolution du nombre de passages à niveau depuis la création de la SNCF.

Type de P.N.	Automatiques			Non gardés	Gardés	Piétons	Privés	Nb. total de P.N. publics
	S.A.L. 0	S.A.L. 2	S.A.L. 4					
Nombre	10 880			3 864	1 767	840	1 108	18 459
S.A.L. = Signalisation Automatique Lumineuse et sonore complétée par 0 (S.A.L. 0), 2 (S.A.L. 2) ou 4 (S.A.L. 4) demi-barrières								

TABLE 1.1 – Répartition des passages à niveau par type.

et les collectivités territoriales dans le cadre d'une politique de suppression des PN, de nombreux points de conflits entre circulation routière et ferroviaire demeurent.

En 1938, année de création de la S.N.C.F., il existait 33 500 PN publics dont 27 000 étaient gardés. En 1973, le nombre de passages à niveau a été réduit à 24 500. Aujourd'hui, la situation s'est profondément modifiée. Le nombre de passages à niveau publics a été ramené à 19 310 en 1995, suite essentiellement aux importants programmes de suppression de PN, concernant en priorité les lignes ferroviaires à grandes vitesses, et à la fermeture des lignes de faible importance. Aujourd'hui, il reste 18 459 passages à niveau (Fig. 1.2). Parmi eux, il existe 17 351 PN ouverts à la circulation publique dont 840 (4,85 %) sont ouverts au seul cheminement des piétons et 16 511 PN (95,15 %) sont ouverts à la circulation des véhicules et des piétons (Table 1.1). 63 % des PN se situent sur des voies communales, 36 % sur des voies départementales et il ne reste que 67 PN sur des routes nationales (1 %), mais ces PN disparaissent progressivement au cours du temps.

Les dernières statistiques communiquées [SET 1996][Got 2012], montrent que la plupart des victimes des accidents aux passages à niveau font partie des habitants qui résident à proximité des passages à niveau, notamment ceux qui sont habitués

à traverser ces zones. Il semble que l'utilisation quotidienne d'un passage à niveau contribue à la baisse de vigilance de ces usagers quotidiens. La figure 1.3 illustre la répartition des accidents mortels par types d'usagers routiers. On note que 56% des accidents impliquent des véhicules légers.

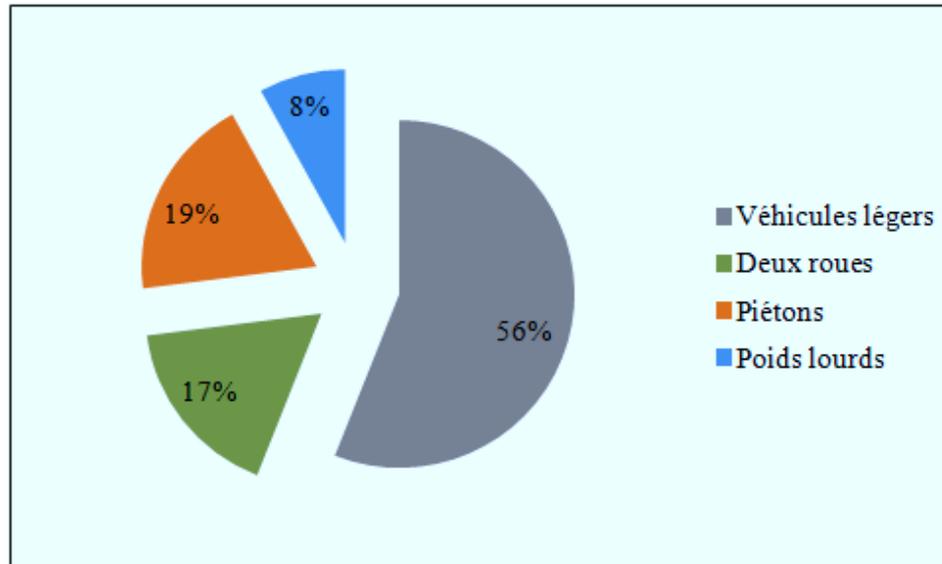


FIGURE 1.3 – Répartition des accidents mortels par types d'usagers routiers [Vic 2012].

Durant les vingt dernières années, 88 % des accidents sont répartis sur les passages à niveau à signalisation automatiques (S.A.L. 2, S.A.L. 4, S.A.L. 0) [Shigeru 1988] et les passages à niveau non gardés. Les passages à niveau à signalisation automatique Lumineuse (S.A.L), représentant 59 % du nombre total de PN et sont concernés par 70 % des collisions usager routier/train. Les PN de type S.A.L.2 sont responsables de 70 % de collisions qui se produisent sur les passages à niveau automatiques. Les causes de collisions sur ce type de passage à niveau (S.A.L.2) se caractérisent principalement par des passages en chicane et par des passages avec des vitesses routières trop élevées à l'approche du PN. En revanche, la cause principale des collisions sur les passages à niveau automatiques de type S.A.L.4 (généralement situés sur les routes à fort trafic) est l'arrêt des véhicules sur le PN ou la présence de remontées de files. Les PN non gardés (représentant 21 % du nombre total de PN), généralement équipés de croix de Saint André, avec ou sans STOP, sont concernés par 18 % des accidents. Les passages à niveau piétons sont concernés par 10 % des accidents. La principale cause des collisions sur ces deux derniers types de PN est le non-respect de la signalisation sur la route. Les passages à niveau gardés subissent une politique de suppression ces dernières années (transformation en passages supérieurs ou inférieurs. . .) et sont en train de disparaître. Le nombre d'accidents sur ces PN est presque nul ces 10 dernières années. La figure 1.4 présente la répartition

du nombre de collisions par type de PN.

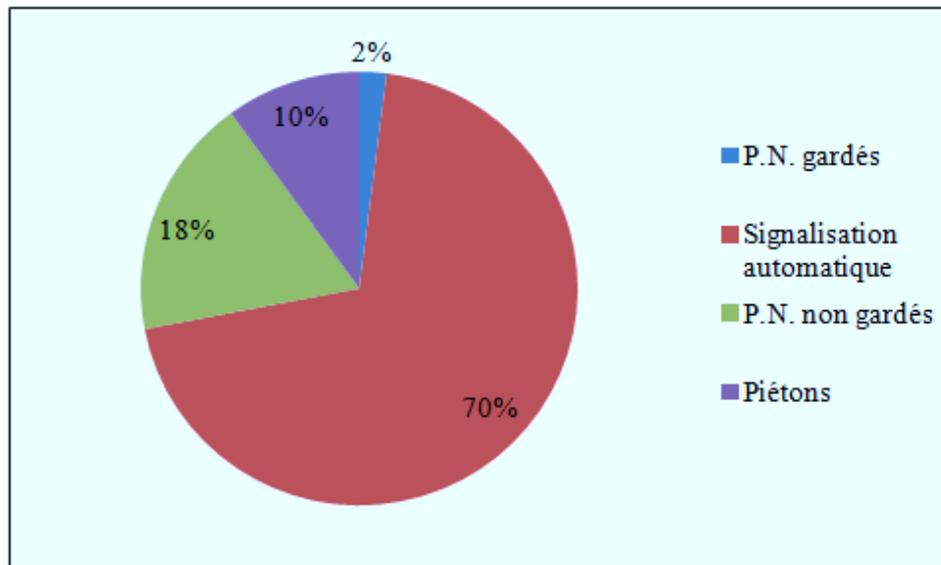


FIGURE 1.4 – Répartition du nombre moyen de collisions moyen par type de PN entre 1990-2010.

Durant la période 1991-1995, on compte une moyenne annuelle de 210 accidents de type collision entre un usager routier et un train, causant en moyenne 61 tués et 26 blessés graves par an. Malgré ces chiffres, on note heureusement un ralentissement du nombre de décès aux passages à niveau depuis une dizaine d'années. Depuis l'année 2000, les points de croisement entre la route et la voie ferrée sont à l'origine en moyenne de 115 collisions par an (une collision tous les trois jours), 14 blessés graves et 40 décès. Ces accidents sont en grande majorité la conséquence d'infractions au code de la route. 50% des collisions entre une voiture particulière et un train causent un mort alors que 5% des "Accidents routiers typiques" sont dans ce cas. La figure 1.5 illustre l'évolution du nombre d'accidents et de victimes durant les 40 dernières années.

1.1.1 Etat actuel de la situation en Europe

Sur la période des 20 dernière années (à partir de 1990), il a été enregistré en Europe 1100 accidents et 400 tués par an en moyenne. Pendant les 4 dernières années, la tendance est à la baisse des collisions et des tués aux passages à niveau (619 accidents et 359 morts en 2010). 28 % des décès sur le réseau ferroviaire européen sont des usagers de PN (Fig. 1.6). 90% des accidents mortels sont dues à une infraction de code de la route. Les campagnes de sensibilisation à la sécurité ferroviaire demeurent incapables à contrôler le mauvais comportement humain et le non-respect des règles de sécurité (60 % des victimes ferroviaires sont des personnes non autorisées, telles que des piétons qui ont traversé les voies malgré l'interdiction).

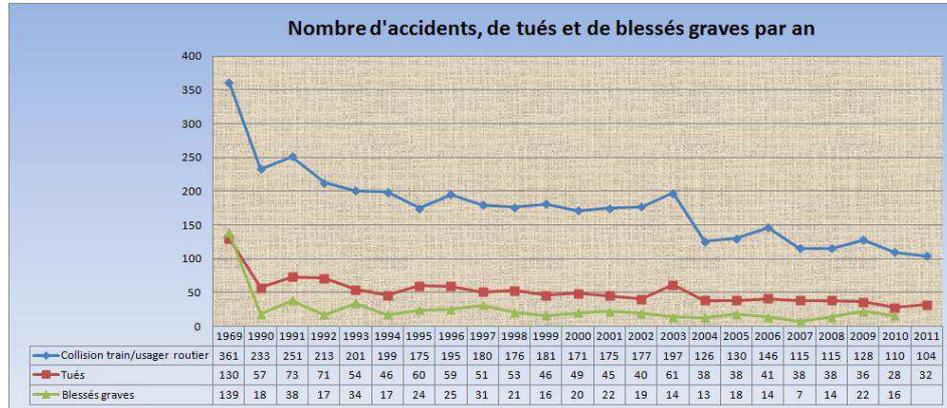


FIGURE 1.5 – Evolution du nombre d'accidents et de victimes.

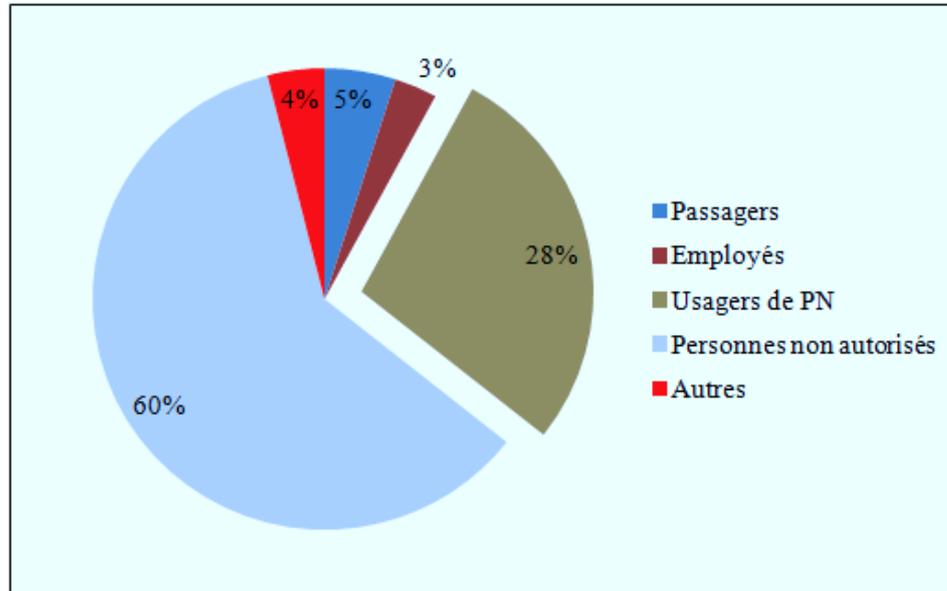


FIGURE 1.6 – Les victimes de chemin de fer en Europe par catégorie (2008-2010) [Agency 2012].

	France	Australie	Belgique	Estonie	Pologne	Royaume-Uni	Etat Unis
Nb. de P.N.	18500	23 500	1902	336	12 977	6452	217 269
Nb. De collisions	110	78	49	17	257	7	2017
Nb. de tués	28	38	9	2	29	4	256

TABLE 1.2 – Répartition des collisions et du nombre de tués par pays pour l'année 2010.

Parmi les 26 pays de l'union européenne qui ont participé à une évaluation statistique de la sécurité aux passages à niveau [ERADIS][Agency 2012], la France est classée 8ième dans cette liste (Fig. 1.7) avec un nombre de décès par 1 million de kilomètres de ligne-train égal à 0.06. Ce chiffre dépasse 0.4 dans quelques pays comme la Roumanie et la Lituanie. La figure 1.7 montre que la France est considérée parmi les pays européens actifs dans la sécurité aux passages à niveau.

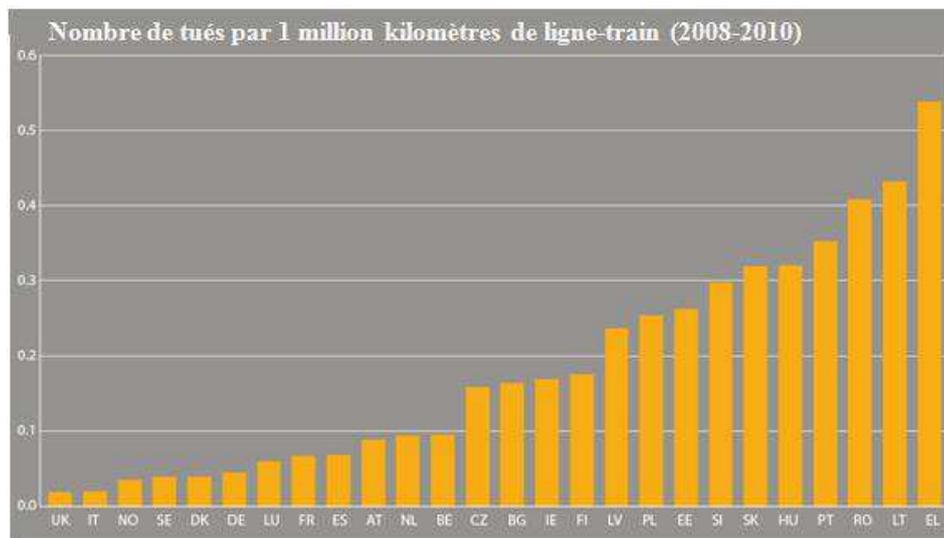


FIGURE 1.7 – Nombre de tués aux passages à niveau dans les pays de l'union européenne (2008-2010) [Agency 2012].

1.1.2 Etat actuel de la situation à l'international

Il existe 1 million de passages à niveau dans le monde dont seulement 123 000 passages en Europe (contre plus de 217 000 PN aux Etats Unis) [Safety][UIC]. Ceci montre que grâce aux programmes de sécurité aux passages à niveau, le problème de la sécurité aux PN en Europe reste mieux contrôlé par rapport à l'échelle mondiale. En 2010, parmi les 6100 morts aux passages à niveau dans le monde, 359 sont européens contre 256 qui sont américains (Table 1.2).

1.1.3 Campagnes pour l'amélioration de la sécurité aux PN

Les données ci-dessus plaident clairement pour une amélioration marquée de la sécurité aux PN. Cette amélioration passe en partie par la conception de systèmes d'aide à la traversée d'un passage à niveau. Ces dernières années, on a vu apparaître plusieurs campagnes et modifications de législation et projets européens sur le thème de la sécurité aux PN, cherchant à trouver des solutions efficaces contre ce fléau. Le paragraphe qui suit fournit un descriptif des initiatives récentes dans le domaine de la sécurité aux PN.

1.1.3.1 ILCAD

Créée en 2009 ILCAD " International Level Crossing Awareness Day " [[Safety](#)] est une communauté qui organise chaque année une journée avec des experts dans le domaine de la sécurité routière aux passages à niveau. Le but de cette journée est de partager les idées et les expériences pour améliorer la sécurité et approfondir la connaissance sur les accidents aux passages à niveau et diffuser aux usagers routiers les règles du franchissement des passages à niveau. Cette journée est diffusée en direct (sur le web de ILCAD) chaque année dans le monde entier. La communauté organise aussi une série d'événements centrés sur le message commun de l'ILCAD : "Aux passages à niveau, priorité à la vie!". 42 pays participent à ILCAD ainsi que des organisations ferroviaires européennes et internationales (UIC, CER, EIM), Operation Lifesaver, le Conseil européen de la sécurité des transports (ETSC), le Forum européen sur les passages à niveau (ELCF), l'Agence ferroviaire européenne (ERA) et la Division Transports de la CEE-ONU. Y participent également l'Association des Chemins de fer d'Amérique latine (ALAF), l'Association australienne des Chemins de fer (ARA) et l'Associations des Chemins de fer américains (AAR). La figure 1.8 dresse la carte des pays participant à la communauté ILCAD.

1.1.3.2 Le Plan Bussereau

Suite à la catastrophe ferroviaire survenue le 2 juin 2008 [[BUSSEREAU 2009](#)] entre un train et un véhicule scolaire au passage à niveau d'Allinges (Haute-Savoie en France) et qui a entraîné la mort de 7 enfants, le gouvernement français par la voix de son premier ministre a demandé à Dominique Bussereau, secrétaire d'Etat aux transports de proposer un plan pour améliorer la sécurité aux passages à niveau. Le plan Bussereau comprend deux objectifs principaux : - 1) renforcer la sensibilisation des conducteurs à respecter le code de la route. - 2) aider les campagnes et les communautés françaises à améliorer la signalisation aux passages à niveau. Pour servir cette politique, un plan comprenant 20 mesures (dont 16 mesures ont été exécutées et 4 mesures sont en cours de mise en oeuvre) a été proposé. Nous vous présentons quelques mesures importantes de ce plan.

Mesure n°1 : Informer les usagers de la route des règles à respecter à l'approche d'un passage à niveau (organiser des actions et des campagnes de sensibilisation. . .).



FIGURE 1.8 – liste des 42 pays participant à la communauté ILCAD [UIC 2012].

Mesure n°4 : Sanctionner les infractions les plus graves commises par les poids lourds.

Mesure n°7 : Implanter des radars de contrôle de vitesse.

Mesures n°8 et 9 : Mise en oeuvre et analyse de la sécurité pour tous les passages à niveau.

Mesure n°15 : Traiter (supprimer ou bien aménager) dans les 10 ans tous les passages à niveau préoccupants (320 PN sont concernés).

Mesure n°18 : Rendre obligatoire les barrières sur tous les passages à niveau où la vitesse des trains est supérieure à 90 km/h.

Ce programme comprend aussi une convention ("Charte pour une politique de sécurité sur les passages à niveau") entre l'Etat, Réseau Ferré de France (RFF) et l'Assemblée des Départements de France (ADF) permettant de partager les travaux et les actions d'amélioration de la sécurité des passages à niveau en termes de responsabilité. Ces actions ont été partagées de la façon suivante : l'Etat sera responsable pour participer (avec RFF) au financement nécessaire au traitement de la sécurité des passages à niveau et faciliter les procédures administratives pour permettre la mise en oeuvre effective des mesures d'amélioration. RFF doit choisir et retenir (avec l'ADF et la SNCF) les techniques nécessaires pour l'amélioration de la sécurité des passages à niveau et installer sur les passages à niveau des panneaux ("Croix de Saint-André" ou bien "panneau Stop"), des feux clignotants et des barrières. En plus de son travail en commun avec le RFF et la SNCF (choix des techniques d'amélioration de la sécurité), l'ADF prend en charge le solde de l'opération.

1.1.3.3 Projet européen Selcat (Safer European Level Crossing Appraisal and Technology) (2006-2008)

Selcat [Slovak 2006][Schnieder 2008] est un projet européen de type PCRD qui s'est déroulé entre 2006 et 2008. Le but principal de ce projet est de contribuer d'une manière effective et active à la réduction des accidents aux passages à niveau. 25 partenaires de 14 pays différents provenant de l'Europe, Asie et Afrique ont participé à ce projet (Fig. 1.9).

Les principales activités et objectifs de Selcat sont les suivants :

- Permettre l'échange d'expertise et la contribution effective entre plusieurs pays (partenaires) pour la réduction des accidents aux passages à niveau.
- Partager et échanger la connaissance de la sécurité des passages à niveau.
- Analyser et discuter les solutions existantes avec leurs résultats pour améliorer le niveau de sécurité.
- Faire un état de l'art sur les recherches existantes sur le thème.
- Explorer des nouvelles technologies puis essayer de les évaluer et les optimiser.

1.1.3.4 Projet PANsafer (2009-2012)

PANsafer "Vers un passage à niveau plus sûr" [Khoudour 2010], dans lequel cette thèse est inscrite, est un des projets soutenus par l'ANR dans le cadre du



FIGURE 1.9 – Les pays impliqués dans le projet Selcat.

programme Véhicules pour les Transports Terrestres 2008. Ce projet est labellisé par les pôles de compétitivité I-Trans et Véhicules du Futur. Les objectifs de PANsafer consistent à :

- Mettre en évidence les principaux facteurs d'accidents sur les passages à niveau (PN).
- Analyser les comportements induits par l'infrastructure et ses modalités d'exploitation.
- Identifier, reconnaître et évaluer des situations dangereuses ou potentiellement dangereuses aux passages à niveau.
- Rechercher de nouvelles solutions techniques pour renforcer le niveau de sécurité des PN.

Coordonné par l'IFSTTAR, les principaux contributeurs au projet PANsafer sont : le Centre d'Etudes sur les Réseaux, les Transports, l'Urbanisme et les constructions publiques (CERTU), l'Ecole Centrale de Lille (EC-Lille), l'Ecole Polytechnique Fédérale de Lausanne (EPFL), la PME INFODIO, Réseaux Ferrés de France (RFF), les Transports Lausannois (TL) et l'Université de Technologie de Belfort Montbéliard (UTBM).

1.2 Problématique générale

Même si les statistiques montrent que le problème d'accidents aux passages à niveau est un problème mineur par rapport aux accidents de la route et, malgré le nombre limité de victimes aux passages à niveau et sa réduction constante durant les dix dernières années, la sécurisation des passages à niveau reste un sujet important pour l'opinion publique à cause des dramatiques accidents qui y surviennent

(accident d'Allinges par exemple). C'est dans cet esprit que les gouvernements (plan Bussereau), les campagnes ferroviaires (RFF, SNCF, ADF) et les campagnes internationales (ILCAD) réagissent pour améliorer la sécurité des usagers de la route à l'approche des passages à niveaux (Projet Selcat, PANsafer...). La plupart des solutions qui sont étudiées et exécutées par les campagnes cherchent d'une part, à améliorer la sensibilisation des acteurs de la route et du fer et, d'autre part, à aménager les infrastructures des passages à niveaux (ajouter des panneaux, des feux clignotants, barrières automatiques...) ou bien à supprimer des passages à niveaux (qui reste la solution la plus efficace mais la plus coûteuse). Pour le moment, les solutions technologiques semblent être compliquées à mettre en oeuvre et présentent un coût élevé. Cependant, les dernières études scientifiques ont montré que l'utilisation des systèmes de perception (comme des caméras de surveillance intelligentes) peuvent être une solution très efficace dans le domaine de la sécurité routière. D'une part l'utilisation des caméras intelligentes n'est pas trop chère par rapport aux autres solutions proposées et, d'autre part, ce type de systèmes permet de détecter et de reconnaître rapidement des scénarios accidentogènes.

1.3 Problématique de la thèse

Si la plupart des campagnes ferroviaires et des études de sécurité aux passages à niveau se focalisent sur le développement des systèmes d'incitation et de sensibilisation chez les acteurs route/fer ou bien sur les systèmes d'aménagement des infrastructures aux passages à niveau les plus préoccupants pendant longtemps, aujourd'hui un vrai besoin se concentre sur le développement de systèmes de contrôle permettant d'évaluer en temps réel le comportement humain et les situations critiques (chicane, remontée de fil, vitesse d'approche élevée...) à l'approche de passages à niveau. Avec la généralisation de l'utilisation des images numériques, l'analyse du mouvement des usagers routiers semble être une approche intéressante pour des applications diverses, notamment l'évaluation et le contrôle de la sécurité routière et ferroviaire. En effet, l'utilisation d'un système de perception visuelle pour la détection et le suivi d'objets évoluant dans l'environnement d'un passage à niveau permettrait d'analyser leur comportement, puis en extraire les situations potentiellement dangereuses. La problématique que cette thèse se propose de traiter est celle de la conception d'un système de perception permettant l'évolution de situations accidentogènes aux passages à niveau (passage en chicane, présence d'obstacles, remontée de file...).

Les aspects fondamentaux de ce travail de thèse sont :

- Une étude sur le calibrage des paramètres extrinsèques des capteurs vidéos vis-à-vis de la scène à surveiller. Il s'agit de déterminer la position et l'orientation des caméras optimisant les champs de vision (communs et totaux). C'est une étape très importante au regard de l'application envisagée.
- La détection d'objets par analyse d'images monoculaires couleur. L'approche développée repose sur deux étapes principales : détection des objets puis leur sépa-

ration.

- Le suivi d'objets basé sur le calcul du flot optique. Ce dernier est estimé par une propagation gaussienne puis rectifié à l'aide d'un filtre de Kalman.

- La détection, reconnaissance et évaluation des situations dangereuses ou potentiellement dangereuses autour d'un passage à niveau et établissement d'un niveau de criticité. Cette étape est mise en oeuvre par modélisation markovienne et utilisation de la théorie de l'évidence.

- L'analyse de scénarios d'accidents réels joués sur un passage à niveau.

1.4 Contribution

Les travaux présentés dans ce mémoire portent sur le développement d'un système de perception pour la détection, la reconnaissance et l'évaluation de situations potentiellement dangereuses autour de l'environnement d'un passage à niveau. Ce système repose sur l'analyse d'images acquises par une caméra fixée sur un support horizontal sur un mat télescopique de 6 mètres de haut, avec une vue plongeante sur la zone de passage à niveau. La figure 1.10 fournit un schéma du système de prise de vue avec deux caméras constituant un capteur stéréoscopique. Dans ce travail, seule une caméra est utilisée dans le cadre d'une analyse monoculaire.

L'objectif du projet est de développer un système de perception et d'analyse de l'environnement d'un passage à niveau. Notre contribution essentielle dans ce travail est de proposer une méthodologie d'évaluation du risque autour d'un passage surveillé en vision monoculaire. Dans le cadre du développement d'un simulateur d'aide à l'évaluation de la sécurité aux passages à niveau à base d'un système de perception multi-vues, nous avons proposé une méthode d'optimisation pour la mise en oeuvre de l'architecture spatiale des capteurs vidéo permettant de couvrir de manière optimale l'environnement d'un passage à niveau. Nous avons développé ensuite un algorithme robuste permettant la détection et le suivi d'objets statiques/dynamiques à travers l'analyse de séquences d'images monoculaires. A l'issue de cette étape, les paramètres caractérisant les objets retenus (position, vitesse, accélération, ...) sont analysés afin d'évaluer les situations dangereuses ou potentiellement dangereuses. Les tests expérimentaux ont été effectués sur des séquences d'images réelles prises lors de campagnes d'acquisition avec des scénarios de situations accidentogènes (passage en chicane, présence d'obstacles, remontée de file) joués sur un passage à niveau mis à disposition par RFF.

1.5 Plan de la thèse

Après l'introduction générale sur la problématique de la sécurité des passages à niveau en France et dans le monde, que nous avons présentée dans ce premier chapitre, le reste du manuscrit se décline de la manière suivante :

Le second chapitre décrit en détail la mise en place de l'architecture spatiale de capteurs vidéo (CCD) dans le cadre du développement d'un simulateur d'aide à

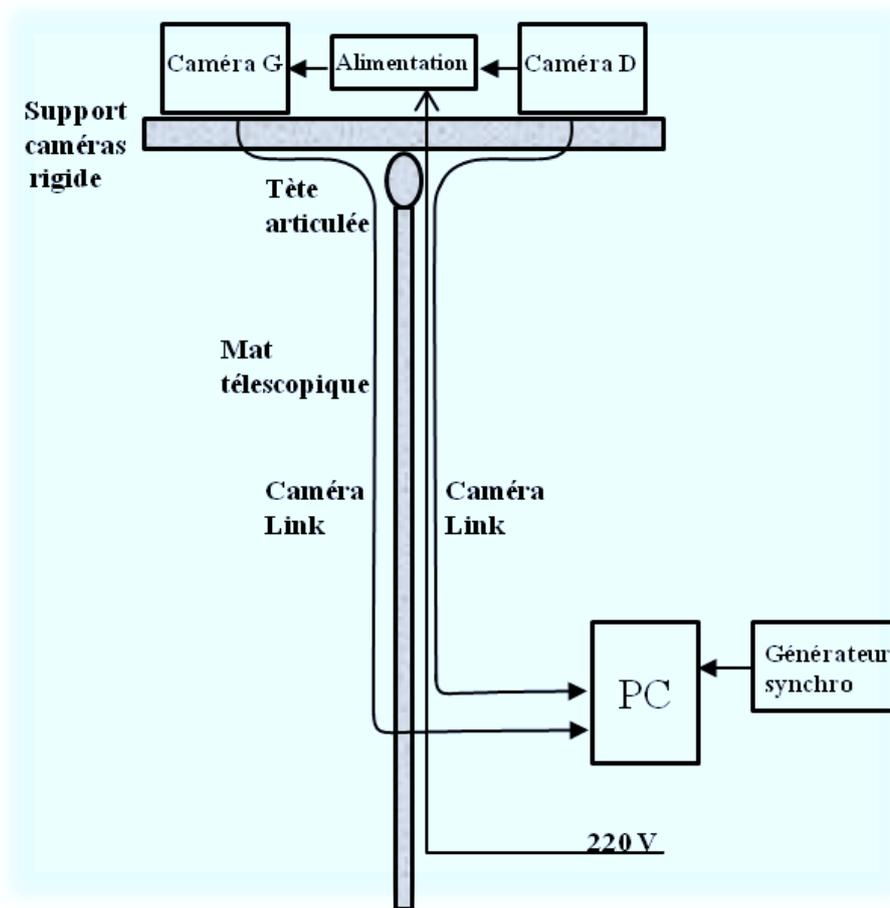


FIGURE 1.10 – Schéma matériel du système de prise de vue.

l'évaluation de la sécurité aux passages à niveau à base d'un système de perception multi-vues. Dans cette partie, nous proposons une méthode pour déterminer automatiquement la position et l'orientation géométriques des caméras pour couvrir de manière optimale l'environnement à percevoir. Cette méthode est basée sur la modélisation du champ de vue des capteurs vidéo dans l'espace afin de déterminer les paramètres et d'optimiser les champs commun et total de la zone à surveiller.

Le troisième chapitre comprend deux parties principales. La première consiste à détecter et séparer les objets fixes et en mouvement dans une image. Cette étape est basée sur deux étapes essentielles : 1) l'étape de cadrage pour la détection des objets et, 2) l'étape de comparaison des vecteurs d'énergies pour la séparation des objets. La deuxième partie de ce chapitre est consacrée au suivi des objets détectés dans des séquences d'images réelles (cadence : 5 images/seconde). Pour pouvoir suivre les objets détectés, nous proposons une technique basée sur le calcul du flot optique, estimé par propagation gaussienne à partir du flot optique des points de Harris, puis rectifié en utilisant un modèle de filtrage de Kalman.

Le quatrième chapitre détaille la technique proposée pour l'évaluation de la sécurité aux passages à niveau selon des scénarios prédéfinis et tournés en situation réelle sur un PN dans le cadre du projet PANsafer. Nous attaquons dans ce chapitre le problème de la détection et de la reconnaissance des scénarios d'accident autour d'un passage à niveau. Dans cette étape, nous avons développé une méthode de reconnaissance en utilisant une modélisation Markovienne et un modèle de calcul de risque basé sur la théorie de l'évidence.

Enfin, nous concluons et nous proposons nos perspectives pour l'amélioration de ces travaux.

Developpement d'un algorithme d'optimisation de positions de cameras dans l'espace

Sommaire

2.1	Introduction	17
2.2	Plateforme de réalité virtuelle pour l'étude des PN	19
2.2.1	Contexte et démarche	19
2.2.2	Description de l'application	21
2.3	Problématique	24
2.3.1	Modélisation du champ de vue	24
2.3.2	Intersection de pyramides : Géométrie descriptive	26
2.3.3	Plans de projection	28
2.4	Processus d'optimisation	31
2.4.1	Optimisation relative au plan horizontal	31
2.4.2	Optimisation relative au plan vertical d'orientation	35
2.4.3	Optimisation relative au plan vertical de projection	36
2.5	Résultats expérimentaux	37
2.6	Conclusion	44

2.1 Introduction

Dans le domaine de la vidéosurveillance [Chen 2005], les architectures à caméras multiples ont connu une très forte expansion ces dernières années [Kettner 1999][Xu 2005]. Non seulement les objets en mouvement doivent être segmentés et suivis, mais la position des caméras joue un rôle très important pour bien reconnaître un même objet filmé dans le champ du système multi-caméra [Vezzani 2010]. Par ailleurs la multiplication des caméras installées dans des espaces publics ou privés, rend de plus en plus difficile l'exploitation de leur champ de vue (surtout en stéréovision) par des opérateurs humains. Pour cela, la plupart de ces systèmes considèrent les caméras d'une manière indépendante les unes des autres.

Ajuster la position et l'orientation de plusieurs caméras est un problème crucial dans de nombreuses applications en vision par ordinateur. L'approche la plus

utilisée pour positionner et orienter les caméras (afin de couvrir l'environnement à surveiller) procède de manière expérimentale et empirique en jouant des configurations géométriques potentielles. Cette approche requiert généralement beaucoup de temps pour ajuster les champs de vue autour de la scène à observer en termes de champ de vue total et champ de vue commun aux caméras. Ceci est dû au fait que la solution optimale appartient à un domaine de variation relativement restreint. En effet, un petit déplacement en position et/ou orientation d'une caméra conduit à une large variation de son champ de vue. De plus, le processus d'ajustement de la position et l'orientation des caméras devient plus complexe lorsqu'il s'agit d'observer des milieux contraignants tels que les passages à niveau.

De nombreuses techniques de surveillance multi-caméra ont été étudiées dans la littérature [Ohta 2005][Fakhfakh 2011b]. La plupart de ces algorithmes se concentrent sur la détection et la cohérence de mouvement d'un même objet observé par les différentes caméras du système de surveillance. Ces systèmes utilisent généralement des descripteurs colorimétriques [Wang 2007] basés sur des modèles d'apparence simplifiés [Ilyas 2010]. Ces descripteurs sont associés pour chaque objet suivi par une ou plusieurs caméras, sans tenir compte d'aucune connaissance à priori du positionnement des caméras les unes par rapport aux autres et par rapport à l'environnement. Le mauvais positionnement spatial des caméras peut entraîner des problèmes d'occultations, d'illumination et de ré-identification des objets trop complexes. Les systèmes de surveillance existants ne sont pas suffisamment robustes pour résoudre la complexité de ce type de problèmes. Pour réduire les effets négatifs liés à ce problème, les experts essaient généralement de positionner les caméras au point le plus haut par rapport à la zone de surveillance [Foresti 1998].

Dans ce chapitre, nous proposons une méthode originale permettant de déterminer automatiquement l'architecture spatiale des caméras par rapport à l'environnement à surveiller. Il s'agit de déterminer la position et l'orientation des caméras dans l'espace pour couvrir d'une manière optimale une zone de surveillance donnée. Une modélisation du champ de vue des caméras et une méthode d'optimisation sont alors proposées pour déterminer la position géométrique désirée du système multi-caméras. Cette méthode a pour objectif de faciliter le processus d'implantation sur site du dispositif de prise de vue en calculant plus précisément la position et l'orientation des caméras par rapport à la scène à observer. Ceci est d'autant plus important lorsqu'il s'agit d'environnements contraints tels que celui d'un passage à niveau. Dans ce chapitre, nous présentons le cas de deux caméras, mais la méthode est développée pour être adaptée à un nombre quelconque de caméras. En se basant sur la géométrie descriptive [Aubert 2003], l'algorithme développé estime les positions et orientations optimales des caméras par rapport à leurs champs de vue projetés dans différents plans horizontaux et verticaux. Dans un premier temps, nous optimisons le domaine horizontal représentant la surface issue de l'intersection du champ de vue pour chaque caméra avec le plan terrestre. Ce processus d'optimisation est effectué en ramenant le centre de gravité du domaine horizontal au centre de la surface de l'environnement à percevoir. Une deuxième optimisation est effectuée dans le domaine vertical représentant la surface issue de l'intersection du

champ de vue pour chaque caméra avec un plan vertical de surveillance. À l'issue de ces deux premières étapes appliquées aux deux caméras, les champs commun et total des deux caméras sont ensuite optimisés en appliquant simultanément le processus décrit précédemment sur les domaines horizontaux et verticaux liés aux deux caméras. Dans cette dernière étape, les domaines verticaux sont obtenus par projection des champs de vue sur le plan vertical. Ce processus d'optimisation des champs de vue commun et total est réalisé en maximisant des surfaces communes (respectivement totales) et en gardant les surfaces totales (respectivement communes) supérieures à une certaine limite (définie par rapport à l'environnement). Cette méthode de positionnement de caméras a été mise en oeuvre dans le cadre du développement d'un simulateur d'aide à l'amélioration de la sécurité aux passages à niveau à base d'un système de perception multi-vue. Avant de se lancer dans la description de la méthode de positionnement des caméras, nous commençons tout d'abord par donner une vue générale de ce simulateur.

2.2 Plateforme de réalité virtuelle pour l'étude des PN

Afin de tester et valider les choix nécessaires à l'étude d'un passage à niveau (PN), une plateforme de simulation à base de réalité virtuelle a été développée par le laboratoire IRTES-SET, dans le cadre du projet PANsafer. L'objectif de cette plateforme est de pouvoir réaliser des fonctions telles que la modélisation 3D du PN, le positionnement des capteurs vidéo pour la perception de l'environnement du PN, l'édition de scénarios et la génération de séquences d'images issues de caméras virtuelles. Dans cette section, nous allons décrire brièvement cette plateforme [Lamotte 2010] en la replaçant dans le contexte de l'application traitée dans le cadre du projet PANsafer. Pour plus de détails, le lecteur intéressé pourra se référer au livrable D4.2 du projet PANsafer [Boukour 2012].

2.2.1 Contexte et démarche

2.2.1.1 Contexte

Afin d'assurer un haut niveau de sécurité et de fiabilité, le domaine ferroviaire est soumis à des contraintes importantes. Ainsi, différentes zones autour des voies dictent des règles d'implantation des équipements. Dans ce cadre, plusieurs critères peuvent être pris en compte. D'une part le volume de l'équipement, afin d'éviter tout risque de collision avec le matériel roulant. Mais également d'autres types de contraintes comme les composants utilisés ou les ondes émises par les équipements électroniques, afin d'éviter toute incompatibilité avec les équipements déjà en place ou les équipements embarqués dans les trains.

Ces contraintes, bien que nécessaires, rendent difficile l'accès aux abords des voies dans le cadre d'expérimentations. En effet, seul le personnel habilité peut travailler en bordure de voie. Et même dans ce cas, la sécurité des personnes est mise en jeu. De plus, toute intervention dans cette zone doit être signalée aux

différents trains (empruntant ces voies), qui doivent alors adapter leurs vitesses en conséquence. Ces activités ont donc inévitablement un impact sur le trafic.

Un des objectifs du projet PANSafer est de proposer une méthode permettant d'améliorer la sécurité des passages à niveau. En effet, faisant cohabiter le rail et la route, les passages à niveau sont aujourd'hui, en termes de sécurité, un des points faibles des réseaux ferrés. Dans ce cadre, le laboratoire IRTES-SET, en collaboration avec l'IFSTTAR, ont développé un système de perception (capteurs vidéo) de l'environnement des PN afin de détecter les situations potentiellement dangereuses (obstacles, passage en chicane, remontée de file, ...).

Plusieurs problèmes pratiques apparaissent, rendant ainsi difficile l'évaluation et la mise au point des algorithmes de traitement et analyse d'images, compte tenu des contraintes que nous avons citées :

- le choix du meilleur positionnement pour la ou les caméras pour la perception de l'environnement du PN. Cela comprend la détermination de la meilleure position pour le mat supportant les capteurs, mais également la détermination de la hauteur et l'orientation des caméras, tout en prenant en compte les obstacles (végétation, bâtiments, etc.).
- la nécessité de pouvoir récolter des jeux de test nécessaires au développement des algorithmes. Ces jeux de données devant être significatifs en considérant les conditions de prise de vues (heure de la journée, météo, trafic, etc.).

Afin de répondre à ces problèmes techniques, et pour permettre un développement plus facile des algorithmes sans prise de risque et sans perturbation du trafic, le laboratoire IRTES-SET a développé un outil informatique basé sur la réalité virtuelle capable de simuler l'environnement réel d'un passage à niveau.

Ce simulateur n'a pas la prétention de totalement remplacer les tests grandeur nature effectués sur le terrain, mais il permet de s'extraire des contraintes fortes que nous avons évoquées pendant toute la phase de recherche et de mise au point des techniques. Il permet également de disposer d'un environnement paramétrable offrant la possibilité de définir des scénarii précis.

2.2.1.2 Intérêt de la réalité virtuelle

Un des principes fondateurs de la réalité virtuelle est d'offrir aux utilisateurs un environnement tenant lieu de réalité, c'est-à-dire un environnement en trois dimensions dans lequel certains phénomènes physiques sont reproduits virtuellement, avec plus ou moins de précision, et qui permet à un ou plusieurs utilisateurs d'accomplir des actions proches du réel. Dans ce contexte, il est difficile de parler de degré de réalisme d'une maquette virtuelle. En effet, bien que certains phénomènes physiques soient simplifiés et peu précis (généralement à cause de leurs complexités), d'autres au contraire iront au-delà du réel. Comme par exemple l'utilisation d'un code couleur sur le sol permettant d'indiquer des zones radioactives (la radioactivité n'étant pas visible à l'oeil nu dans la réalité) sans pour autant en ressentir les effets nocifs.

Deux objectifs sont fréquemment recherchés dans les systèmes à base de réalité virtuelle :

- une économie de moyens, et donc de budget,
- le maintien d'un niveau de sécurité fort.

Le projet PANSafer ne déroge pas à cette règle puisque l'utilisation des univers virtuels dans les phases de recherche et de validation des algorithmes de traitement et d'analyse d'images permet d'éviter plusieurs tests en grandeurs nature dans un environnement potentiellement dangereux.

Toutefois, les univers virtuels offrent un second avantage important par rapport au réel : ils sont paramétrables. En effet, les univers virtuels étant " fabriqués " informatiquement, ils peuvent être configurés à souhait. Couplés à des outils de simulation ou d'animation, ils vont pouvoir être peuplés d'entités virtuelles aux comportements proches du réel (piétons, voitures, etc.).

Enfin, un dernier avantage que nous pouvons citer est que ces univers sont complètement connus géométriquement. Il est ainsi possible de les utiliser comme source d'information pour la validation par exemple.

2.2.1.3 Démarche générale

La démarche entreprise repose, dans un premier temps, sur l'application informatique spécifiquement développée dans le cadre de ce projet. En effet, celle-ci permet de simuler l'environnement réel pendant toutes les phases de conception des algorithmes de traitement d'images. Cet outil permet ainsi d'obtenir un premier niveau de validation de ces algorithmes. Dans un second temps, la validation finale des algorithmes se fera sur le terrain, avec toutes les contraintes que nous avons déjà décrites précédemment.

Cette démarche a d'une part l'objectif de limiter au maximum les expérimentations sur le terrain. Mais, l'utilisation du simulateur, permet également de paramétrer l'environnement simplement et de façon précise, en donnant ainsi la possibilité de tester différentes configurations significatives (éclairage faible, trafic important, etc.).

Pour pouvoir satisfaire à cette démarche, l'outil de simulation doit donc disposer des fonctionnalités suivantes :

- produire des images réalistes reproduisant des passages à niveau afin d'alimenter les algorithmes de traitement d'images,
- permettre de placer dans l'environnement virtuel des caméras dont les caractéristiques et les contraintes sont proches des caméras réelles,
- permettre de créer des scénarii dans lesquels des entités mobiles se déplacent (voitures, camions, motos, vélos, piétons, ...).

2.2.2 Description de l'application

Ici, nous nous attachons à décrire les étapes indispensables permettant l'utilisation de l'application. L'ensemble des fonctionnalités de l'application sont décrites dans le livrable D4.2 du projet PANsafer [Boukour 2012]

Comme le montre le schéma de la figure 2.1, l'utilisation de l'application informatique peut se résumer en quatre étapes principales.

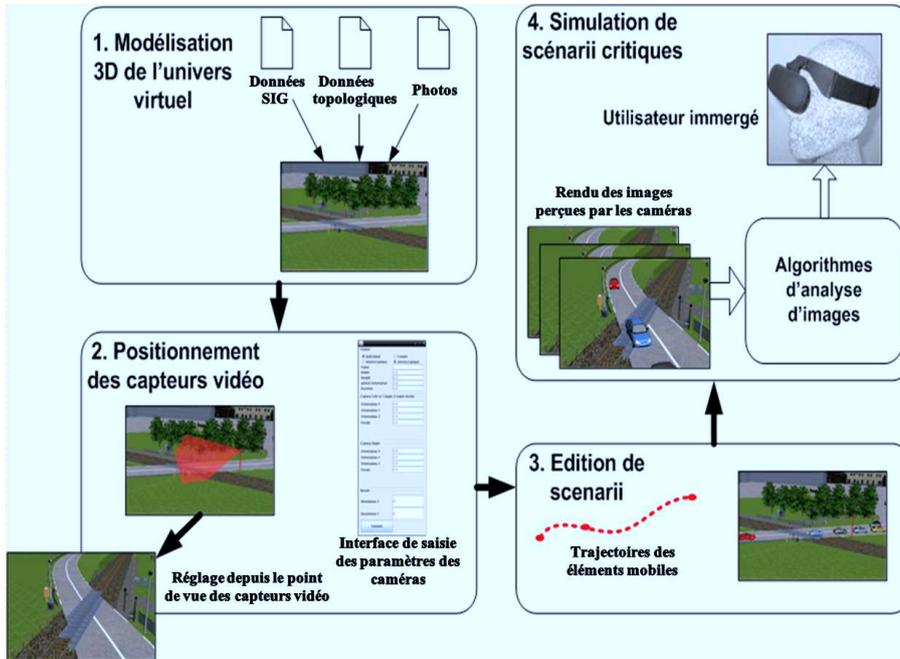


FIGURE 2.1 – Etapes pour l'utilisation du simulateur.

2.2.2.1 Modélisation 3D de l'univers virtuel

La première étape concerne la modélisation de l'univers virtuel. Cette étape préliminaire est indispensable. De la qualité de cet univers virtuel va dépendre la qualité de l'ensemble des résultats escomptés. La modélisation 3D consiste à reproduire sous forme d'objets 3D l'ensemble des éléments constituant l'univers. Cela revient à discrétiser par un ensemble de points significatifs les formes géométriques des éléments de l'univers réel. Ces points sont ensuite assemblés en facettes, généralement triangulaires, afin de constituer le maillage des objets 3D.

Pendant cette étape, deux éléments importants sont à surveiller. D'une part la justesse des géométries et d'autre part, le poids du modèle. Il est fréquent de rencontrer des univers virtuels visuellement beaux et semblant se rapprocher de la réalité, mais qui présentent des erreurs géométriques parfois importantes. Par exemple, la hauteur d'une façade ou la largeur d'une route pourrait ne pas correspondre à la réalité. Pour des applications de type visite virtuelle ou présentation architecturale, ces erreurs ne constituent pas forcément un problème grave. Elles sont même parfois voulues et accentuées afin de mettre en avant un élément important de l'environnement. Toutefois, dans notre cas, comme généralement avec

les applications de simulation, l'univers doit être géométriquement juste afin de permettre l'obtention de résultats probants. Par exemple, comment déterminer le positionnement des caméras si des différences importantes existent entre l'univers virtuel et la réalité. Le poids du modèle est également un point essentiel. Un univers trop lourd, c'est-à-dire comportant trop de facettes sera certes plus précis (ce qui ne veut pas dire forcément plus juste), mais il ne sera pas utilisable en temps réel de part sa complexité. Un compromis doit donc être trouvé entre précision et complexité (nombre de facettes).

Le laboratoire IRTES-SET possède une certaine expérience dans la modélisation d'univers virtuels destinés à la simulation. Un certain nombre d'algorithmes ont été développés afin d'automatiser une partie de la génération automatique de modèles 3D en utilisant les données des systèmes d'informations géographiques. Les univers ainsi générés sont donc géométriquement justes et constituent une première version qui peut ensuite être reprise manuellement par les infographistes pour les compléter ou les embellir.

2.2.2.2 Positionnement des capteurs vidéo

Bien que décrite en seconde position, cette étape peut toutefois être reprise plus tard pendant l'utilisation de l'application. En effet, elle constitue un des points essentiels de l'outil : permettre de simuler les caméras réelles dans l'univers virtuel. Dans cette étape, l'utilisateur va donc devoir positionner et paramétrer les caméras virtuelles qui permettront à terme de générer les images de synthèses transmises aux algorithmes de traitement d'images. Il est également envisageable que les positions des caméras soient déterminées de manière automatique. L'un des objectifs de ce travail de thèse est de proposer un algorithme d'optimisation permettant de positionner les capteurs vidéo de manière à percevoir de manière optimale l'environnement à surveiller. La présentation de l'algorithme d'optimisation proposé débute à partir de la section 2.3 de ce chapitre.

2.2.2.3 Edition de scénarii

A ce niveau, nous disposons d'un univers virtuel en trois dimensions dans lequel nous avons placé des caméras virtuelles. Toutefois, celui-ci est complètement statique. Il n'abrite aucune entité animée, aucune "vie". Dans cette troisième étape, sont gérés différents scénarii de déplacement d'entités.

Le principe du déplacement d'une entité est le même quelque soit son type (voiture, piéton, camion, etc.). Il repose sur la saisie de courbes définissant les trajectoires à suivre, et un ensemble d'informations temporelles décrivant la dynamique avec laquelle les entités vont suivre ces trajectoires. La multiplication des entités et des trajectoires permet ainsi la création de scénarii complexes et représentatifs de la réalité. L'objectif est d'animer l'environnement afin que les images générées par les caméras virtuelles soient suffisamment significatives et correspondent à des images réelles.

2.2.2.4 Simulation de scénarii critiques

La dernière étape concerne l'utilisation des images générées par l'application. En effet, une fois l'environnement complètement configuré, c'est-à-dire disposant de ses caméras virtuelles et d'un ensemble d'entités animées, les images de synthèses peuvent être générées.

Dans cette étape, deux approches peuvent être considérées. Une connexion directe entre l'application et les algorithmes de traitement d'images. Ce qui offre la possibilité d'obtenir un retour temps réel des traitements pour un éventuel retour d'information vers un utilisateur immergé (utilisation d'un casque de réalité virtuelle par exemple). Dans la seconde approche, il n'existe pas de connexion directe entre l'application et les algorithmes. Les images sont sauvegardées sous forme de séquences qui peuvent ensuite être utilisées par les algorithmes. L'avantage de cette seconde approche est la possibilité de réutiliser ces séquences d'images plusieurs fois, ce qui peut être très utile pendant les phases de mise au point des algorithmes. Bien entendu, il est envisageable, dans une certaine mesure, de mixer ces deux approches et ainsi de pouvoir enregistrer la séquence d'images en même temps que de les transférer aux algorithmes. Toutefois, on peut vite se retrouver limité par la puissance des machines, et par conséquent perdre l'aspect temps réel de l'application.

2.3 Problématique

Considérons un système composé de deux caméras. L'objectif est de déterminer la position et l'orientation optimales des caméras pour couvrir la scène désirée. Le processus d'optimisation concerne à la fois le champ de vue total et le champ de vue commun aux deux caméras.

La première étape consiste à maximiser séparément le champ de vue de chaque caméra par rapport à la scène à observer. Ensuite, en considérant simultanément les deux caméras, on peut maximiser leur champ de vue total en minimisant leur champ de vue commun, ou inversement, on peut minimiser leur champ de vue total en maximisant leur champ de vue commun (Fig. 2.2). Ainsi, si les deux caméras sont proches l'une de l'autre, alors l'objectif est de maximiser le champ de vue total, en gardant le champ de vue commun supérieur à une certaine limite. A l'inverse, si les deux caméras sont distantes l'une de l'autre, alors l'objectif est de maximiser le champ de vue commun, en gardant le champ de vue total supérieur à une certaine limite.

2.3.1 Modélisation du champ de vue

Le champ de vue de chaque caméra est modélisé par une pyramide (Fig. 2.3) [Burke 1996][Jain 2009]. Le sommet de cette pyramide représente le centre optique du capteur vidéo. La ligne joignant le sommet de la pyramide au centre de la surface de base de la pyramide (axe pyramidal) forme l'axe optique du capteur. Le champ de vision de la caméra est déterminé par la largeur, la hauteur ou la diagonale du

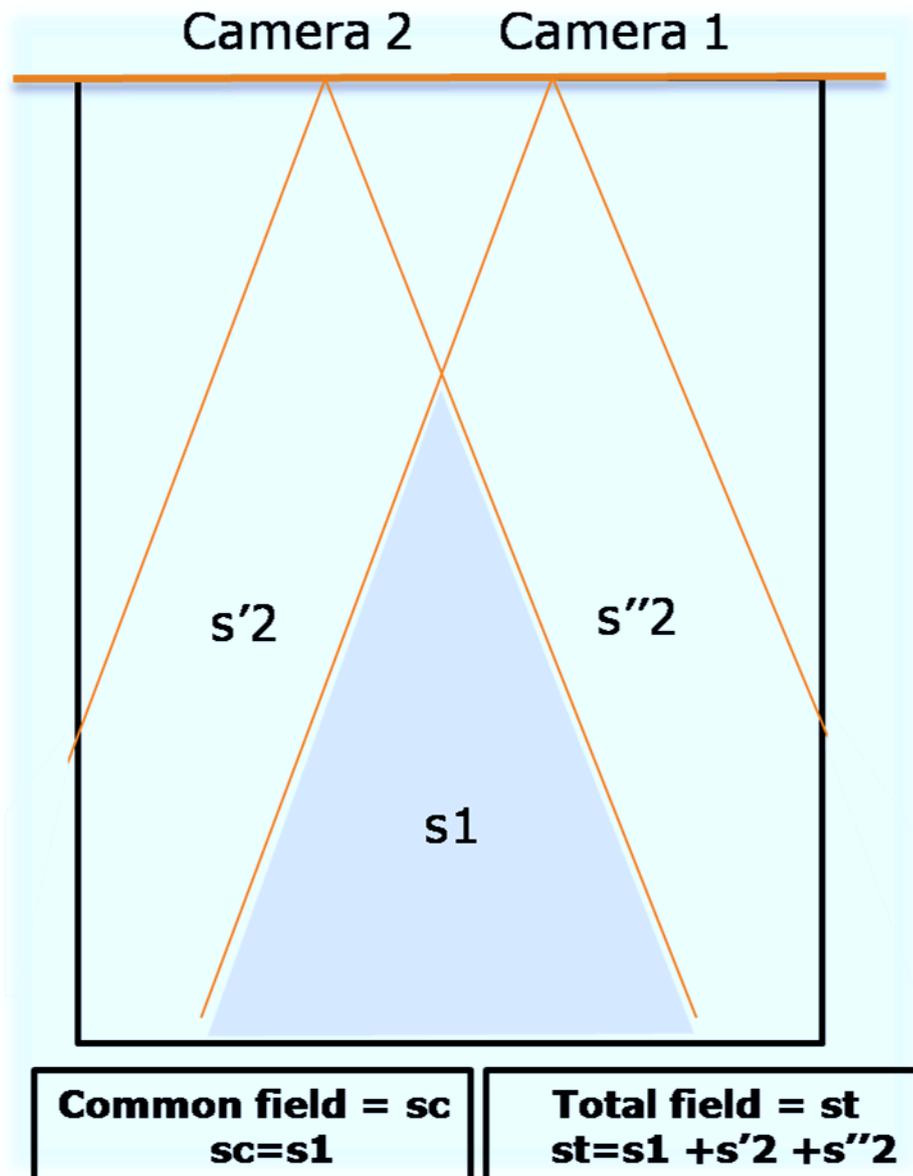


FIGURE 2.2 – Contraintes d'optimisation (champ commun / total).

plan image et la distance par rapport à l'objectif (distance focale) est :

$$fov_{di} = 2 \times \arctan \frac{di}{2f} \approx \frac{180di}{\pi f} \text{degr} \quad (2.1)$$

$$fov_h = 2 \times \arctan \frac{h}{2f} \approx \frac{180h}{\pi f} \text{degr} \quad (2.2)$$

$$fov_v = 2 \times \arctan \frac{v}{2f} \approx \frac{180v}{\pi f} \text{degr} \quad (2.3)$$

où fov_{di} , fov_h , fov_v sont respectivement les champs de vue diagonal, horizontal et vertical. di , h , v représentent respectivement les dimensions diagonale, horizontal et vertical du plan image. f est la distance focale du capteur vidéo. Pour contrôler le déplacement des caméras dans l'espace (Fig. 2.3), nous utilisons 3 paramètres de translation (t_x, t_y, t_z) et trois paramètres de rotation (φ, θ, ψ) qui sont les angles d'Euler :

$$P_N = R \times P_i + T \quad (2.4)$$

$$R = Rot(z, \varphi) \times Rot(x, \theta) \times Rot(z, \psi) \quad (2.5)$$

$$T = [t_x, t_y, t_z]^T \quad (2.6)$$

où P_N , P_i sont respectivement les nouvelles positions et les positions initiales des points de la pyramide. T représente le vecteur de translation. $Rot(z, \varphi)$, $Rot(x, \theta)$, $Rot(z, \psi)$ sont les rotations d'Euler. Pour un petit déplacement, la matrice de rotation R est approximée de la façon suivante :

$$R = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.7)$$

$$R \approx \begin{bmatrix} 1 - \psi\varphi & -(\varphi + \psi) & \varphi\theta \\ \varphi + \psi & 1 - \psi\varphi & -\theta \\ \theta\psi & \theta & 1 \end{bmatrix} \quad (2.8)$$

2.3.2 Intersection de pyramides : Géométrie descriptive

La géométrie descriptive [Ciblac 2004] est un outil de résolution graphique de problèmes d'intersections entre plusieurs objets définis de façon géométrique dans l'espace à trois dimensions. Elle consiste à présenter un ou plusieurs objets de l'espace en un certains nombre de projections planaires tout en conservant le maximum

de propriétés utiles. Nous utilisons cette approche pour reconstruire le volume d'intersection entre les champs de vue de deux caméras. La méthode consiste à déterminer l'intersection entre les faces triangulaires des deux pyramides représentant les champs de vue des deux caméras.

On se propose alors d'étudier l'intersection des deux faces triangulaires ABC et $A'B'C'$ (Fig. 2.4). Nous nous plaçons dans le cas où les bases BC et $B'C'$ des deux triangles (ABC) et ($A'B'C'$) appartiennent à un même plan P (plan terrestre). Les points B, C, B' et C' correspondent à des points d'intersection des champs de vue des caméras avec la terre (plan horizontal terrestre P). Nous supposons que les projections des points A, B, B', A', B' et C' sont connues. Le problème consiste à construire la droite d qui est l'intersection des plans ABC et $A'B'C'$ et à chercher sur cette droite un segment commun aux deux triangles ABC et $A'B'C'$. L'hypothèse de coplanarité de BC et $B'C'$ nous donne un premier point d'intersection I_0 sur la droite d ($I_0 = (BC) \cap (B'C')$). Soit O le point d'intersection de la droite (AA') avec le plan terrestre P . Ensuite, soit P' un plan auxiliaire contenant la droite (AA'). Ce plan P' coupe le plan P suivant une droite D passant par le point O . La droite D coupe respectivement (BC) et ($B'C'$) aux points w et w' . Les droites (Aw) et ($A'w'$) appartiennent au même plan P' . Soit I leur point d'intersection ($I = (Aw) \cap (A'w')$). Comme les droites (Aw) et ($A'w'$) appartiennent aussi aux plans (ABC) et ($A'B'C'$), alors leur point d'intersection I appartient à la droite d d'intersection des plans (ABC) et ($A'B'C'$). La droite d'intersection d correspond donc à la droite (II_0). Finalement, il reste à faire tourner la droite D autour du point O pour trouver le segment d'intersection entre les triangles ABC et $A'B'C'$. De la même manière, on peut tracer les segments d'intersection entre tous les faces latérales des deux pyramides, et ainsi, on peut construire le volume d'intersection des deux champs de vue des caméras (Fig. 2.5).

2.3.3 Plans de projection

La forme pyramidale du champ de vue des caméras est une forme convexe. L'intersection de deux pyramides est aussi une forme convexe [Grünbaum 2003]. Ainsi, en utilisant la géométrie descriptive [Aubert 2003], la méthode proposée est décomposée de manière à ce que le processus d'optimisation opère dans des domaines horizontaux et verticaux, tout en préservant la plupart des propriétés. Ces domaines 2D sont obtenus par projection/intersection des champs de vue des caméras sur/avec des plans horizontaux et verticaux. Par conséquent, optimiser les champs de vue total et commun des deux caméras revient à optimiser les surfaces des domaines 2D horizontaux et verticaux. Trois plans sont utilisés : plan horizontal, plan vertical d'orientation et plan vertical de projection.

Le plan horizontal représente tout simplement le plan de la terre. L'intersection de ce plan avec le champ de vue d'une caméra correspond à la base de vue de la caméra, qui définit une surface horizontale (Fig. 2.6).

Avant de définir le plan vertical d'orientation, nous supposons que, dans le cadre de notre application, les objets à détecter et à suivre entrent dans la scène par deux

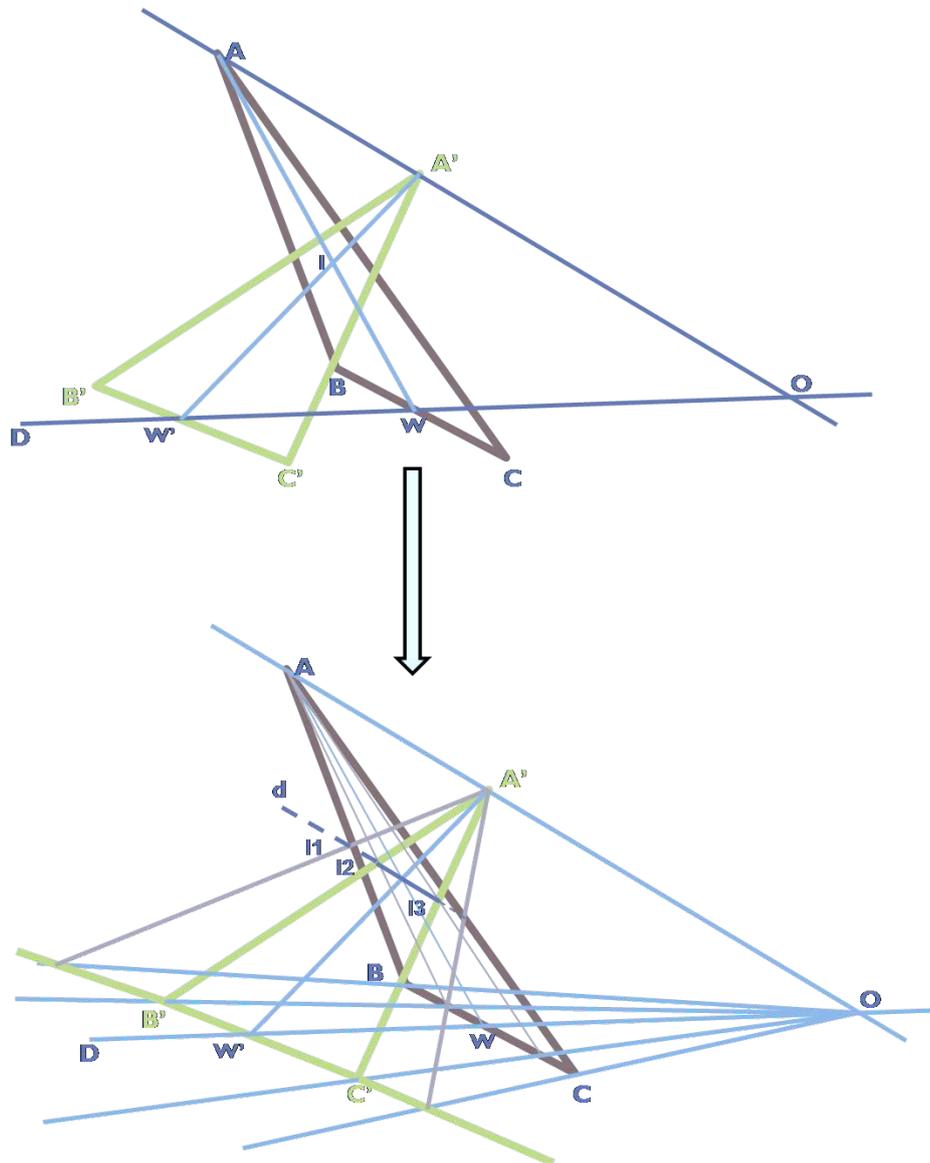


FIGURE 2.4 – Intersection entre deux faces latérales de pyramides.

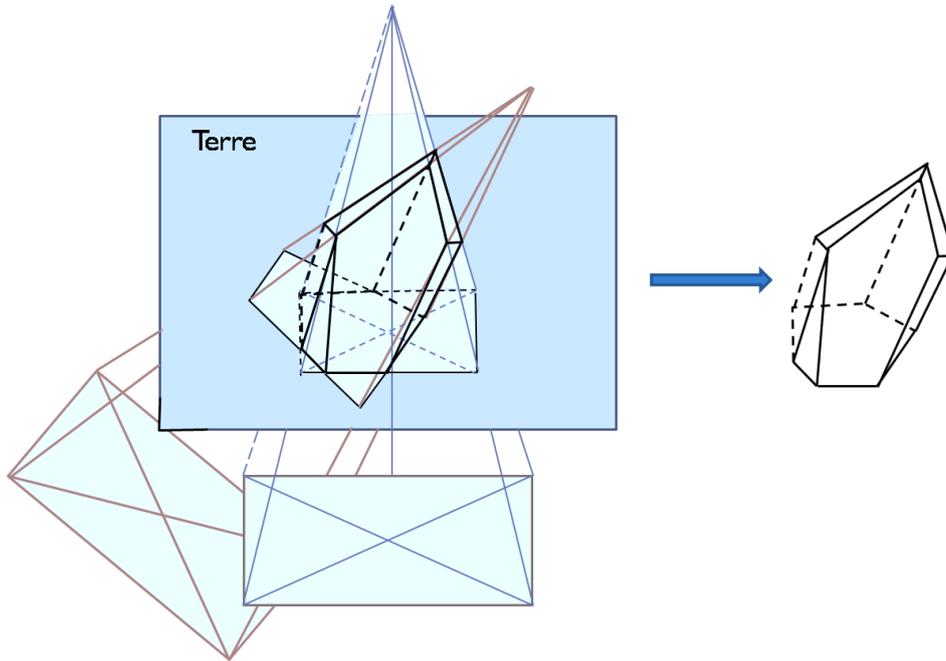


FIGURE 2.5 – Représentation du champ de vue commun de deux caméras (forme d'un polyèdre convexe).

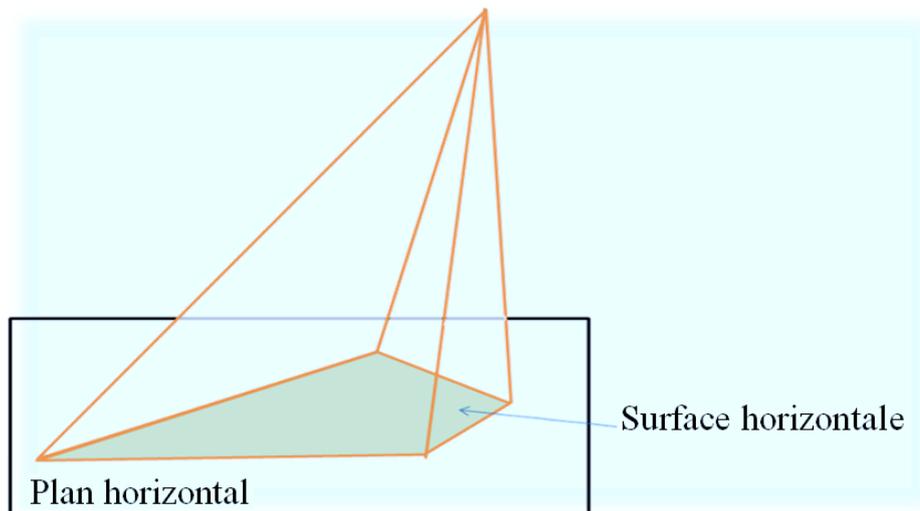


FIGURE 2.6 – Plan horizontal.

côtés opposés (dans notre cas, côtés route dans l'environnement d'un passage à niveau). Le plan vertical d'orientation correspond au plan vertical représentant la limite de la zone à surveiller de chaque côté (Fig. 2.7). Ici, l'objectif est d'orienter séparément les caméras vers le même plan vertical. Initialement, ce plan peut être loin du champ de vue de la caméra sans aucune intersection. Pour atteindre ce plan, des plans verticaux intermédiaires sont alors considérés. Le premier plan intermédiaire et le champ de vue de la caméra doivent avoir une intersection. La caméra est soumise ensuite à une procédure d'orientation pour chacun des plans intermédiaires : une fois la caméra orientée en utilisant un plan intermédiaire, la procédure continue de la même manière avec un autre plan intermédiaire (qui coupe le champ de vue de la caméra), jusqu'à atteindre le plan vertical désiré. L'intersection du champ de vue de la caméra avec un plan vertical produit un polygone vertical, qui définit une surface verticale.

Le plan vertical de projection constitue un plan vertical sur lequel les champs de vue total et commun des deux caméras sont projetés orthogonalement (Fig. 2.8). Une hauteur minimale est considérée afin de prendre en compte la détection d'objets de grande taille, comme les camions par exemple, et pouvant traverser le passage à niveau. En considérant cette hauteur limite, la projection des champs de vue total et commun des deux caméras sur le plan vertical de projection produit deux polygones verticaux, l'un définissant une surface totale verticale et l'autre définissant une surface commune verticale.

2.4 Processus d'optimisation

A partir des positions initiales des deux caméras, le processus d'optimisation procède en trois étapes (Fig. 2.9). Les deux premières étapes sont appliquées à chacune des deux caméras séparément. La troisième étape est appliquée en considérant les deux caméras simultanément. Basée sur l'utilisation du plan horizontal, la première étape commence par centrer la base de vue de chaque caméra autour d'une position désirée. Ensuite, la surface horizontale est maximisée en gardant la base de vue de la caméra centrée autour de la position désirée. La deuxième étape consiste tout d'abord à déplacer en haut (selon une position désirée) le polygone résultant de l'intersection du champ de vue de la caméra et le plan vertical d'orientation. Ensuite, la surface du polygone est maximisée. Dans cette étape également, la base de vue de la caméra doit rester centrée autour de la position désirée. Basée sur le plan vertical de projection, la troisième étape concerne l'optimisation des surfaces totales et communes en considérant les deux caméras simultanément. Les surfaces horizontales (resp. verticales) sont produites à partir du plan horizontal (resp. plan vertical de projection).

2.4.1 Optimisation relative au plan horizontal

Cette étape est appliquée à chacune des caméras séparément. Considérons l'intersection du champ de vue d'une caméra et le plan horizontal (Fig. 2.10). Cor-

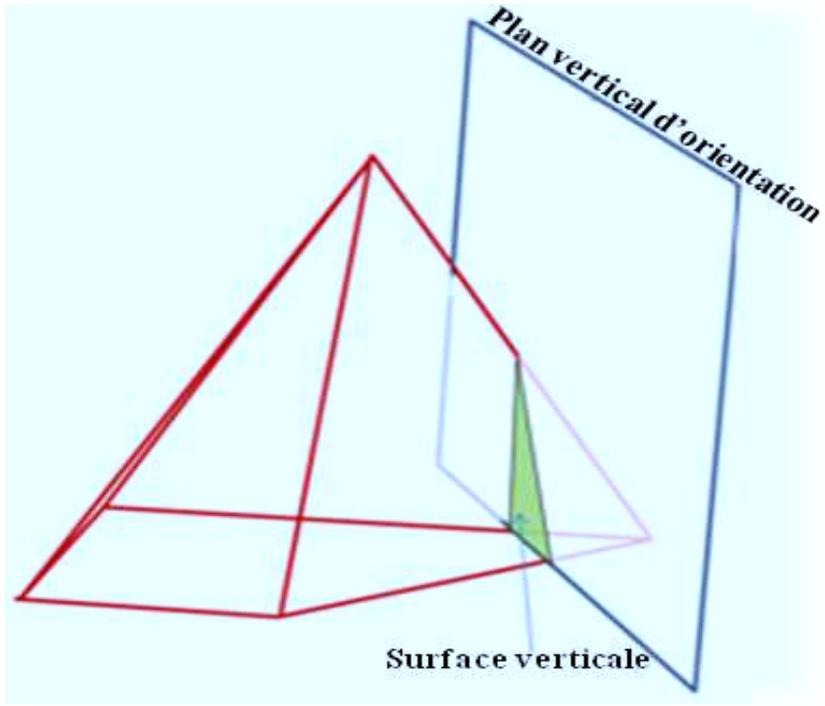


FIGURE 2.7 – Plan vertical d'orientation.

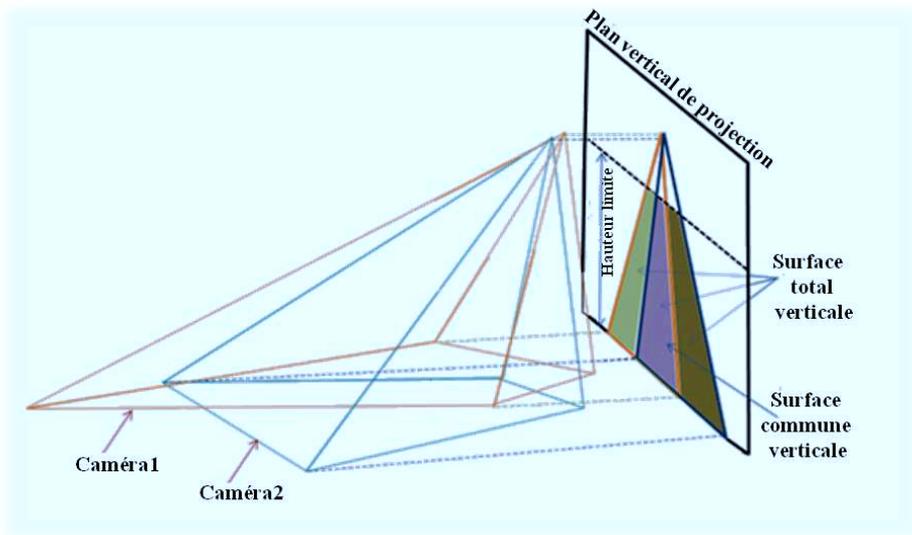


FIGURE 2.8 – Plan vertical de projection.

respondant à la base de vue de la caméra, cette intersection est un quadrilatère $ABCD$. Soient G_h et A_h respectivement le centre de gravité et la surface de la base de vue $ABCD$. En utilisant un espace ortho-normal fixé, G_h et A_h peuvent être exprimés de la manière suivante :

$$\begin{cases} x_G = \frac{1}{2} \sum_{i=0}^n (x_i - x_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \\ y_G = \frac{1}{2} \sum_{i=0}^n (y_i - y_{i+1})(x_i y_{i+1} - x_{i+1} y_i) \\ A_h = \frac{1}{2} \sum_{i=0}^n (x_i y_{i+1} - x_{i+1} y_i) \end{cases} \quad (2.9)$$

où x_G et y_G sont les coordonnées du centre de gravité G_h . $n = 4$ est le nombre de côtés du quadrilatère $ABCD$, tel que $x_{i=0..4} = x_A, x_B, x_C, x_D, x_A$ et $y_{i=0..4} = y_A, y_B, y_C, y_D, y_A$. Comme G_h et A_h dépendent de la position et de l'orientation de la caméra, ils peuvent être réécrits comme suit :

$$\begin{cases} x_G = x_G(\varphi, \theta, \psi, t_x, t_y, t_z) \\ y_G = y_G(\varphi, \theta, \psi, t_x, t_y, t_z) \\ A_h = A_h(\varphi, \theta, \psi, t_x, t_y, t_z) \end{cases} \quad (2.10)$$

où (t_x, t_y, t_z) représente le vecteur translation de la caméra, et (φ, θ, ψ) définit son orientation avec la représentation d'Euler.

L'optimisation de la surface horizontale commence par centrer la base de vue de la caméra autour d'un point désiré $P_{désiré}$, qui correspond au centre de la surface horizontale de la zone à observer. En utilisant la technique de Levenberg-Marquardt [Lourakis 2005], la procédure de centrage est assurée en minimisant la quantité suivante :

$$F(\beta) = \|G_h(\beta) - P_{désiré}\|^2 \quad (2.11)$$

où $\beta = (\varphi, \theta, \psi, t_x, t_y, t_z)$

A l'issue de la procédure de centrage, la surface du quadrilatère $ABCD$ est maximisée, tout en gardant la base de vue de la caméra centrée autour du point $P_{désiré}$. Pour ce faire, la technique de Levenberg-Marquardt est utilisée pour minimiser la quantité suivante :

$$F(\beta) = \|M_h(\beta) - M_{désiré}\|^2 \quad (2.12)$$

où M_h est un vecteur composé du centre de gravité G_h et la surface A_h . $M_{désiré}$ est un vecteur composé du point de centrage $P_{désiré}$ et la surface désirée $A_{désiré}$.

Pour minimiser la fonction F , nous proposons de résoudre l'équation suivante :

$$\left[\mu \times Id + DF_{x_k}^T DF_{x_k} \right] d_k = -DF_{x_k}^T F_{x_k} \quad (2.13)$$

où μ est le facteur d'amortissement. Ce facteur est ajusté à chaque nouvelle itération lors de l'exécution de l'algorithme de Levenberg-Marquardt (Annexe A). Id est une matrice identité. DF_{x_k} est la matrice Jacobienne par rapport à β et $DF_{x_k}^T DF_{x_k}$ représente la matrice Hessienne. d_k est un paramètre d'incrément ajouté à β et calculé à chaque nouvelle itération k de l'algorithme.

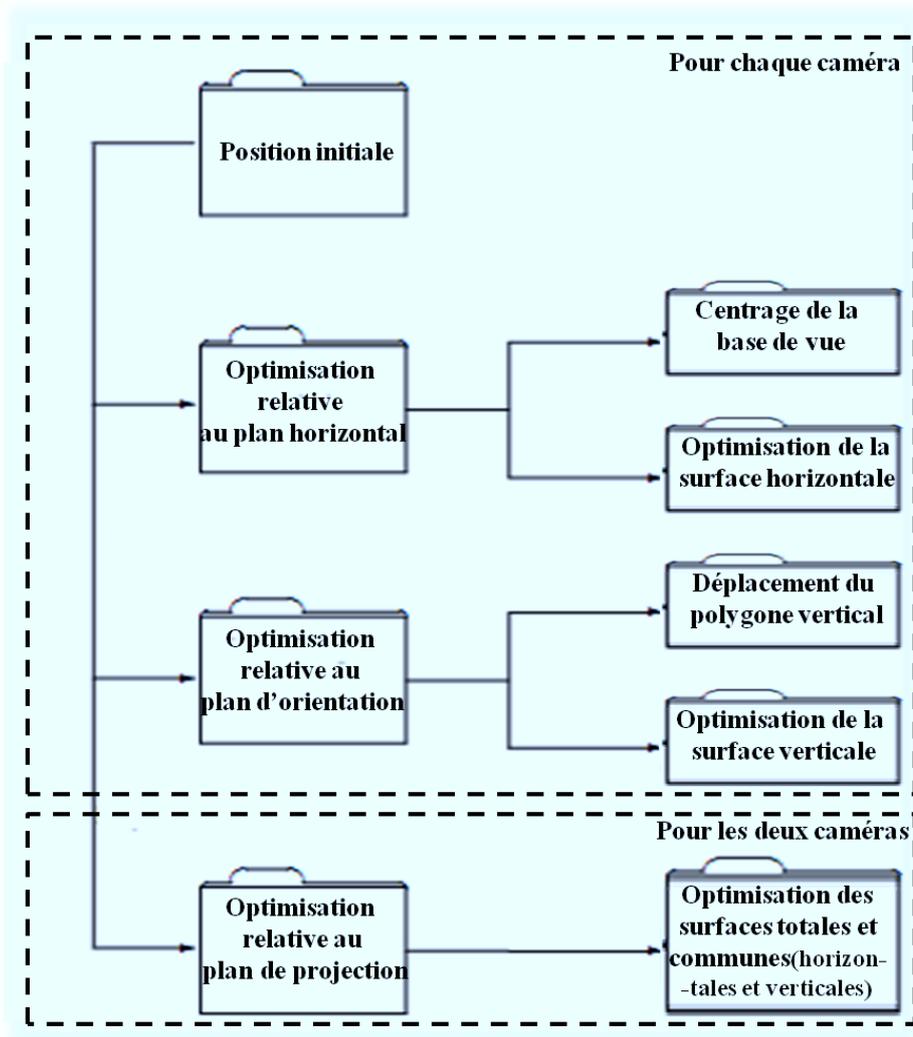


FIGURE 2.9 – Etapes du processus d'optimisation.

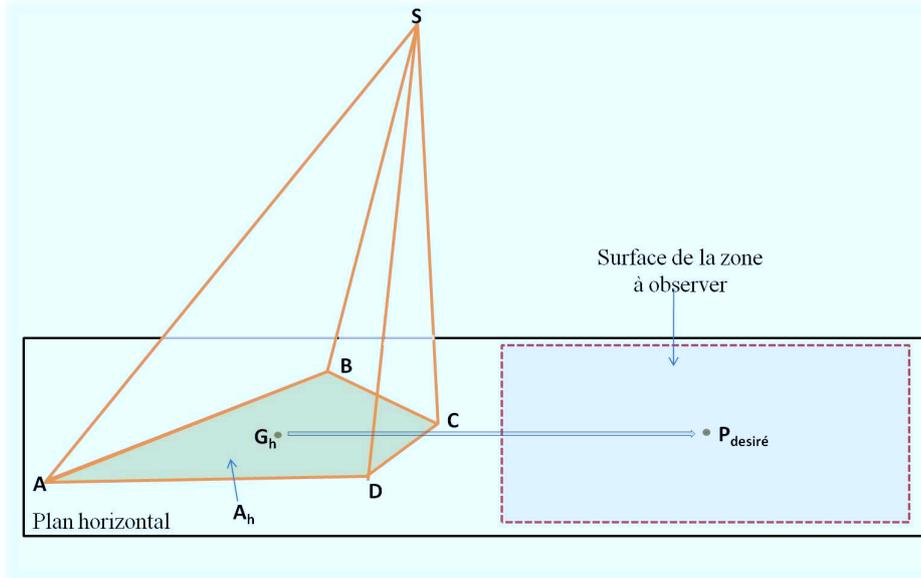


FIGURE 2.10 – Optimisation relative au plan horizontal.

2.4.2 Optimisation relative au plan vertical d'orientation

Comme pour l'étape précédente, cette phase est considérée pour chaque caméra indépendamment l'une de l'autre. Elle consiste à orienter une caméra vers le plan vertical d'orientation. Rappelons que ce plan représente une limite verticale de la zone à observer. Etant donné un plan vertical d'orientation, coupant le champ de vue de la caméra (Fig. 2.11), l'intersection de ce plan avec le champ de vue de la caméra est un polygone, avec une forme de triangle, quadrilatère ou pentagone. Ce polygone définit une surface verticale à maximiser. Afin de forcer le champ de vue de la caméra à couvrir de manière optimale le plan d'orientation, nous proposons dans un premier temps de déplacer verticalement le centre de gravité du polygone vers le point correspondant à la projection orthogonale de l'apex de la pyramide (représentant le champ de vue de la caméra) sur le plan d'orientation. Ceci est appliqué, tout en gardant la base de vue de la caméra centrée autour du point désiré $P_{désiré}$ (eq.2.11).

Initialement, le plan vertical d'orientation peut être éloigné du champ de vue de la caméra, avec aucune intersection entre eux. Pour atteindre ce plan, nous considérons des plans verticaux intermédiaires (Fig. 2.11). Le premier plan intermédiaire doit couper le champ de vue de la caméra. La caméra est soumise à la procédure d'orientation pour chaque plan intermédiaire : une fois la caméra orientée par rapport à un plan intermédiaire, un autre plan intermédiaire coupant le champ de vue de la caméra est considéré pour exécuter la même procédure, jusqu'à atteindre le plan vertical d'orientation désiré.

Lorsque le plan vertical d'orientation est atteint, une procédure de maximisation

de la surface, définie par le polygone par l'intersection de ce plan et du champ de vue de la caméra, est mise en oeuvre en utilisant la technique de Levenberg-Marquardt (Annexe A). Le processus de maximisation est assuré, tout en gardant encore une fois la base de vue de la caméra centrée autour du point désiré $P_{désiré}$.

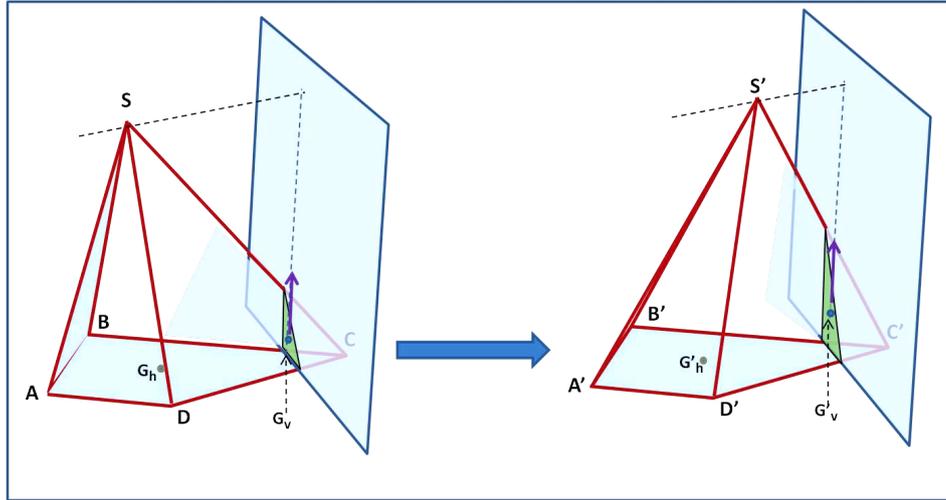


FIGURE 2.11 – Optimisation relative au plan vertical d'orientation.

2.4.3 Optimisation relative au plan vertical de projection

Cette dernière étape est appliquée aux deux caméras simultanément. Elle prend en compte à la fois le plan horizontal et le plan vertical de projection (Fig. 2.12). Comme indiqué auparavant, une hauteur limite est considérée sur le plan vertical de projection pour pouvoir détecter des objets de grande taille tels que des camions. Dans cette étape, il s'agit d'optimiser les surfaces totales et communes (horizontales et verticales) représentant l'intersection du champ de vue des caméras avec le plan horizontal et le plan vertical de projection. Selon la position des caméras, calculée lors des deux premières étapes (optimisation relative aux plans horizontal et vertical), l'une des deux situations suivantes se présente : maximiser les surfaces totales (horizontale et verticale) en minimisant les surfaces communes (horizontale et verticale), ou l'inverse. Ainsi, si les caméras sont proches l'une de l'autre, l'objectif est de maximiser les surfaces totales (horizontale et verticale), tout en gardant les surfaces communes (horizontale et verticale) supérieures à une certaine limite. Inversement, si les caméras sont distantes l'une de l'autre, l'objectif est de maximiser les surfaces communes (horizontale et verticale), tout en gardant les surfaces totales (horizontale et verticale) supérieures à une certaine limite. Comme pour les étapes précédentes, la procédure d'optimisation est implémentée en utilisant la technique de Levenberg-Marquardt (Annexe A).

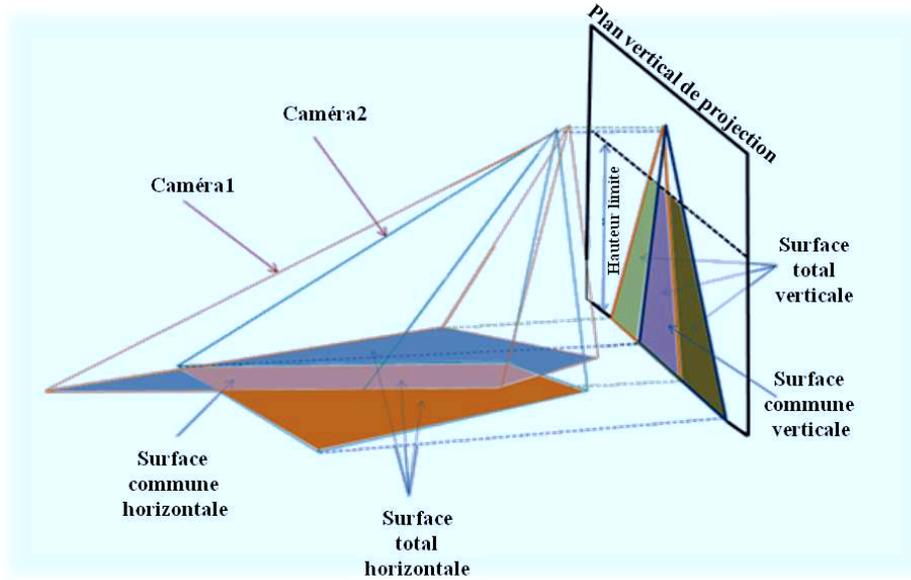


FIGURE 2.12 – Optimisation relative au plan horizontal et au plan vertical de projection.

2.5 Résultats expérimentaux

Dans cette section, nous présentons quelques résultats de simulation de la méthode d'optimisation proposée. Correspondant aux données d'un capteur vidéo réel, les paramètres intrinsèques des deux caméras sont les suivants :

- 4 mm pour la longueur focale.
- 4.8 mm, 6.8 mm et 8 mm respectivement pour la taille horizontale, verticale et diagonale du capteur vidéo.
- 77.3° , 61.9° et 90° respectivement pour l'angle de vue horizontal, vertical et diagonal.

Les dimensions de la scène à observer sont illustrées sur la Fig. 2.13. Les surfaces horizontale et verticale à observer sont respectivement $A_h = 64m^2$ et $A_v = 16m^2$. La position désirée autour de laquelle la base de vue de chacune des caméras doit être centrée est $P_{désiré} = (6, 2)$. La hauteur limite sur le plan vertical de projection est égale à 2m. Les surfaces totales horizontale et verticale désirées sont respectivement $STHD = 64m^2$ et $STVD = 16m^2$. Les surfaces communes horizontale et verticale désirées sont respectivement $SCHD = 22m^2$ et $SCVD = 5m^2$.

Dans ce qui suit, nous présentons quelques simulations pour montrer l'évolution des champs de vue des caméras durant les différentes étapes d'optimisation. La figure 2.14 montre la position initiale des deux caméras par rapport à l'environnement à surveiller. Pour simplifier la présentation, nous considérons le champ de vue d'une seule caméra (Fig. 2.15, Fig. 2.18, Fig. 2.19). Durant cette phase de

mise en place des caméras, plusieurs optimisations ont été réalisées. Une première optimisation consiste à ramener toujours le centre de gravité G_0 de la surface horizontale $A_0B_0C_0D_0$ (Fig. 2.15) vers la position désirée G_d . G_d représente le centre du domaine horizontale de l'environnement à percevoir. Cette étape sera suivie par une maximisation de la surface horizontale $A_0B_0C_0D_0 \rightarrow A_1B_1C_1D_1 \rightarrow A_2B_2C_2D_2$. La figure 2.15 illustre les étapes d'évolution du champ de vue (forme, rotation et translation) durant cette phase d'optimisation. À l'issue de cette dernière étape, la surface de surveillance horizontale est passée de $25m^2$ à $30m^2$. La figure 2.16 illustre le processus de centrage de la base de vue de chacune des deux caméras. Initialement, les bases de vue sont centrées respectivement autour des points $(0, 0)$ et $(4, -6)$. Nous constatons que la procédure de centrage conduit bien les bases de vue des deux caméras à un centrage autour de la position désirée $P_{désiré}$. La figure 2.17 montre l'évolution des surfaces horizontales des bases de vue des deux caméras durant le processus d'optimisation relative au plan horizontal. On constate que les surfaces augmentent, tout en gardant les bases de vue des deux caméras centrées autour la position désirée.

Une deuxième optimisation consiste à orienter le champ de vue vers un plan vertical de surveillance. Durant cette étape, le champ de vue sera déplacé continuellement par l'intermédiaire de plusieurs plans auxiliaires verticaux jusqu'à orienter le champ de vue vers le plan vertical de surveillance (Fig. 2.18). Le rôle de ces plans auxiliaires est de forcer le champ de vue à avoir toujours une intersection avec ces plans verticaux. Cette stratégie oblige la caméra à prendre une telle orientation pour couvrir ces plans. Cette étape d'orientation sera suivie par une autre étape de maximisation de la surface verticale $I_3J_3K_3L_3 \rightarrow I_4J_4K_4L_4$ ainsi que la surface horizontale $A_3B_3C_3D_3 \rightarrow A_4B_4C_4D_4$. Les figure 2.18 et figure 2.19 illustrent le processus d'évolution du champ de vue durant cette phase d'optimisation. À l'issue de cette phase, la surface de surveillance horizontale a augmenté de $30m^2$ à $49m^2$ et la surface de surveillance verticale a augmenté de $0m^2$ à $7m^2$. La figure 2.20 montre l'évolution des surfaces verticales (intersection du plan vertical d'orientation et les champs de vue des deux caméras) durant le processus d'optimisation relative au plan vertical d'orientation. Nous constatons une croissance des surfaces verticales, tout en gardant les bases de vue des deux caméras centrées autour de la position désirée. La figure 2.20 montre également l'évolution des surfaces horizontales des bases de vue des deux caméras. Après une courte décroissance, due à une mauvaise orientation des caméras par rapport au plan vertical d'orientation au début du processus, nous obtenons une augmentation des surfaces horizontales des bases de vue des deux caméras.

La figure 2.21 illustre le processus d'optimisation relative simultanément au plan horizontal et au plan vertical de projection. Dans cet exemple, le processus d'optimisation maximise les surfaces totales horizontale (STH) et verticale (STV), en minimisant les surfaces communes horizontale (SCH) et verticale (SCV). Nous pouvons constater que STH et STV convergent respectivement vers des valeurs ($60.3m^2$ et $11.9m^2$) proches des surfaces désirées $STHD = 64m^2$ et $STVD = 16m^2$. Par ailleurs, les surfaces communes horizontale (SCH) et verticale (SCV)

atteignent des valeurs ($23.3m^2$ et $5.6m^2$) supérieures respectivement aux limites fixées $SCHD = 22m^2$ et $SCVD = 5m^2$.

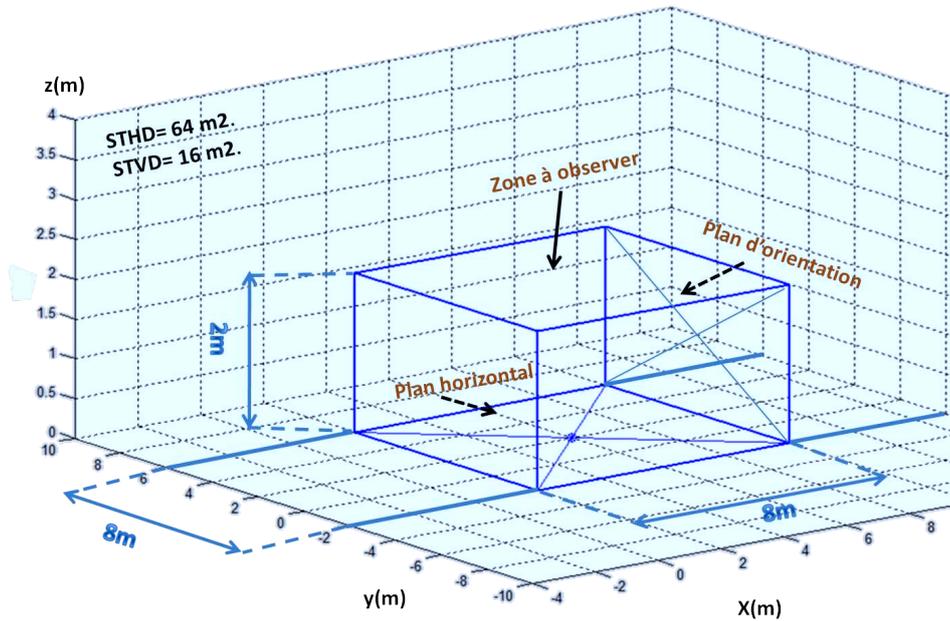


FIGURE 2.13 – Dimensions de la scène à observer.

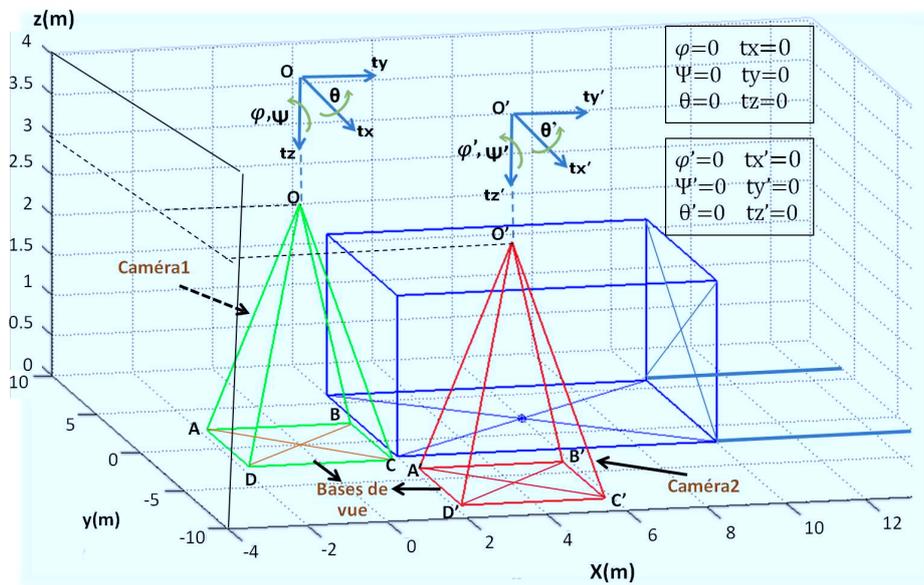


FIGURE 2.14 – Position initiale des caméras par rapport à l'environnement à surveiller.

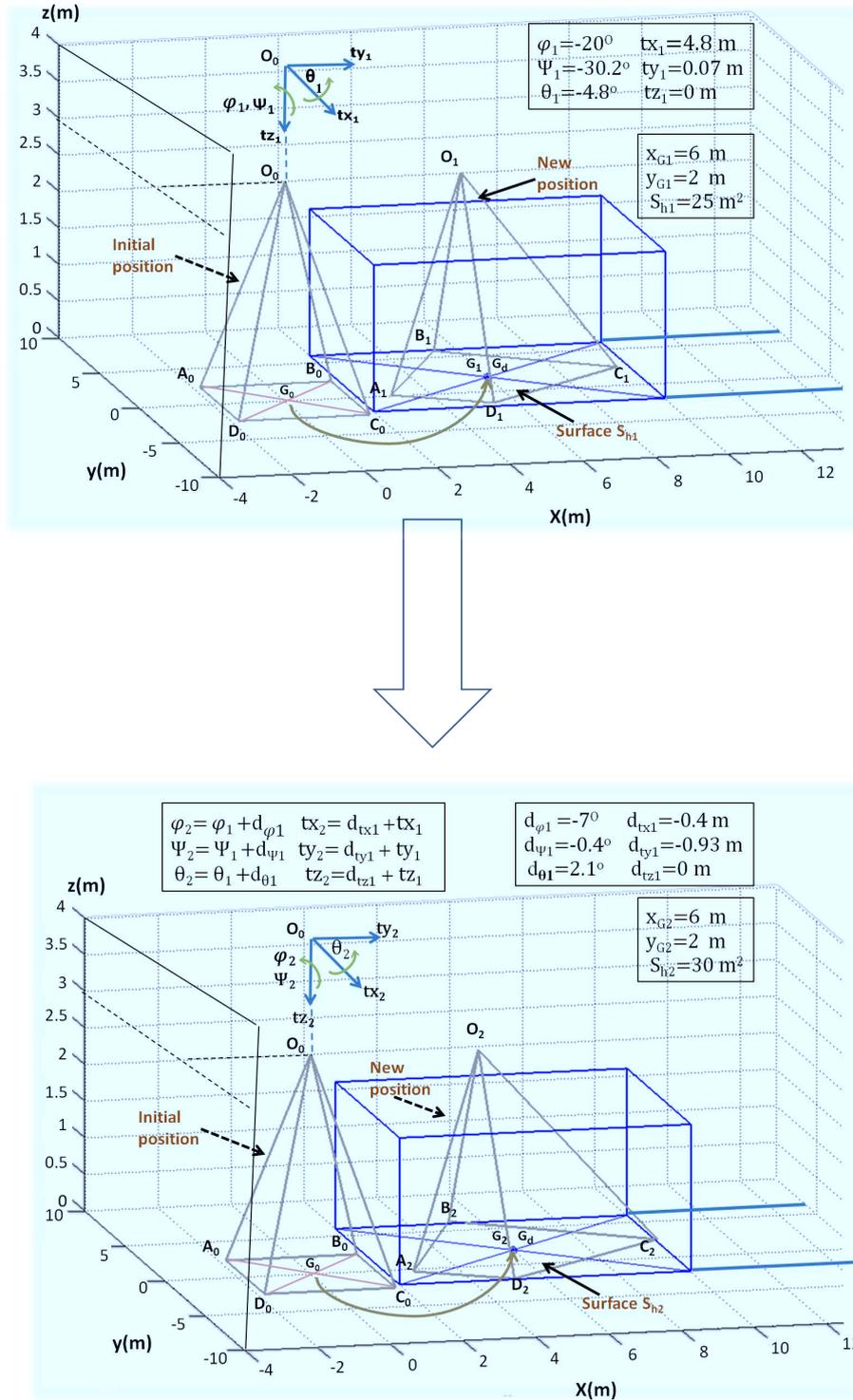


FIGURE 2.15 – Phase d'optimisation par rapport au plan horizontal : centrage et optimisation de la surface horizontale de la base de vue d'une caméra.

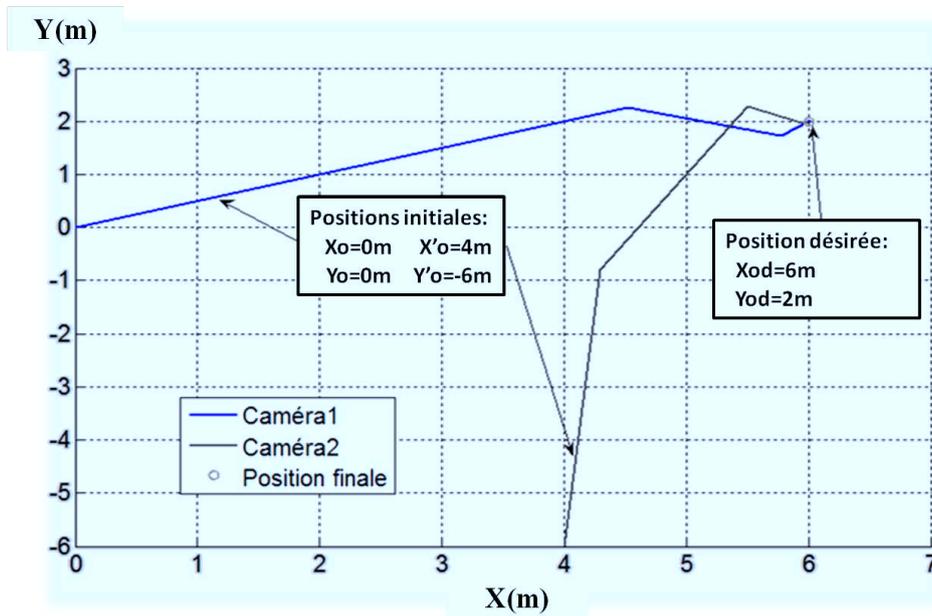


FIGURE 2.16 – Centrage de la base de vue.

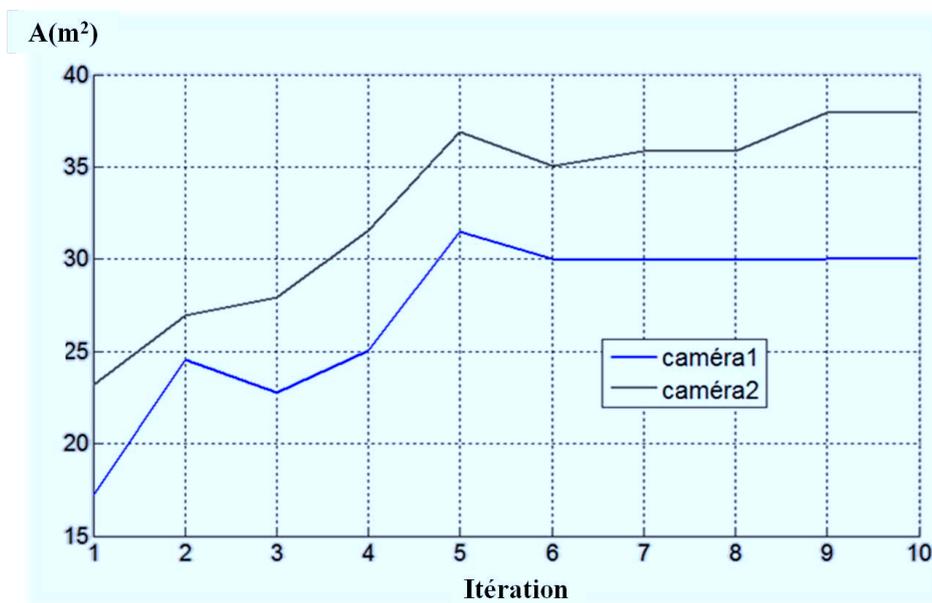


FIGURE 2.17 – Evolution des surfaces durant l'optimisation relative au plan horizontal.

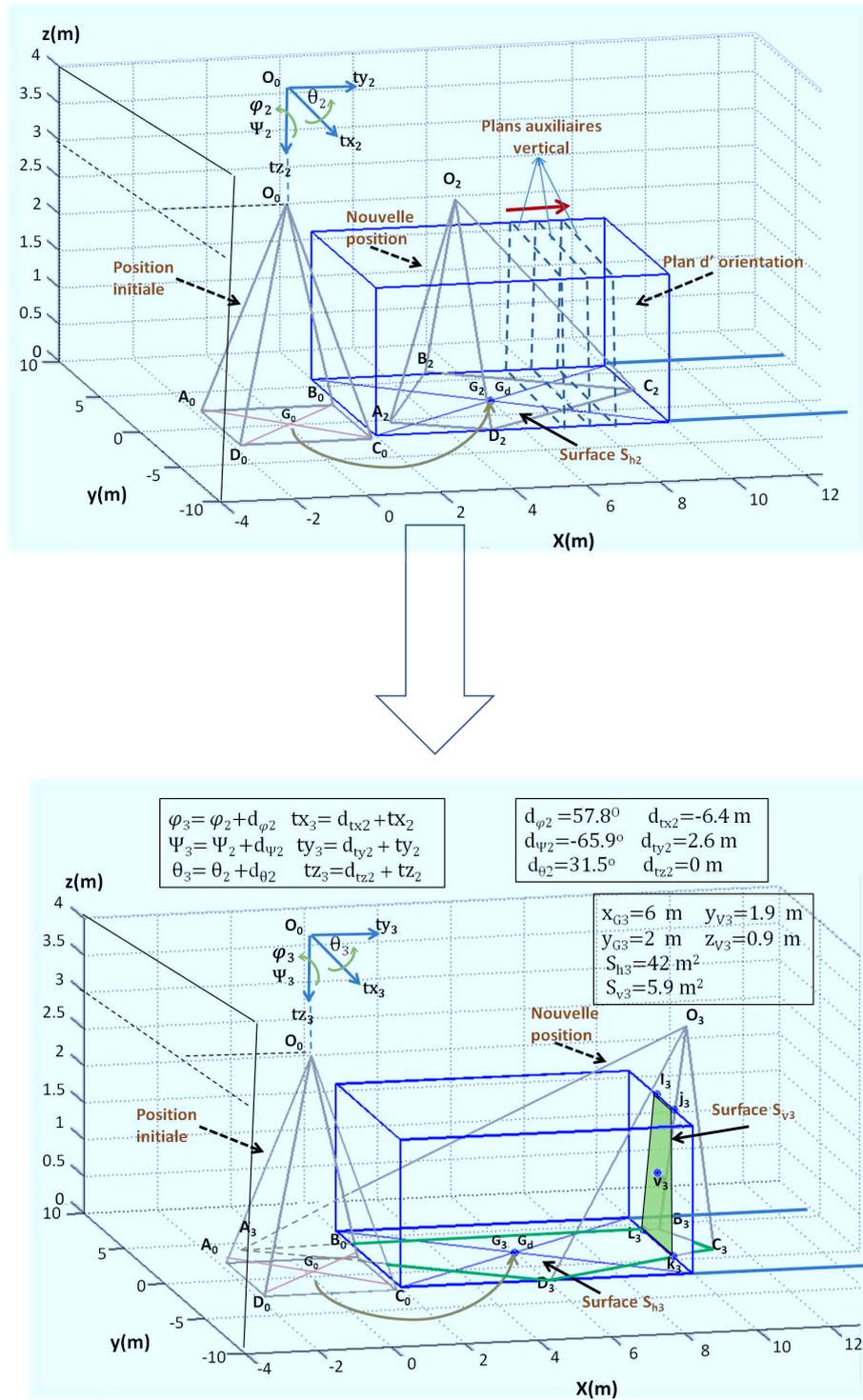


FIGURE 2.18 – Phase d'optimisation par rapport au plan vertical d'orientation : orientation du champ de vue d'une caméra.

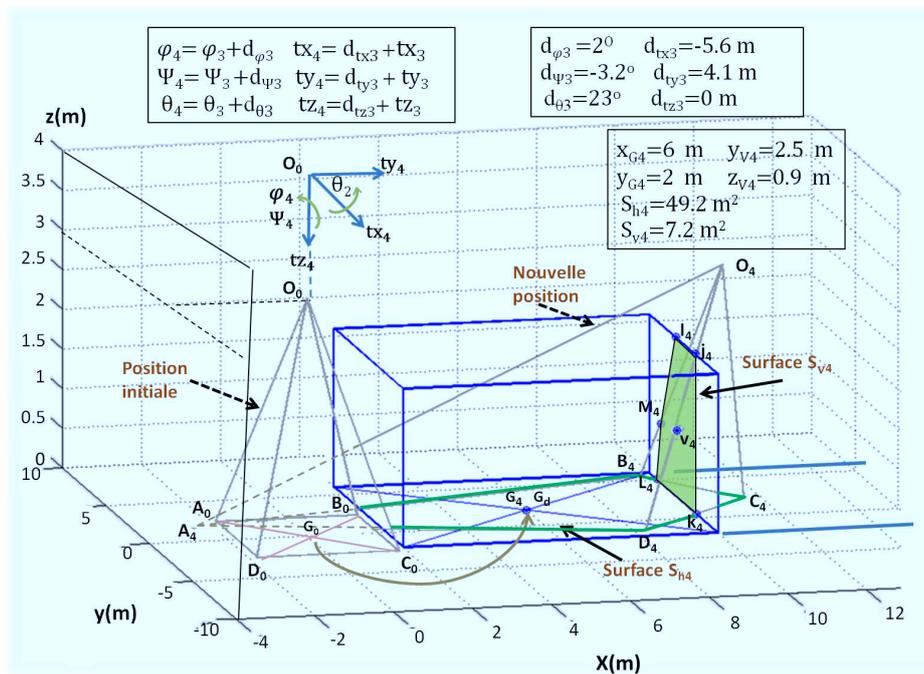


FIGURE 2.19 – Phase d’optimisation par rapport au plan horizontal et au plan vertical de projection : optimisation des surfaces horizontales et verticales (totales et communes) des champs de vue des deux caméras.

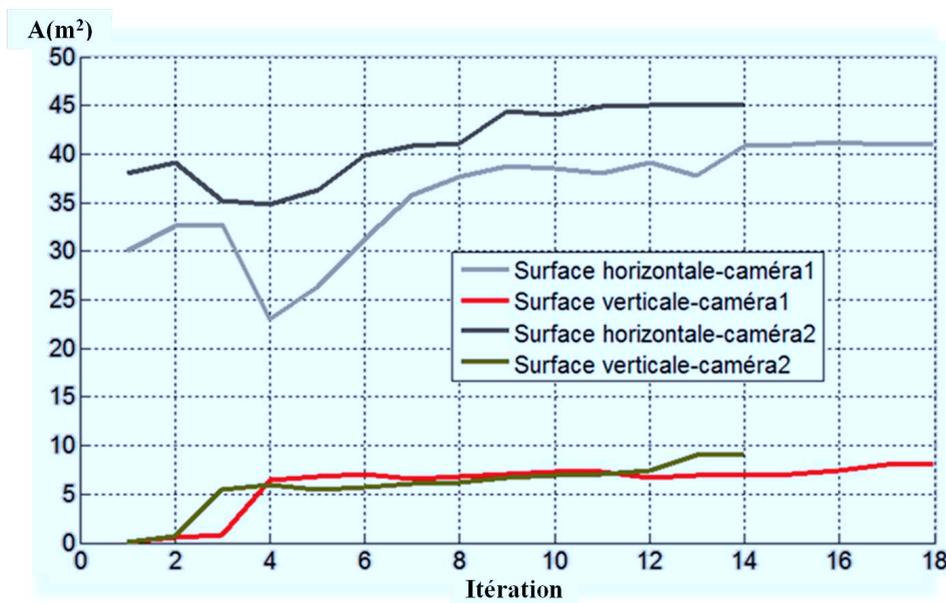


FIGURE 2.20 – Evolution des surfaces durant le processus d’optimisation relative au plan d’orientation.

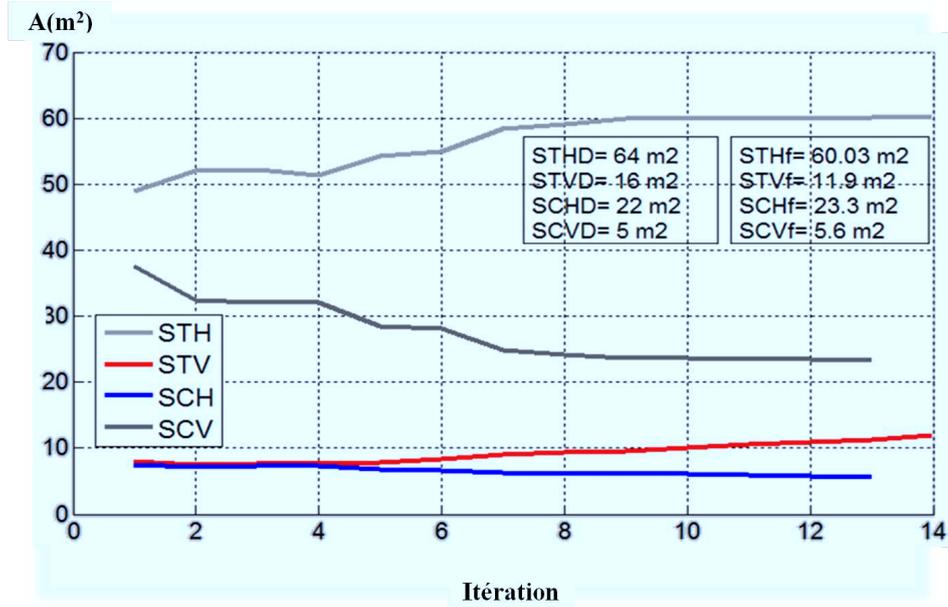


FIGURE 2.21 – Evolution des surfaces durant le processus d'optimisation relative au plan horizontal et au plan vertical de projection.

2.6 Conclusion

Dans ce chapitre, nous nous sommes intéressés au problème du calibrage extrinsèque des capteurs vidéo vis-à-vis d'une scène à surveiller. Nous avons développé une méthode d'optimisation originale permettant de déterminer la position géométrique optimale du système vidéo multi-vues, destiné à la perception de l'environnement d'un passage à niveau. Il s'agit de déterminer la position et l'orientation des caméras pour obtenir des champs de vue (commun et total) optimaux. Cette méthode modélise le champ de vue des caméras dans l'espace afin de déterminer les paramètres permettant de le déplacer autour de la zone de surveillance. Basée sur la géométrie descriptive, la méthode proposée suit les champs de vue sur différents plans de projection au lieu de les suivre dans l'espace. Les propriétés géométriques extraites de ces plans de projections sont utilisées pour déplacer et orienter le champ de vue dans l'espace. Tous les paramètres (positions et orientations des caméras) nécessaires pour fixer les capteurs vidéos sur un terrain réel (ou virtuel) sont alors définis.

Cette méthode a été mise en oeuvre dans le cadre du développement d'un simulateur d'aide à l'étude des passages à niveau, notamment l'étude de positionnement de capteurs vidéo pour la surveillance de l'environnement de passages à niveau à partir d'un système de perception multi-vues.

La deuxième partie de la thèse est consacrée au développement d'un système de perception permettant l'évaluation de situations potentiellement dangereuses

aux passages à niveau. Au début de la thèse, nous avons orienté ces travaux vers l'utilisation d'un système stéréoscopique à base de deux caméras. Au regard des difficultés de mise en place d'un tel système dans les conditions de l'application visée (processus de calibrage, mise en correspondance des images acquises dans des conditions d'éclairage et de changement de luminosité différentes, ...) nous avons opté finalement pour un système de perception monoculaire.

Le chaîne de traitement de notre système d'évaluation de situations potentiellement dangereuses est composée de 3 étapes principales : 1) suivi d'objets en mouvement, 2) modélisation des trajectoires 2D des objets, 3) analyse des trajectoires 2D pour l'évaluation du danger qu'elles engendrent.

Dans le chapitre suivant, nous présentons une méthode permettant le suivi des objets en mouvement à partir d'une séquence d'images monoculaires. La méthode proposée commence par détecter puis séparer les objets en mouvement. Ce processus de suivi est ensuite mis en oeuvre grâce à une estimation robuste du flot optique.

Suivi d'objets en mouvement à partir d'une séquence d'images

Sommaire

3.1	Introduction	47
3.2	État de l'art sur le suivi d'objets en mouvement	49
3.2.1	Détection des objets	49
3.2.2	Séparation des objets	51
3.2.3	Suivi des objets	52
3.2.4	Méthode proposée	53
3.3	Détection et séparation des objets	53
3.3.1	Introduction	53
3.3.2	Processus de cadrage pour la détection des objets fixes et mobiles	54
3.3.3	Méthode de séparation des objets	55
3.4	Suivi dense des objets	58
3.4.1	Introduction	58
3.4.2	Suivi d'objets à partir d'une séquence d'images monoculaires	58
3.5	Résultats expérimentaux	69
3.5.1	Expérimentation 1 : Suivi de piétons	69
3.5.2	Expérimentation 2 : Evaluation avec des courbes gaussiennes	75
3.5.3	Expérimentation 3 : Suivi d'un objet dans une séquence d'images simulées	78
3.6	Conclusion	83

3.1 Introduction

Un système de vidéosurveillance se compose généralement d'un réseau de caméras dont la fonction première consiste à enregistrer et à transmettre des images vers un centre de contrôle ou bien simplement vers un opérateur humain [Dubois-Arber 2006].

Depuis quelques années, les systèmes de vidéosurveillance contribuent significativement dans le domaine de la sécurité [Monroe 2003]. Il s'agit de surveiller des lieux publics ou privés, observer une zone ou un processus continuellement, analyser un événement donné, . . .

L'opérateur humain étant incapable de maintenir son attention pendant de longues heures, l'introduction de la détection automatique d'événements dans la vidéosurveillance permet de répondre de manière efficace aux besoins de surveillance et de sécurité. Cela engendre des solutions techniques permettant de créer des systèmes automatiques ou semi-automatiques de vidéosurveillance. Ces systèmes doivent être robustes et applicables en temps réel. Ils doivent également disposer d'une procédure de calibrage des caméras (automatique ou semi-automatique) afin de maintenir les valeurs et la cohérence des différents paramètres en cas de déplacement accidentel ou décisionnel des capteurs optiques.

Un système de vidéosurveillance automatique doit être capable de :

- détecter et séparer les objets dans les séquences vidéo ;
- suivre les objets en mouvements ;
- analyser, évaluer et reconnaître les événements en situations définies.

La vidéosurveillance d'un passages à niveau (PN) [Foresti 1998][Sacchi 2000] à des fins de sécurité consiste à positionner un certain nombre de capteurs optiques surveillant l'enceinte du PN. Un système de vidéosurveillance de PN consiste alors à détecter et identifier des événements potentiellement dangereux qui peuvent survenir sur le PN. En effet, les objets empruntant les passages à niveau (voitures, piétons, deux roues, trains, ...) peuvent créer des situations dangereuses lorsqu'ils entrent en interaction les uns avec les autres ou avec l'infrastructure PN elle-même. Il s'agit donc de repérer dans les images ces différents objets, de les suivre et d'analyser " leur comportement " par rapport aux règles de sécurité inhérentes aux PN.

Dans ce chapitre, nous présentons une méthode robuste pour le suivi des objets en mouvement dans une séquence d'images monoculaires. Dans un premier temps, les objets sont extraits du fond en utilisant une technique basée sur l'analyse des images de référence et courante. Ce processus est basé sur deux étapes principales : une étape de cadrage pour la détection des objets, et une étape de séparation des objets.

Par la suite, pour pouvoir suivre les objets détectés, nous proposons un filtre basé sur le calcul du flot optique. Le principe consiste tout d'abord à déterminer le flot optique des points de Harris qui présentent une invariance en termes de translation, rotation et changement d'échelle. En supposant qu'il suit une distribution gaussienne dont les paramètres sont extraits du flot optique des points de Harris, le flot optique des autres pixels de l'objet est ensuite estimé par propagation du flot optique des points de Harris. Le flot optique obtenu est enfin rectifié en utilisant un modèle de filtre de Kalman.

Ce chapitre est structuré de la manière suivante. La section 2 présente un état de l'art relatif au suivi d'objets en mouvement. Dans la section 3, nous présentons notre méthode de détection et de séparation d'objets. L'étape de suivi des objets en mouvement est décrite dans la section 4. Avant de conclure, la section 5 présente les résultats expérimentaux de l'approche proposée.

3.2 État de l'art sur le suivi d'objets en mouvement

En se référant à la littérature, les modèles de suivi d'objets sont généralement basés sur les trois étapes suivantes [Leibe 2008][Nan 2011] : détection, séparation et suivi. Ces trois étapes principales font appel à la mise en place d'algorithmes adaptés à l'application considérée.

3.2.1 Détection des objets

Ce paragraphe a pour but de fournir un bref état de l'art sur les méthodes de détection d'objets en mouvement basées sur une image de référence. Pour construire cette image de référence, aussi appelée modèle de fond, une séquence d'observations est utilisée pour faire apprendre au système de détection des caractéristiques telles que l'intensité lumineuse, la couleur, ou certaines caractéristiques de texture. Le processus de détection consiste alors à classer chaque pixel de manière indépendante dans les classes fond et objet, en fonction des observations courantes. Les méthodes présentées ci-dessous nécessitent des périodes d'apprentissage plus ou moins longues.

3.2.1.1 Modèle à valeur médiane

Ce modèle adaptatif a été proposé par Greenhill [Greenhill 2004] pour l'extraction d'objets en mouvement dans des conditions d'illumination dégradées. En se référant aux différents états de chaque pixel durant une période d'apprentissage, un modèle du fond est ainsi élaboré. Chaque pixel du modèle correspond à la valeur (intensité, couleur, ...) minimisant la distance avec chacune des valeurs possibles. Le modèle du fond est constamment mis à jour pour chaque nouvelle image de telle sorte qu'un vecteur V de valeurs médianes est construit sur les $N/2$ dernières images, sachant que N correspond au nombre d'images utilisées durant l'étape d'apprentissage. Une distance notée D correspond à la différence entre la valeur minimale et la valeur maximale du vecteur V . La classification fond/objet se fait simplement par seuillage sur la distance entre la valeur du pixel à classer et son correspondant dans le modèle du fond. Afin de tenir compte des changements d'illumination, le seuil tient compte de la distance D et d'un facteur de corrélation.

3.2.1.2 Codebook

Dans [Kim 2005], Kim propose de modéliser le fond en se basant sur une séquence d'observations de la scène sans mouvement durant une période de plusieurs minutes. Les occurrences similaires d'un pixel donné sont mises sous forme d'un vecteur appelé Codeword. Chacun de ces Codewords contient des informations comme la valeur moyenne du pixel, les limites d'intensité ou encore son nombre d'occurrences. Deux Codewords sont différents si la distance, dans l'espace vectoriel, dépasse un certain seuil. Un Codebook, qui est un ensemble de Codewords, est construit pour chaque pixel pendant la phase d'apprentissage. La classification

fond/objet d'un pixel donné, sur une nouvelle image dans la phase de recherche de mouvement, se base sur une simple différence entre sa valeur courante et chacun des Codewords correspondants.

3.2.1.3 Distributions gaussiennes

Dans [Wren 1997], Wren et al. proposent d'adapter le seuil en chaque pixel en modélisant la distribution des intensités lumineuses de chaque pixel par une loi Gaussienne. Ce modèle peut s'adapter à des changements lents de la scène, tels que des changements d'illumination progressifs, en se mettant à jour de manière récursive grâce à un filtre adaptatif. Différentes extensions de ce modèle ont été développées en changeant les caractéristiques utilisées au niveau pixel. Gordon et al. [Gordon 1999] représentent dans chaque pixel par un vecteur composé de quatre composantes indépendantes qui sont les trois composantes couleur et la profondeur, déterminé à l'aide de deux caméras.

3.2.1.4 Mélange de gaussiennes

Une amélioration du modèle précédent consiste à modéliser l'évolution temporelle des pixels par un mélange de Gaussiennes. Stauffer et Grimson, dans [Stauffer 1999] et [Stauffer 2000], modélisent la couleur de chaque pixel par un mélange de n gaussiennes. Le nombre de gaussiennes doit être ajusté en fonction de la complexité de la scène observée. Afin de simplifier les calculs, la matrice de covariance est supposée diagonale, ce qui revient à considérer les trois canaux couleur de manière indépendante. Le modèle de mélange de gaussiennes est mis à jour à chaque nouvelle image en utilisant l'algorithme des k-means. Harville et al. [Harville 2001] proposent d'utiliser le modèle de mélange de gaussiennes dans un espace combinant la profondeur et l'espace couleur YUV. Ils améliorent la méthode en modulant le taux d'apprentissage du modèle en fonction de l'activité de la scène. Ce taux est un paramètre qui permet à l'algorithme de réagir rapidement à l'apparition d'une perturbation. Toutefois, elle est très sensible aux variations brusques du fond, tel que le changement global d'illumination. Ainsi, un compromis doit être trouvé entre le taux d'apprentissage et l'adaptation aux changements du fond.

Un faible taux d'apprentissage produira énormément de fausses détections lors d'un changement d'éclairage, tandis qu'un fort taux d'apprentissage aura tendance à inclure les objets en mouvement dans le modèle du fond.

3.2.1.5 Modèle de Markov

Pour considérer l'évolution temporelle des pixels, l'ordre d'arrivée des niveaux de gris en ce pixel est une information utile. Une solution consiste alors à modéliser l'évolution des niveaux de gris en chaque pixel par une chaîne de Markov. Rittscher et al. [Rittscher 2000] emploient une chaîne de Markov à 3 états : objet, fond et ombre. Tous les paramètres de la chaîne, qui sont la probabilité initiale, la probabilité de transition et la probabilité d'observation, sont estimés hors ligne, sur

une séquence d'apprentissage. Stenger et al. [Stenger 2001] ont proposé une amélioration, puisque après un court apprentissage préalable, le modèle de la chaîne et ses paramètres continuent à être mis à jour. Cette mise à jour, réalisée durant la période de détection, permet de mieux gérer les états non-stationnaires dus par exemple à de brusques changements d'illumination.

3.2.1.6 Analyse en composantes indépendantes (ACI)

Une approche très prometteuse et récemment développée par N. Fakhfakh dans [Fakhfakh 2011a] consiste à utiliser une analyse en composantes indépendantes [Herault 1986] pour détecter les objets dans un fond de scène.

Cette technique a été récemment exploitée pour l'extraction des régions en mouvement à partir d'une séquence d'images [Tsai 2009][Zhang 2006]. L'ACI se base sur les méthodes traitant des observations vectorielles afin d'en extraire des composantes linéaires qui soient aussi indépendantes que possible [Oja 2000]. Elle permet la séparation d'un signal source complexe, qui est un mélange de signaux inconnus a priori, en une combinaison de plusieurs signaux, dits signaux estimés. L'ACI est définie par un modèle génératif permettant l'estimation d'un ensemble de signaux à partir d'un signal donné. Le signal observé est supposé être une combinaison linéaire des signaux à estimer. Ce modèle d'ACI peut être résolu en définissant une fonction d'énergie et un algorithme d'optimisation comme par exemple FastACI [Oja 2000]. Il s'agit d'un algorithme itératif permettant une convergence vers des paramètres optimaux. Pour initialiser l'algorithme FastACI, une simple soustraction entre l'image et le fond [Panahi 2008] est suffisante pour déclencher l'algorithme. Cet algorithme a permis l'utilisation de l'ACI dans plusieurs applications grâce à sa rapidité. L'algorithme se divise en deux étapes : une étape d'apprentissage et une étape de détection.

La première étape consiste à estimer le modèle de bruit ainsi que la matrice de séparation, notée W , obtenue par analyse en composantes indépendantes. La deuxième étape se divise en deux parties : la première partie consiste à approximer les régions affectées par du mouvement en utilisant la matrice de séparation estimée pendant l'étape d'apprentissage. La deuxième partie permet de raffiner les régions extraites en introduisant une étape de lissage, effectuée en minimisant une énergie dans un cadre de champs de Markov aléatoires.

3.2.2 Séparation des objets

Quelle que soit la méthode utilisée, l'étape de détection ne permet que de donner une indication sur les pixels affectés par du mouvement dans l'image et ne retourne qu'un seul masque binaire des pixels représentant les objets en mouvement. L'exploitation de la détection de pixels des objets en mouvement nécessite alors l'ajout d'une étape d'extraction d'un masque pour chaque objet.

Une des solutions fondamentales utilisées pour la séparation des objets détectés consiste à labelliser différentes classes de pixels connectés [Singh 2010]. Ces classes

représentant les différents objets en mouvement dans l'image sont alors séparés par la gestion d'une table d'équivalence [Suzuki 2003]. Une autre méthode intéressante consiste à associer une fonction d'énergie pour chaque objet segmenté [Boykov 2001]. Des algorithmes de flot-max/coupe-min [Boykov 2004] sont alors utilisés dans cette méthode pour achever un processus de minimisation d'énergie et pour extraire de l'image l'ensemble des classes qui représente l'ensemble des objets en mouvement. Le principal problème avec ces techniques est qu'elles ne marchent correctement que dans le cas d'objets colorés et texturés. On peut aussi faire appel à des méthodes d'association qui, à partir d'observations caractéristiques, les objets sont filtrés puis associés aux cibles existantes ou nouvelles [Lin 2004]. Les performances de ces méthodes dépendent de la distance entre les objets et leur taille. D'autres méthodes sont basées sur l'échantillonnage [Smith 2004]. Dans ces méthodes, un modèle d'observation pour chaque objet mobile est créé et le problème du suivi multi-objets est ensuite résolu avec le principe de l'exclusion probabiliste [MacCormick 1999]. Cette approche est intéressante mais nécessite le traitement de nombreux problèmes pour une séparation complète des objets. Le traitement des images multi spectrales et polarimétriques [Peng 2009] peut fournir des informations supplémentaires et spécifiques aux caractéristiques des objets détectés en mouvement. La fusion de ces propriétés polaires [Gevers 2003] permet une séparation plus robuste des objets mobiles. Ces techniques multi-spectrales sont plus efficaces pour le traitement de scènes d'intérieur que pour le traitement de scènes d'extérieur. En effet, les conditions d'éclairage jouent un rôle très important lors de la séparation des objets.

3.2.3 Suivi des objets

Concernant le processus de suivi d'objets, deux approches principales ont montré leur efficacité et leurs performances dans les systèmes de suivi vidéo : l'approche par localisation (Les méthodes de " Contour ", " Blob ", " Kernel " ...) et l'approche par filtrage (Les filtres particuliers, les filtres de Kalman ...). En particulier, François [Francois 2004] propose un algorithme de suivi basé sur la segmentation de l'objet dans des images successives. Li [Li 2010] utilise un filtre de Kalman adaptatif combiné à l'algorithme de Mean Shift, pour suivre les cibles candidates à travers leur centre, grâce à une distribution couleur normalisée. Blake [Blake 1998] propose de suivre les objets en utilisant les contours actifs et une fusion temporelle basée sur le filtre de Kalman. Yang [Yang 2005] propose de calculer une distribution état-espace des cibles en utilisant un filtre à particules. Yilmaz [Yilmaz 2011] développe un système de noyau asymétrique " asymmetric Kernel " avec une nouvelle technique de sélection de l'échelle et de l'orientation de ce noyau (kernel). Le but de cette technique est de représenter avec performance la forme de l'objet détecté. Zhang [Zhang 2008] s'appuie sur des méthodes d'association des données et sur des techniques d'optimisation dans un réseau de flots optiques, pour trouver les trajectoires des objets détectés.

Les méthodes présentées regroupent plusieurs idées intéressantes pour la détection et le suivi des objets en mouvement. Cependant, aucune parmi ces méthodes

n'est apte à suivre localement les pixels des objets dans des environnements réels et complexes.

3.2.4 Méthode proposée

En adoptant la chaîne de traitement « détection, séparation, suivi », nous proposons dans ce chapitre une méthode de suivi d'objets en mouvement dans une zone de surveillance à partir d'une caméra fixe. Nous avons développé différents algorithmes permettant la détection, la séparation et le suivi d'objets mobiles. Nous avons exploité des séquences vidéo, dans des conditions réelles, pour suivre de manière précise et robuste le plus grand nombre possible de pixels des objet en mouvement. Le but de la méthode est la détection et le suivi local de pixels représentant les objets en mouvement dans des scènes réelles et relativement complexes.

Nous avons mis en place une technique de détection basée sur la différence inter-images. Nous utilisons cette technique pour indiquer les pixels en mouvement dans l'image et comme étape de prétraitement pour séparer les objets en mouvement. Durant cette étape, nous représentons les pixels connectés dans l'image à l'intérieur de plusieurs cadres. Nous supposons ensuite que les dimensions de ces cadres évoluent d'une manière continue lors de l'apparition des objets dans la zone de surveillance. A l'étape suivante, nous prenons en compte le nombre d'objets possibles détectés dans l'image. A chaque itération de l'algorithme, nous attribuons le pixel observé à l'un des objets détectés. Nous définissons alors un vecteur d'énergie pour chaque pixel à assigner afin de distinguer les pixels en mouvement. La dernière étape consiste à suivre les pixels des objets détectés à l'étape précédente. Etant donné un objet détecté, nous effectuons une extraction des points de Harris. Nous calculons ensuite le flot optique des points de Harris en utilisant la technique itérative de Lucas-Kanade [Bouguet 2002]. Nous supposons ensuite que le flot optique des points de l'objet suit une distribution gaussienne dont les paramètres sont estimés à partir du flot optique des points de Harris. Le flot optique est enfin rectifié grâce à un filtre de Kalman adaptatif. Pour montrer la faisabilité de l'approche proposée, différents tests sur des séquences d'images réelles et différentes évaluations sont alors présentées dans ce chapitre.

3.3 Détection et séparation des objets

3.3.1 Introduction

La détection des objets en mouvement joue un rôle important pour la phase de suivi. Elle permet de définir la position initiale des objets, et par conséquent facilite le processus de suivi.

Un certain nombre de méthodes de la littérature pour la détection des objets ont été présentées précédemment. La plupart d'entre elles présentent des difficultés à s'adapter aux problèmes de changement de luminosité, aux problèmes liés aux occultations, inhérents aux environnements extérieurs où l'on cherche à tester notre

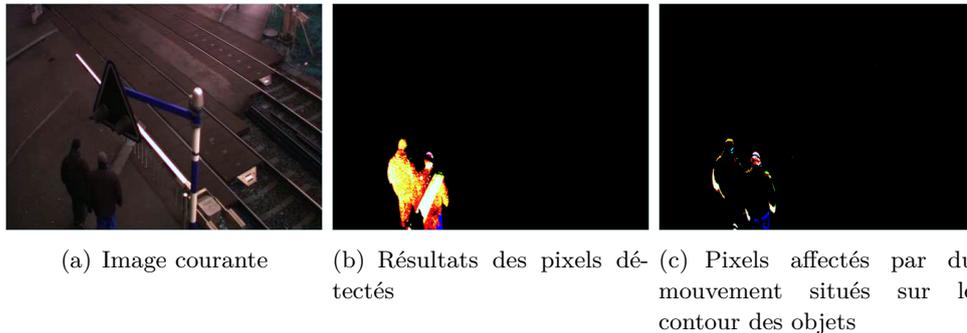


FIGURE 3.1 – Détection des pixels affectés par du mouvement.

approche. Pour résoudre ces problèmes, nous avons développé un nouvel algorithme de détection des objets en mouvement bien adapté aux environnements réels. Cet algorithme comporte deux étapes principales :

- La première étape est basée sur un processus de cadrage et consiste à déterminer le nombre de cibles possibles dans l'image. L'extraction des régions connexes composées de pixels en mouvement permet de faciliter le traitement ultérieur.
- La deuxième étape est basée sur un processus de séparation et consiste à distinguer les objets connectés entre eux et à déterminer la taille finale de chaque objet détecté. Cette procédure de séparation permet de répondre aussi à de nombreux problèmes comme : la représentation des formes des objets détectés, le changement d'orientation et la variabilité de la taille des objets en mouvement.

3.3.2 Processus de cadrage pour la détection des objets fixes et mobiles

Se concentrer uniquement sur l'extraction des régions connexes composées de pixels affectés par du mouvement permet de réduire le temps de traitement de l'algorithme et de faciliter le processus de détection et de séparation des cibles mobiles dans l'image. Nous définissons une cible dans l'image par un ensemble de pixels proches les uns des autres et affectés par du mouvement. Ces pixels représentent la cible se trouvant à l'intérieur d'un cadre rectangulaire. A la lumière de cette définition, nous procédons selon deux étapes essentielles.

La première étape a pour but de détecter tous les pixels des objets en mouvement. Elle consiste à effectuer une soustraction entre l'image courante et une image de référence (Fig. 3.1(b)). De la même manière, il est possible de soustraire deux images successives pour obtenir les pixels affectés par du mouvement situés sur les contours des objets dans l'image courante (Fig. 3.1(c)).

La deuxième étape a pour but de déterminer le nombre de cibles possibles avec leurs cadres correspondants dans l'image courante. Pour réaliser cette tâche, nous prenons en compte la connectivité entre les pixels affectés par du mouvement sur les contours des objets. Un cadre est alors associé à chaque groupe de pixels connectés. A l'intérieur d'un même cadre, les pixels affectés par du mouvement situés sur le

contour des objets sont complétés par les autres pixels affectés par du mouvement. L'ensemble des pixels intérieurs et aux bord des objets sont ainsi regroupés dans le même cadre. Chaque pixel en mouvement à l'intérieur d'un cadre donné peut représenter soit une nouvelle cible, soit une cible existante. Cela dépend fortement de l'intersection entre le cadre dans l'image courante et toutes les cibles existantes dans l'image précédente. Tous les cadres dans l'image courante ne possédant aucune intersection avec les cibles existantes sont alors considérés comme des nouvelles cibles potentielles. La figure 3.2 montre quelques résultats d'extraction de cibles entre deux images successives.

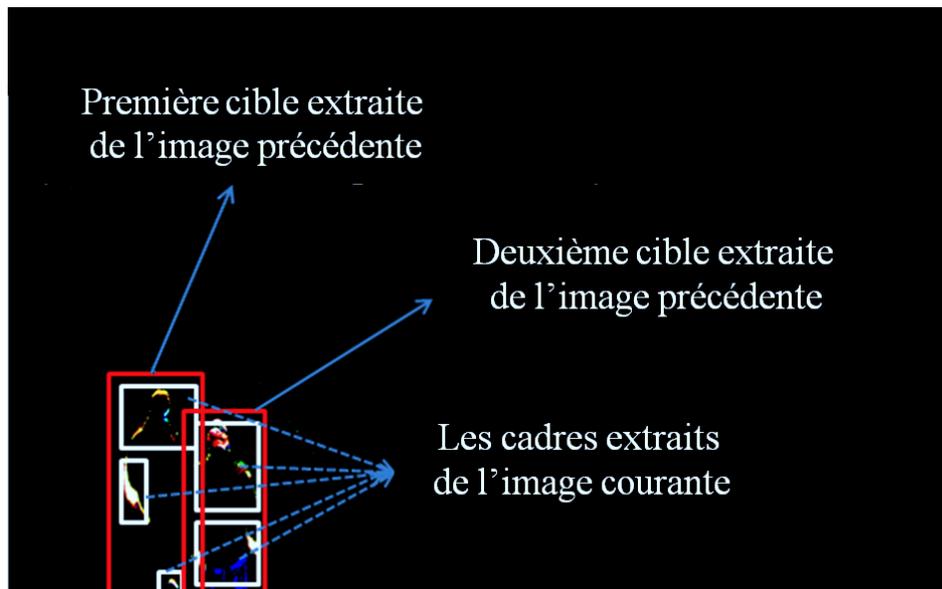


FIGURE 3.2 – Extraction de cibles entre deux images successives.

3.3.3 Méthode de séparation des objets

Après avoir déterminé le nombre de cibles possibles dans la zone de surveillance, le but de l'étape suivante est de mettre en oeuvre une stratégie de séparation si un ou plusieurs cadres obtenus possèdent une intersection avec plusieurs cibles existantes. Pour ce faire, deux vecteurs d'énergie sont définis. Le premier est relatif à chaque cible. Au début du traitement de chaque image de la séquence, ce vecteur est initialisé à zéro. Il est ensuite mis à jour durant le traitement de l'image en question. Notons par E_{target}^i le vecteur d'énergie associé à la cible numéro i . Le deuxième vecteur d'énergie est défini pour chaque pixel à assigner. Ce vecteur d'énergie, calculé par rapport à chaque cible, dépend de plusieurs caractéristiques : flot optique, gradient, intensité et distance. Ce vecteur d'énergie est utilisé pour comparer différents critères permettant de distinguer les cibles entre elles. Etant donné un pixel de position (x, y) dans l'image et une cible numéro i , le vecteur

d'énergie du pixel E_{pixel}^i est donné par :

$$E_{pixel}^i = [E_G^i, E_I^i, E_F^i, E_D^i]^T \quad (3.1)$$

$$E_G^i = W_i * \left(\Delta I_{(x,y)}^2 + \left(\vec{\nabla}(I_{(x,y)}) \cdot \hat{v}_i \right)^2 \right) \quad (3.2)$$

$$E_I^i = W_i * I_{(x,y)}^2 \quad (3.3)$$

$$E_F^i = W_i * F_{(x,y)}^2 \quad (3.4)$$

$$E_D^i = W_i * \left(D_{(x,y)}^i \right)^2 \quad (3.5)$$

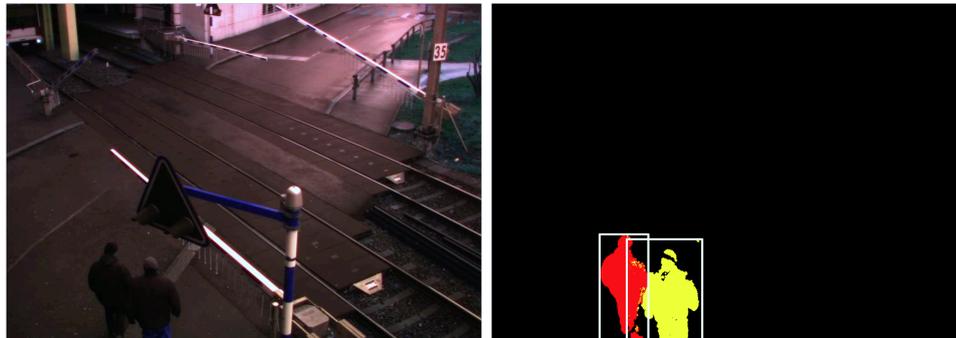
$$W_i = e^{-\left(\frac{D_{(x,y)}^i}{\sigma} \right)^2} \quad \sigma = \frac{d_{i_1}^{i_2}(x,y)}{2} \quad (3.6)$$

où $E_G^i, E_I^i, E_F^i, E_D^i$ sont respectivement les énergies relatives au gradient, à l'intensité, au flot optique et à la distance. $I_{(x,y)}$, $\vec{\nabla}(I_{(x,y)})$ et $F_{(x,y)}$ sont respectivement la fonction intensité de l'image, le gradient de l'image et le flot optique (calculé par l'algorithme de Lucas-Kanade [Bouguet 2002]) au point (x, y) . $\Delta I_{(x,y)}$ est le Laplacien de l'image au point (x, y) . \hat{v}_i est le vecteur unitaire d'origine le point (x, y) et dirigé vers le centre de la cible numéro i . $D_{(x,y)}^i$ représente la distance entre le pixel (x, y) et la ligne de fuite verticale de l'image, passant par le centre de la cible numéro i . Exprimé sous la forme d'une fonction Gaussienne centrée, W_i est un coefficient de pondération relatif à la cible numéro i . La déviation σ de la fonction Gaussienne est donnée par $d_{i_1}^{i_2}(x, y)$, qui représente la distance entre les centres des deux cibles les plus proches au pixel (x, y) .

Afin d'assigner le pixel (x, y) à une cible, les vecteurs d'énergies E_{pixel}^i et E_{target}^i sont comparés composante par composante, pour chaque cible. Le pixel (x, y) est alors assigné à la cible qui met en évidence un nombre maximum de composantes proches. Lorsqu'une situation de conflit se présente, c'est-à-dire lorsque deux cibles ou plus répondent à ce critère de comparaison, alors le pixel (x, y) est assigné à la cible pour laquelle l'énergie de distance (4^{ème} composante du vecteur d'énergie) est la plus proche de celle du pixel. Notons par p le numéro de la cible à laquelle le pixel est assigné. Après cette procédure de l'assignation, le vecteur d'énergie de la cible numéro p est mis à jour de la manière suivante :

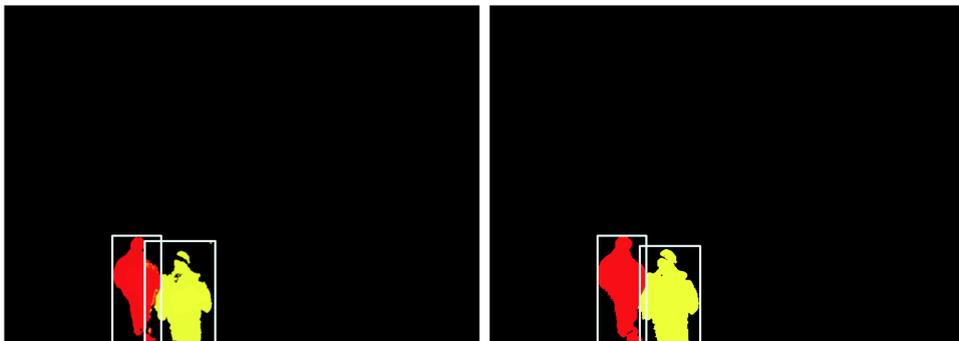
$$E_{target}^p = \left(\frac{N * E_{target}^p + E_{pixel}^i}{N + 1} \right) \quad (3.7)$$

où N est le nombre de pixels dans la cible numéro p , avant l'ajout du pixel (x, y) . A la fin du traitement de l'image en question, le résultat de séparation des objets est filtré puis classé en utilisant l'algorithme du mean shift [Cheng 1995]. Les figure 3.3(a), figure 3.3(b), figure 3.3(c) et figure 3.3(d) représentent respectivement l'image originale, le résultat de séparation des objets, la séparation des objets après



(a) Image originale (deux piétons en mouvement)

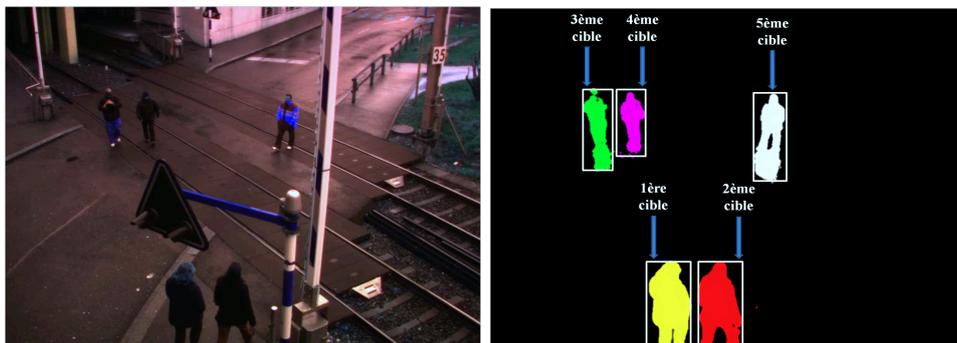
(b) Séparation d'objets



(c) Séparation d'objets après filtrage (mean-shift)

(d) Séparation d'objets après classification (mean-shift)

FIGURE 3.3 – Résultats de l'algorithme de séparation d'objets.



(a) Image originale (cinq piétons en mouvement)

(b) Résultat de séparation d'objets

FIGURE 3.4 – Séparation multi-objets.

filtrage par l'algorithme du mean shift, puis la séparation des objets après classification par l'algorithme du mean shift. Nous pouvons constater à partir des résultats de la figure 3.3(b) (avant filtrage) que la méthode proposée permet une très bonne séparation dans la zone critique autour de la frontière entre les deux objets, grâce à l'algorithme d'assignation basé sur la comparaison des vecteurs d'énergie. Cependant, des erreurs d'assignation apparaissent à l'intérieur des cibles. Ces erreurs peuvent être traitées et éliminées en utilisant un filtrage (Fig. 3.3(c)), puis une classification (Fig. 3.3(d)) basés sur l'algorithme du mean shift. La figure 3.4 montre le résultat de séparation de cinq piétons en mouvement.

3.4 Suivi dense des objets

3.4.1 Introduction

Le suivi d'objets en temps réel occupe une place très importante dans l'analyse des séquences vidéo. En effet, construire un algorithme robuste de suivi d'objets pour des applications temps réel demeure un challenge important. Parmi ces applications, on peut citer la surveillance, l'asservissement visuel, l'imagerie médicale. Nous présentons, dans cette partie, la méthode développée pour le suivi d'objets. Cette méthode suppose que les cibles sont extraites à l'issue de l'étape de séparation des objets. Etant donnée une cible, l'objectif est de calculer le flot optique de tous les pixels de cette cible. Pour ce faire, nous proposons une méthode basée sur la propagation du flot optique des points de Harris. Pour prendre en compte les effets du bruit, le flot optique estimé est ensuite filtré pour ne retenir que les pixels suivis de manière précise et robuste. Le principe consiste tout d'abord à déterminer le flot optique des points de Harris, qui présentent une invariance en termes de translation, rotation et changement d'échelle. Le flot optique des points de l'objet est supposé suivre une distribution gaussienne dont les paramètres sont extraits du flot optique des points de Harris. Suite à son initialisation par propagation, le flot optique rectifié est ensuite calculé en utilisant un filtre de Kalman et un algorithme d'optimisation basé sur une approche par différence d'intensité. La figure 3.5 montre schématiquement les différentes étapes de la méthode proposée.

3.4.2 Suivi d'objets à partir d'une séquence d'images monoculaires

3.4.2.1 Suivi des points d'intérêt de Harris

La détection de points d'intérêt est une étape préliminaire à de nombreux processus de vision par ordinateur. En général, ces points sont choisis car ils ont la propriété d'être parmi les points les plus stables dans les images [Shi 1994]. Leur détection est basée sur une analyse locale de l'image. Celle-ci consiste à retenir les points qui présentent une double discontinuité (dans les deux directions principales) de la fonction intensité de l'image. Plusieurs techniques de détection de points d'intérêt ont été proposées dans la littérature. Elles se caractérisent généralement par le

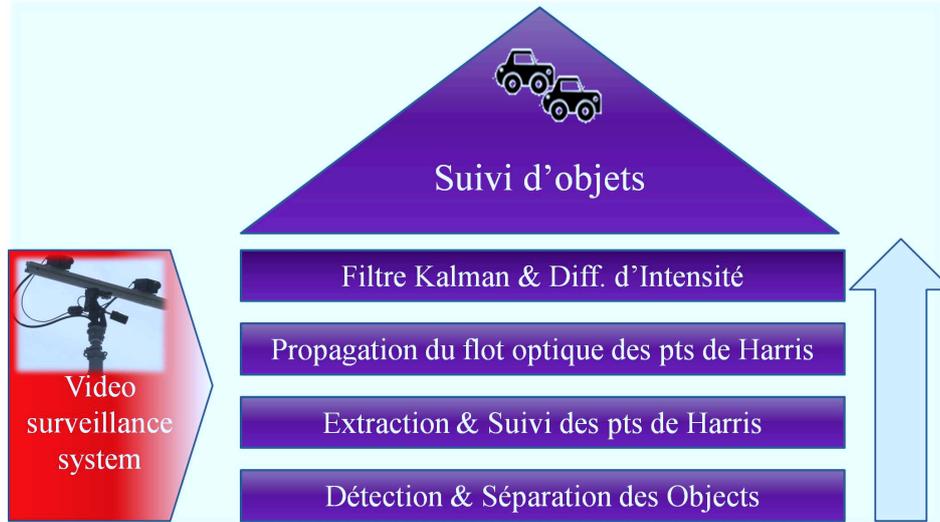


FIGURE 3.5 – Méthodologie de suivi d'objets.

type de l'opérateur de dérivation utilisé [Schmid 2000]. Dans ce travail, nous avons choisi le détecteur de Harris [Harris 1988][Derpanis 2004].

Pour suivre les points de Harris, nous avons choisi d'utiliser l'algorithme itératif du flot optique de Lukas-Kanade [Bouguet 2002][Tamgade 2009]. Soit I_t et I_{t+1} deux images consécutives. Supposons qu'un objet est détecté à l'image I_t et que nous disposons des points de Harris sur cet objet. Le but est de trouver la position de ces points à l'instant $t + 1$. Soit (u_t, v_t) un point de Harris extrait à l'instant t . La position de ce pixel à l'instant $t + 1$ peut s'exprimer alors de la manière suivante :

$$u_{t+1} = u_t + d_u \quad (3.8)$$

$$v_{t+1} = v_t + d_v \quad (3.9)$$

où (d_u, d_v) est le vecteur déplacement calculé en pixels/seconde entre les instants t et $t + 1$. Nous définissons la fonction d'erreur $\varepsilon(d_u, d_v)$ sur une fenêtre centrée sur le pixel (u_t, v_t) :

$$\varepsilon(d_u, d_v) = \sum_{u_t-w_u}^{u_t+w_u} \sum_{v_t-w_v}^{v_t+w_v} (I_t(u_t, v_t) - I_{t+1}(u_t + d_u, v_t + d_v))^2 \quad (3.10)$$

où (w_u, w_v) est la taille de la fenêtre centrée sur le point (u_t, v_t) . La méthode itérative de Lukas-Kanade consiste alors à calculer le flot optique (d_u, d_v) en minimisant la fonction d'erreur $\varepsilon(d_u, d_v)$ [Tamgade 2009]. La figure 3.6 montre un exemple de suivi des points de Harris sur un objet en mouvement. L'image I_t (à gauche) illustre les

points de Harris à l'instant t . L'image I_{t+1} (à droite) montre la position de ces points, calculée à l'instant $t + 1$.



FIGURE 3.6 – Suivi des points de Harris.

3.4.2.2 Propagation gaussienne du flot optique des points de Harris

Le suivi d'un nombre maximum de pixels assurera un suivi plus précis et efficace d'un objet extrait de la zone de surveillance. Cependant, le problème principal qui se pose ici est le bruit porté par le flot optique des pixels de l'objet détecté. L'objectif de cette partie est de proposer une solution pour calculer le flot optique de tous les points de l'objet. Nous partons de l'hypothèse que les points de Harris représentent l'information la plus stable par rapport au déplacement de l'objet dans l'image. Nous proposons alors de propager ce flot optique aux points restants de l'objet. Pour ce faire, nous supposons que le flot optique de tous les pixels suit une distribution gaussienne. A la lumière de cette hypothèse, nous proposons deux types de distribution de flot optique des points de l'objet. Dans la partie expérimentation, nous verrons l'influence de chacune de ces deux propositions sur les résultats de distribution du flot optique.

3.4.2.2.1 Première proposition Dans cette proposition, nous considérons que le flot optique des points de Harris est stable et que le flot optique de tous les points de l'objet est distribué normalement. La valeur moyenne et l'écart type du flot optique des point de Harris sont les paramètres de cette distribution normale, donnée par :

$$P(ft(x)) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{ft(x) - m_x}{\sigma_x} \right)^2} \quad (3.11)$$

où $P(ft(x))$ exprime la distribution du flot optique $ft(x)$ au point considéré sur l'axe x de l'image (on utilise la même expression pour l'axe y), m_x est le flot optique moyen des points de Harris sur l'axe x . L'écart type σ_x est calculé à partir de la

combinaison de σ_{hx} , qui représente l'écart type du flot optique des points de Harris sur l'axe x , et l'écart type d_{σ_x} d'un bruit inconnu, mais limité dans un intervalle connu :

$$\sigma_x = d_{\sigma_x} + \sigma_{hx} \quad (3.12)$$

La déviation d_{σ_x} est fixée expérimentalement à 4 pixels.

On peut alors définir le domaine de recherche du flot optique d'un pixel (x, y) de l'objet par un cercle de centre (m_x, m_y) et de rayon σ égal au module du vecteur (σ_x, σ_y) (Fig. 3.7).

3.4.2.2 Deuxième proposition Dans la deuxième proposition, nous procédons selon le même principe de propagation utilisé dans la première proposition en remplaçant cette fois la valeur moyenne m_x par le flot optique F_x^i du point de Harris numéro i le plus proche (en terme de mesure de distance dans l'image) du pixel considéré (idem pour l'axe y) :

$$P(ft(x)) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{ft(x) - F_x^i}{\sigma_x} \right)^2} \quad (3.13)$$

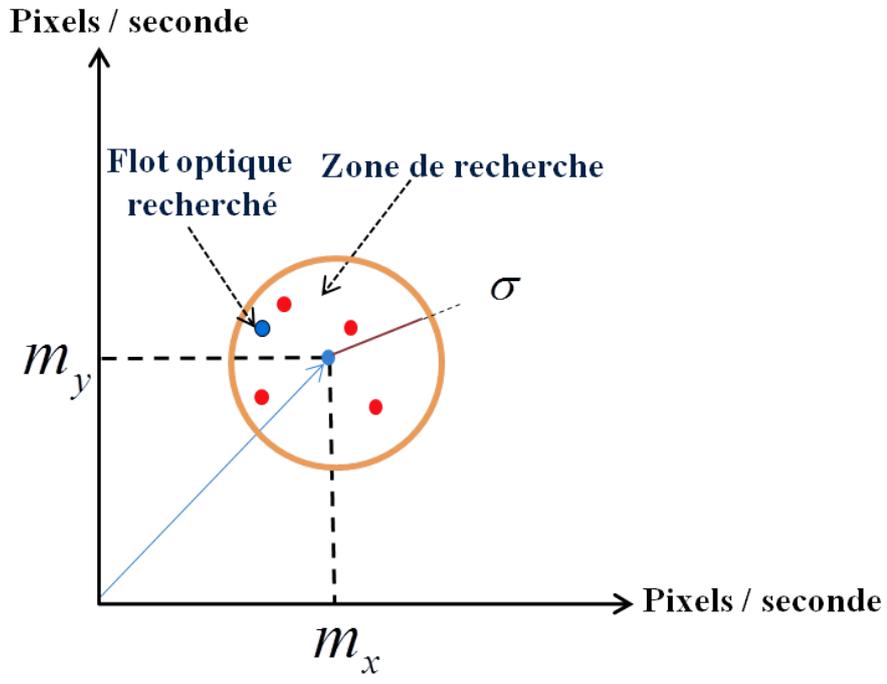
Contrairement à la première proposition où le flot optique des pixels de l'objet suit une seule distribution gaussienne, la deuxième proposition permet d'avoir une distribution gaussienne du flot optique pour chaque ensemble de pixels défini par un point de Harris. Un ensemble lié à un point de Harris est composé des pixels les plus proches à ce point.

On peut alors définir, avec cette proposition, le domaine de recherche du flot optique pour chaque pixel (x, y) de l'objet par un cercle de centre (F_x^i, F_y^i) et de rayon σ égal au module du vecteur (σ_x, σ_y) (Fig. 3.8).

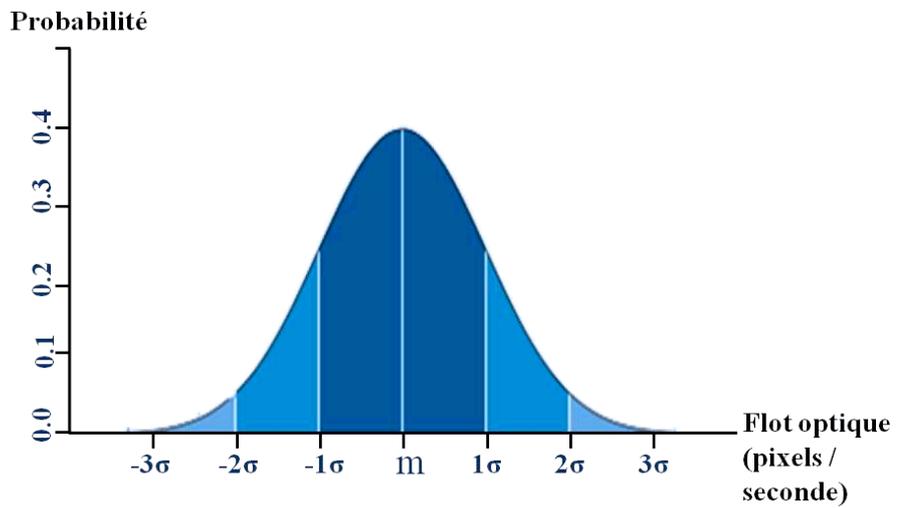
3.4.2.3 Rectification du flot optique par un modèle de filtre de Kalman itératif

Le filtre de Kalman est l'un des filtres les plus utilisés en statistique pour rectifier les mesures d'un modèle. Pour faire face au problème du bruit présent dans le flot optique estimé par propagation gaussienne, un filtre de Kalman itératif est mis en oeuvre pour fournir plusieurs solutions. Seule la solution se trouvant dans le domaine de recherche, défini précédemment, et vérifiant une contrainte de similarité est retenue. En revanche, si aucune solution ne vérifie ces deux critères, le flot optique du pixel considéré est ignoré. Avant de commencer à décrire le modèle de filtre de Kalman proposé, nous commençons tout d'abord par présenter la contrainte de similarité utilisée.

3.4.2.3.1 Contrainte de similarité La contrainte de similarité que nous proposons correspond à une fonction qui mesure le changement d'intensité entre les positions courante p_t et la position précédente p_{t-1} du pixel considéré :

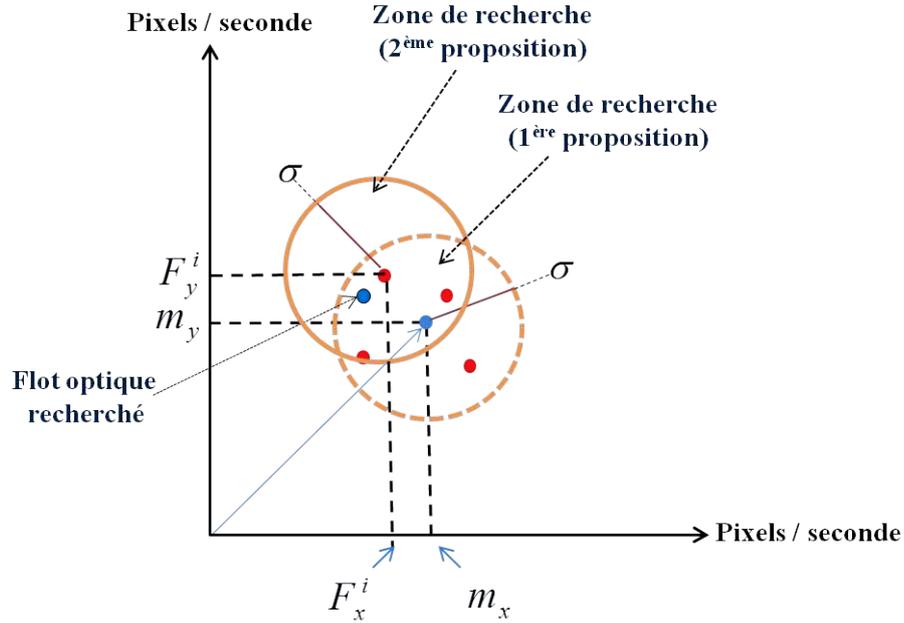


(a) Domaine de recherche



(b) Distribution gaussienne

FIGURE 3.7 – Zone de recherche gaussienne (1^{ère} proposition).

FIGURE 3.8 – Zone de recherche gaussienne (2^{ème} proposition).

$$S(p_t, p_{t-1}) = \sqrt{(r_{(p_t)} - r_{(p_{t-1})})^2 + (v_{(p_t)} - v_{(p_{t-1})})^2 + (b_{(p_t)} - b_{(p_{t-1})})^2} \quad (3.14)$$

avec :

$$r_{(p_i)} = \frac{R_{(p_i)}}{R_{(p_i)} + V_{(p_i)} + B_{(p_i)}}$$

$$v_{(p_i)} = \frac{V_{(p_i)}}{R_{(p_i)} + V_{(p_i)} + B_{(p_i)}}$$

$$b_{(p_i)} = \frac{B_{(p_i)}}{R_{(p_i)} + V_{(p_i)} + B_{(p_i)}}$$

où $R_{(p_i)}$, $V_{(p_i)}$, $B_{(p_i)}$ représentent respectivement les couleurs d'intensité suivant les composantes rouge, vert et bleu de l'espace RVB.

Une correction est validée si la valeur de la fonction de similarité est inférieure à un seuil fixé expérimentalement à 0.15, sachant que la fonction de changement de couleur normalisée varie entre 0 et 1.

3.4.2.3.2 Modèle de filtre de Kalman proposé Pour assurer un suivi robuste des pixels de l'objet, un filtre de Kalman itératif est mis en oeuvre afin de rectifier le flot optique de chaque pixel. Rappelons que le flot optique initial est estimé par propagation gaussienne à partir du flot optique des points de Harris. Le

modèle de filtre de Kalman proposé produit l'état corrigé X_t à partir de deux entrées, l'une représentant l'état précédent du système X_{t-1} et l'autre est la nouvelle mesure du système Y_t :

$$X_{t-1} = [p_x^{t-1}, p_y^{t-1}, ft_x^{t-1}, ft_y^{t-1}]^T \quad (3.15)$$

$$Y_t = [p_x^t, p_y^t, Ft_x^{t-1}, Ft_y^{t-1}]^T \quad (3.16)$$

où p_x^{t-1} et p_y^{t-1} sont les coordonnées du pixel à l'instant $t - 1$. p_x^t et p_y^t sont les coordonnées du pixel mesurées à l'instant t . ft_x^{t-1} et ft_y^{t-1} sont respectivement les flots optiques horizontaux et verticaux à l'instant $t - 1$. Ft_x^{t-1} et Ft_y^{t-1} sont respectivement les flots optiques horizontaux et verticaux mesurés à l'instant $t - 1$.

En poursuivant l'idée de propagation du flot optique, nous considérons que l'état du système X_{t-1} et la mesure Y_t suivent la même loi gaussienne (eq.3.11). L'état X_{t-1} représente l'état à corriger. Il est estimé à partir du processus de propagation du flot optique des points de Harris à l'instant $t - 1$, en utilisant un générateur de nombres gaussiens (RNG) :

$$ft_x^{t-1} = RNG(m_x, \sigma_x) \quad ft_y^{t-1} = RNG(m_y, \sigma_y) \quad (3.17)$$

où m_x et σ_x (resp. m_y et σ_y) sont les paramètres de la distribution normale du flot optique sur l'axe x (resp. y) à l'instant $t - 1$.

L'état Y_t correspond à un état mesuré à partir d'une observation virtuelle. Il est calculé à partir du processus de propagation du flot optique à l'instant $t - 1$, comme suit :

$$Ft_x^{t-1} = RNG(m_x, \sigma_x) \quad Ft_y^{t-1} = RNG(m_y, \sigma_y) \quad (3.18)$$

$$p_x^t = p_x^{t-1} + Ft_x^{t-1}dt \quad p_y^t = p_y^{t-1} + Ft_y^{t-1}dt \quad (3.19)$$

Nous prédisons ensuite l'état à l'instant t par :

$$X_t = FX_{t-1} \quad (3.20)$$

avec :

$$F = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3.21)$$

où F est la matrice de transition et dt est le temps séparant deux images successives.

À l'instant t , nous supposons que toutes les mesures jusqu'à l'instant $t - 1$ sont connues. Nous associons à toute cette période la matrice de covariance de l'erreur

d'estimation P_{t-1} . L'initialisation de P_{t-1} dépend de la confiance dont on dispose sur les mesures durant cette période. Plus la connaissance sur ces mesures est faible, plus on augmente les termes à l'intérieur de P_{t-1} . L'erreur de prédiction à l'instant t est alors :

$$P_t = FP_{t-1}F^T + Q_t \quad (3.22)$$

où Q_t représente la matrice de covariance du bruit de l'état estimé. En cas de bruit blanc, Q_t ne dépend que du bruit de variance calculée sur l'accélération de l'objet.

L'étape suivante consiste à mettre à jour l'état estimé, connaissant la mesure à l'instant t :

$$K_t = P_t H_t^T (H_t P_t H_t^T + R_t)^{-1} \quad (3.23)$$

$$P_t = P_t - K_t H_t P_t$$

$$X_t = X_t + K_t (Y_t - H_t X_t)$$

où K_t est le gain du filtre de Kalman, H_t est la matrice d'observation, R_t est la matrice de covariance du bruit de mesure. En cas de bruit blanc, R_t dépend du bruit de variance calculée sur la mesure de position et de vitesse de l'objet.

3.4.2.3.3 Adaptation des paramètres Dans notre approche, nous considérons que chaque pixel appartenant à un objet en mouvement possède son propre flot optique et ses propres mesures. Par conséquent, les paramètres du filtre de Kalman doivent être adaptés pour tenir compte des changements au cours du temps du flot optique de chaque pixel de l'objet. Le modèle de filtre de Kalman est utilisé alors pour chaque pixel avec des paramètres adaptés (Fig. 3.9).

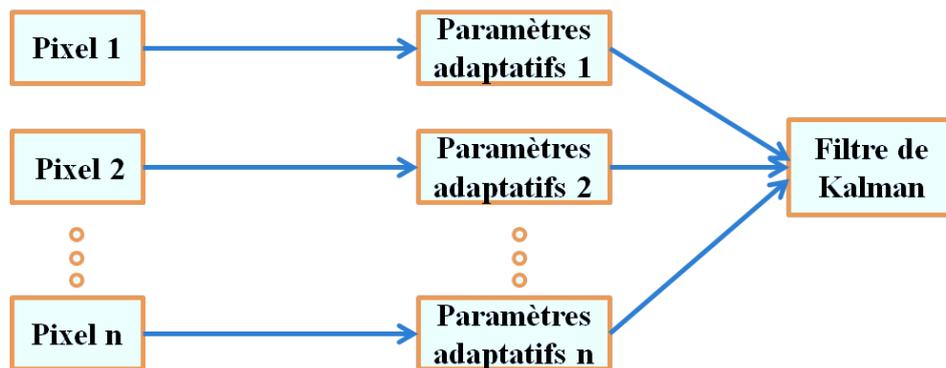


FIGURE 3.9 – Filtre de Kalman adaptatif.

Les paramètres concernés par l'adaptation est la matrice de covariance du bruit de process Q_t et la matrice de covariance du bruit de mesure R_t . La stratégie d'adaptation proposée est la suivante.

Pour la matrice de covariance du bruit de process Q_t , le calcul se fait par rapport au temps dt entre deux images successives et à l'accélération γ du centre de gravité des points de Harris à l'instant t :

$$Q_t = \begin{bmatrix} \frac{dt^2}{2} & 0 & 0 & 0 \\ 0 & \frac{dt^2}{2} & 0 & 0 \\ 0 & 0 & dt & 0 \\ 0 & 0 & 0 & dt \end{bmatrix} * \gamma \quad (3.24)$$

Pour ajuster la matrice de covariance du bruit de mesure R_t , nous prenons en compte les variances de mesure $\sigma_{p_x}^2, \sigma_{p_y}^2$ de la position du pixel de l'objet considéré et les variances du flot optique des points de Harris $\sigma_{f_x}^2, \sigma_{f_y}^2$:

$$R_t = \begin{bmatrix} \sigma_{p_x}^2 & 0 & 0 & 0 \\ 0 & \sigma_{p_y}^2 & 0 & 0 \\ 0 & 0 & \sigma_{f_x}^2 & 0 \\ 0 & 0 & 0 & \sigma_{f_y}^2 \end{bmatrix} \quad (3.25)$$

$$\sigma_{p_x} = \frac{D * f_x}{D_{max}} \quad \sigma_{p_y} = \frac{D * f_y}{D_{max}} \quad (3.26)$$

où D est la distance entre la position du pixel considéré et le point de Harris le plus proche de ce pixel. D_{max} est la distance entre les deux points les plus éloignés de l'objet. f_x et f_y sont les coordonnées du flot optique du pixel considéré.

3.4.2.4 Modèle global de suivi

La figure 3.10 illustre le modèle global de suivi d'objets. Il commence par détecter les objets en mouvement en utilisant la technique de différence inter-image et le processus de cadrage des objets. A l'issue de la détection des objets, la procédure de séparation est appliquée pour extraire les cibles. Ensuite, pour chaque cible extraite, le flot optique des pixels est estimé en utilisant le processus de propagation gaussienne du flot optique des points de Harris. Le processus itératif de filtrage de Kalman est ensuite appliqué. Le flot optique en sortie du filtre de Kalman peut être entâché de bruit. Dans le but d'obtenir un flot optique plus robuste, nous proposons une procédure de rectification. Celle-ci est conduite de la manière suivante.

Le flot optique en sortie du filtre de Kalman est ensuite soumis à un test par rapport au domaine de recherche défini par une distribution gaussienne du flot optique (dont les paramètres sont définis par rapport au flot optique des points de Harris) (eq.3.11). Si ce test est positif, le flot optique est soumis à un second test de similarité (eq.3.14). Si ce second test est positif, le flot optique est retenu. Sinon, le filtre de Kalman est à nouveau sollicité pour effectuer un nouveau filtrage. Si le test sur le domaine de recherche est négatif, nous appliquons un algorithme d'optimisation basé sur la comparaison des différences d'intensité entre le pixel considéré et le point de Harris de l'objet entre deux instants successifs. A l'issue

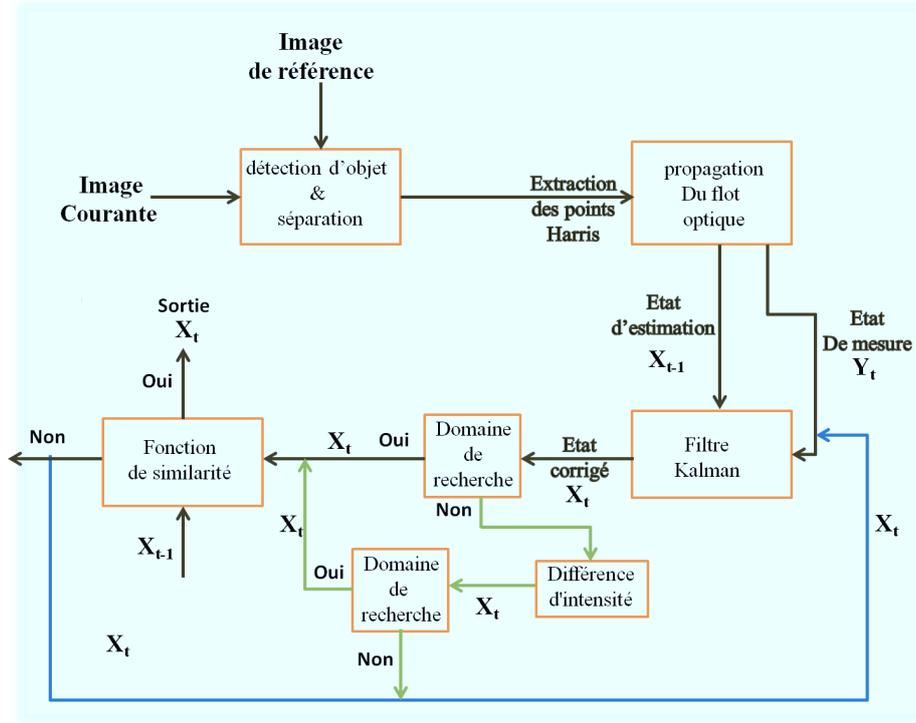


FIGURE 3.10 – Synoptique du suivi des objets.

de cet algorithme d'optimisation, le flot optique calculé est à nouveau soumis aux tests sur le domaine de recherche et la similarité.

Quatre itérations suffisent à l'algorithme de filtrage pour essayer de proposer une correction acceptable pour chaque pixel de l'objet détecté.

3.4.2.5 Algorithme d'optimisation

L'objectif de cet algorithme d'optimisation est d'améliorer la précision et la robustesse de la sortie du filtre de Kalman lors du processus de rectification du flot optique. L'idée principale de cet algorithme est de rectifier le flot optique en comparant les différences d'intensité du pixel considéré et des points de Harris entre deux instants successifs. Nous considérons que le flot optique optimal devrait permettre d'obtenir des différences d'intensité quasi-identiques. Etant donné le flot optique correspondant à l'état X_t du pixel en sortie du filtre de Kalman, le but est de calculer le vecteur de déplacement $d = [d_x, d_y]^T$ permettant de minimiser la fonction d'appariement ε défini par :

$$\varepsilon(fl) = \sum_{i=0}^{N_h} \left((I_{t-1}(x, y) - I_{t-1}^i) - (I_t(x + fl_x + d_x, y + fl_y + d_y) - I_t^i) \right)^2 \quad (3.27)$$

où N_h est le nombre de points de Harris extraits de l'objet considéré. $I_{t-1}(x, y)$, est l'intensité du pixel (x, y) à l'instant $t-1$. $I_t(x + fl_x + d_x, y + fl_y + d_y)$ est l'intensité du pixel estimé à l'instant t , à partir du pixel (x, y) à l'instant $t-1$, avec le flot optique $[fl_x + d_x, fl_y + d_y]$. I_{t-1}^i est l'intensité du point de Harris numéro i à l'instant $t-1$. I_t^i est l'intensité du pixel estimé à l'instant t , à partir du point de Harris numero i à l'instant $t-1$ avec le flot optique de ce point de Harris. $fl = [fl_x \ fl_y]^T$ est le vecteur flot optique estimé par le filtre de Kalman.

À l'optimum, la dérivée première de la fonction d'erreur ε par rapport à fl sera égal à zéro :

$$\begin{bmatrix} \frac{\partial \varepsilon}{\partial x} & \frac{\partial \varepsilon}{\partial y} \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix} \quad (3.28)$$

$$\begin{bmatrix} \frac{\partial \varepsilon}{\partial x} & \frac{\partial \varepsilon}{\partial y} \end{bmatrix} = -2 * \sum_{i=0}^{N_h} \left((I_{t-1}(x, y) - I_{t-1}^i) - (I_t(x + fl_x + d_x, y + fl_y + d_y) - I_t^i) \right) \cdot \begin{bmatrix} \frac{\partial I_t}{\partial x} & \frac{\partial I_t}{\partial y} \end{bmatrix} \quad (3.29)$$

Après le développement de Taylor à l'ordre 1 de $I_t(x + fl_x + d_x, y + fl_y + d_y)$, nous obtenons :

$$\frac{1}{2} * \begin{bmatrix} \frac{\partial \varepsilon}{\partial x} & \frac{\partial \varepsilon}{\partial y} \end{bmatrix} = - \sum_{i=0}^{N_h} \left((I_{t-1}(x, y) - I_{t-1}^i) - (I_t(x + fl_x, y + fl_y) - \begin{bmatrix} \frac{\partial I_t}{\partial x} & \frac{\partial I_t}{\partial y} \end{bmatrix} * \begin{bmatrix} d_x \\ d_y \end{bmatrix} - I_t^i) \right) \cdot \begin{bmatrix} \frac{\partial I_t}{\partial x} & \frac{\partial I_t}{\partial y} \end{bmatrix} \quad (3.30)$$

La forme simplifiée de l'équation (eq.3.30) est :

$$\frac{1}{2} * \begin{bmatrix} \frac{\partial \varepsilon}{\partial x} & \frac{\partial \varepsilon}{\partial y} \end{bmatrix}^T = G * \begin{bmatrix} d_x & d_y \end{bmatrix}^T - B \quad (3.31)$$

$$G = N_h * \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (3.32)$$

$$B = \sum_{i=0}^{N_h} \left(\begin{bmatrix} D_i * I_x \\ D_i * I_y \end{bmatrix} \right) \quad (3.33)$$

$$D_i = (I_{t-1}(x, y) - I_{t-1}^i) - (I_t(x + fl_x, y + fl_y) - I_t^i) \quad (3.34)$$

Finalement, le vecteur de déplacement d est donné par :

$$\begin{bmatrix} d_x & d_y \end{bmatrix}^T = G^{-1} * BD \quad (3.35)$$

3.5 Résultats expérimentaux

Dans ce chapitre nous présentons trois expériences de suivi d'objets dans des séquences d'images. Ces expérimentations permettent de mettre en évidence l'efficacité de la méthode de suivi proposée. Dans la première expérimentation, la méthode développée est testée sur des séquences d'images réelles. Dans la deuxième expérimentation, nous présentons la distribution gaussienne des flots optiques des pixels de l'objet suivi par rapport aux deux approches de répartition du flot optique proposées (voir § 3.3.2). Nous analysons par la suite ces courbes gaussiennes pour évaluer notre méthode. La troisième expérimentation concerne le suivi d'un objet virtuel. Le but de cette expérimentation est de montrer la capacité de suivi de la méthode proposée en comparant les résultats du suivi grâce à une base de données connue dans la littérature.

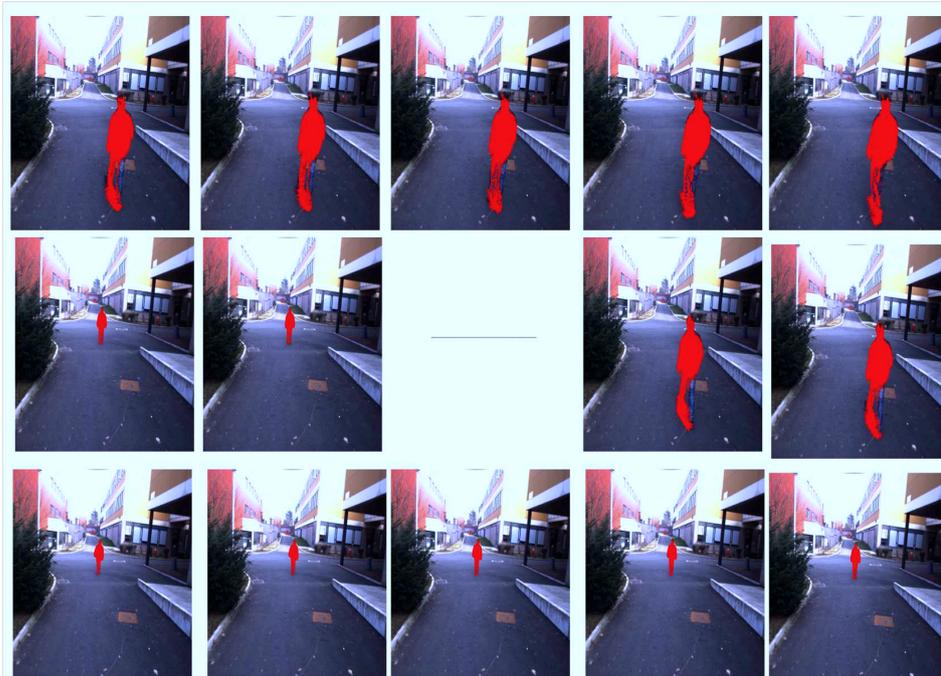
3.5.1 Expérimentation 1 : Suivi de piétons

Dans un premier temps, nous nous focalisons sur le suivi d'un seul objet afin d'évaluer les performances de la méthode proposée en termes de suivi local. Les figure 3.11(a) et figure 3.11(b) montrent respectivement les résultats de suivi avec l'algorithme de Lucas-Kanade et la méthode proposée. Seuls les pixels avec un suivi précis sont labellisés en rouge. Nous constatons que notre méthode donne de très bons résultats, y compris en présence de changement de luminance et de régions homogènes dans l'objet suivi. De plus, comparée à la méthode de Lucas-Kanade, la technique proposée permet de suivre un nombre élevé de pixels. En effet, notre méthode est capable de suivre de manière précise au-delà de 60% des pixels de l'objet, alors que l'algorithme de Lucas-Kanade perd la majorité des pixels durant la séquence (ce qui nécessite pour la méthode de Lucas-Kanade de procéder régulièrement à une phase de détection alors que pour notre méthode, la détection est faite qu'une seule fois au début de la séquence). Ce résultat est confirmé par la figure 3.12, qui présente l'évolution du nombre de pixels suivis pour notre méthode et celle de Lucas-Kanade. Dans cette expérience, le piéton est suivi même lorsqu'il se trouve éloigné de la caméra. Ceci explique la diminution de nombre de pixels appartenant à l'objet suivi au cours de temps. En termes de temps de calcul, le processus de suivi des objets tourne en moyenne à une cadence de 4 images/seconde.

La figure 3.13 illustre un suivi multi-objets par la méthode proposée. Nous obtenons de très bons résultats en termes de séparation d'objets et de suivi local. Par ailleurs, la méthode est capable de résoudre des situations difficiles avec des mouvements compliqués, comme dans le cas où un des piétons passe en dessous de la barrière d'un passage à niveau. Les problèmes liés aux occultations sont également pris en compte. Pour une occultation partielle, les objets sont suivis par leurs pixels visibles. Pour une occultation totale, l'étape de détection est nécessaire pour continuer le suivi, en utilisant le fait que les objets entrent et sortent de la scène à surveiller par la zone d'entrée/sortie.



(a) Résultats de suivi par l'algorithme de Lucas-Kanade



(b) Résultats de suivi par la méthode proposée

FIGURE 3.11 – Résultats de suivi avec un objet en mouvement : (a) méthode de Lucas-Kanade (b) méthode proposée.

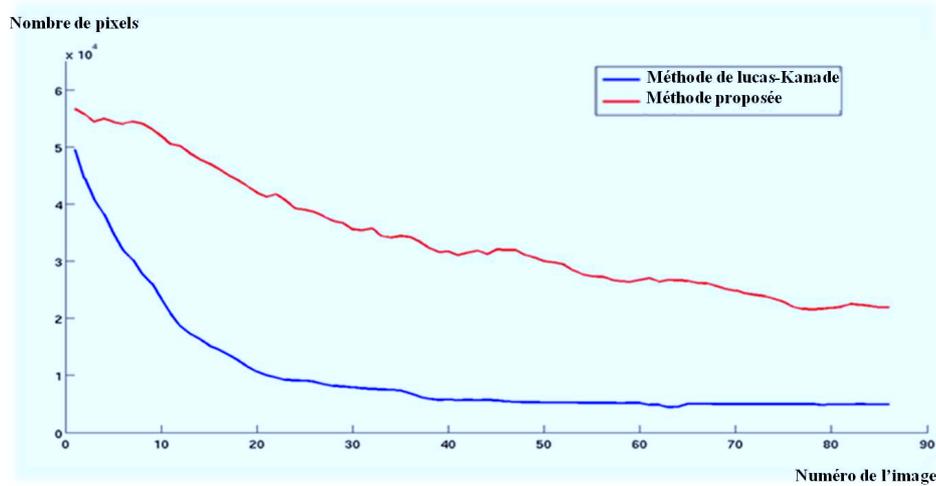


FIGURE 3.12 – Evolution du nombre de pixels suivis par notre méthode et celle de Lucas-Kanade.

3.5.1.1 Évaluation des résultats de suivi obtenus

Dans cette partie, nous effectuons quelques tests pour évaluer les résultats de suivi obtenus. Les figure 3.14(a) et figure 3.14(b) montrent la répartition de l'erreur d'intensité sur l'objet piéton de la séquence présentée sur la figure 3.11, en utilisant respectivement la méthode de Lucas-Kanade et la méthode proposée. L'erreur d'intensité exprime la différence d'intensité (en module) des pixels affectés par du mouvement entre deux images successives. Nous pouvons remarquer que la plupart des pixels de l'objet suivi par la méthode proposée font apparaître un faible niveau d'erreur d'intensité. D'autre part, la plupart des pixels suivis par la méthode de Lucas-Kanade présentent un niveau d'erreur d'intensité élevé.

Les figure 3.15 et figure 3.16 fournissent une représentation 2D de l'erreur d'intensité en utilisant respectivement la méthode proposée et celle de Lucas-Kanade. L'axe horizontal définit la distance entre le centre géométrique des points de Harris de l'objet et la position du pixel considéré. L'axe vertical représente l'erreur d'intensité du pixel considéré. Pour mieux lire les résultats de ce test, nous traçons la courbe médiane (valeurs médianes) de l'erreur d'intensité pour chaque méthode. Lorsque l'on compare les résultats des deux méthodes, on constate que notre méthode réduit de manière significative l'erreur, en comparaison avec la méthode Lucas-Kanade. En effet, sur la figure 3.16, où nous avons tracé les deux courbes médianes correspondant aux deux méthodes, nous pouvons constater que la courbe médiane issue de la méthode proposée est toujours en dessous de celle obtenue par la méthode de Lucas-Kanade, notamment pour les pixels qui sont loins du centre géométrique des points de Harris.

Sur la figure 3.17, sont illustrées les valeurs médianes des erreurs d'intensité (par intervalle de taille 25 sur l'axe des distances) obtenues par plusieurs méthodes de

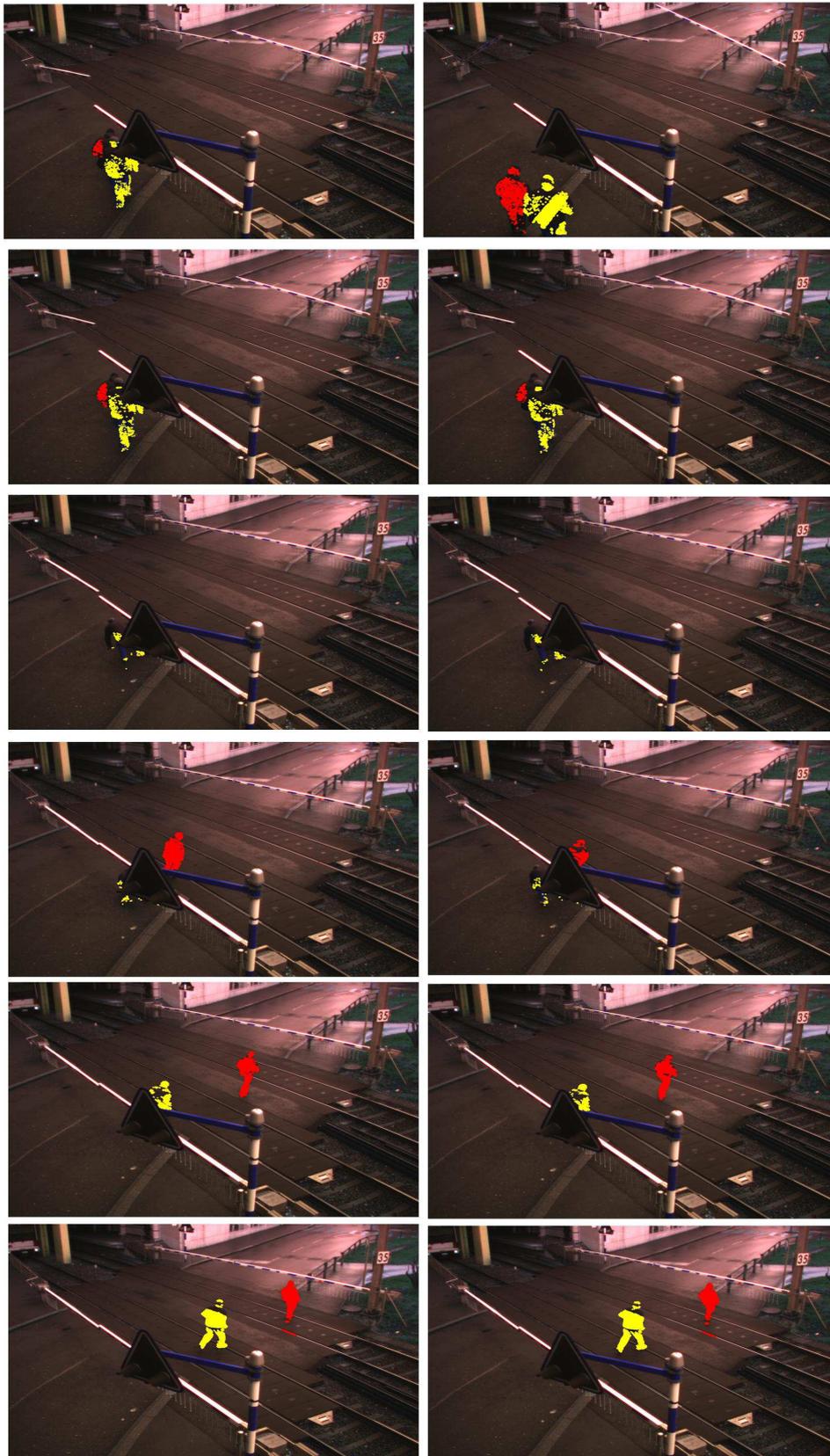
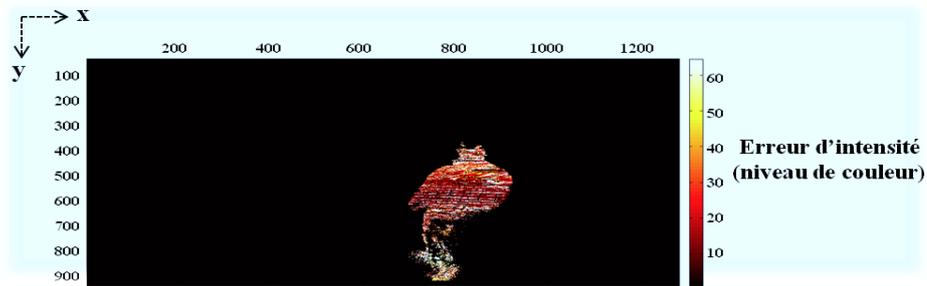
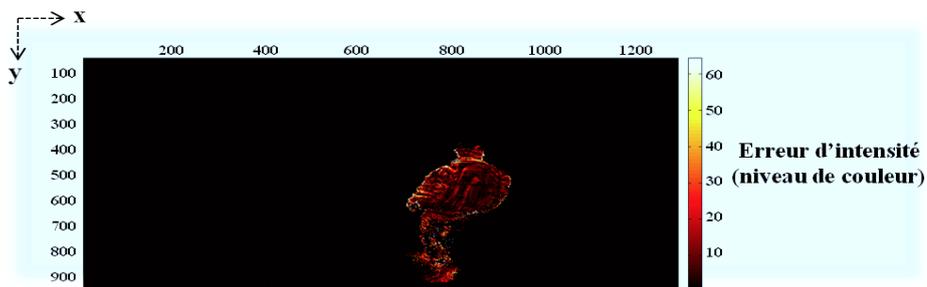


FIGURE 3.13 – Résultats de suivi multi-objets par la méthode proposée.



(a) niveau d'erreur de l'intensité (méthode Lucas-Kanade)



(b) niveau d'erreur de l'intensité (méthode proposée)

FIGURE 3.14 – répartition de l'erreur d'intensité.

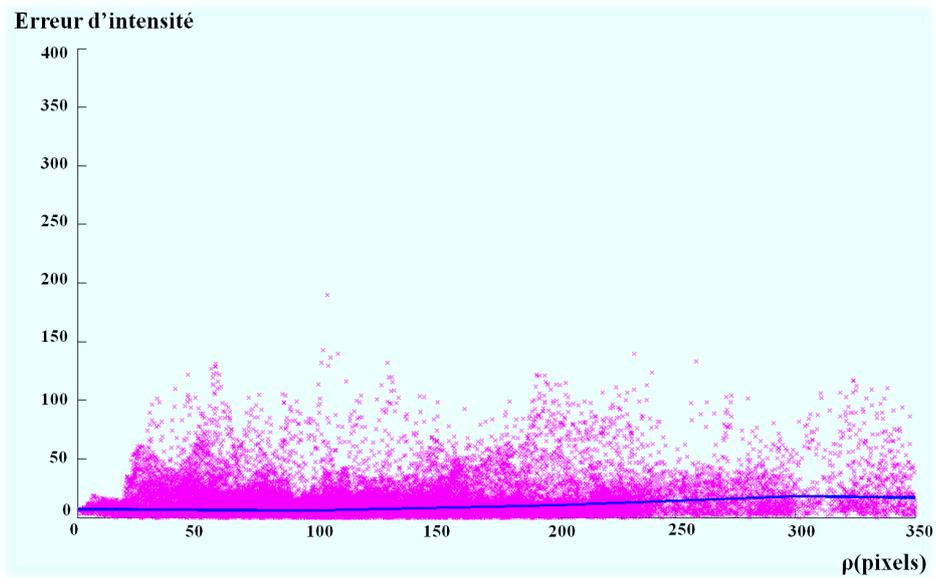


FIGURE 3.15 – Représentation de l'erreur d'intensité par la méthode proposée.

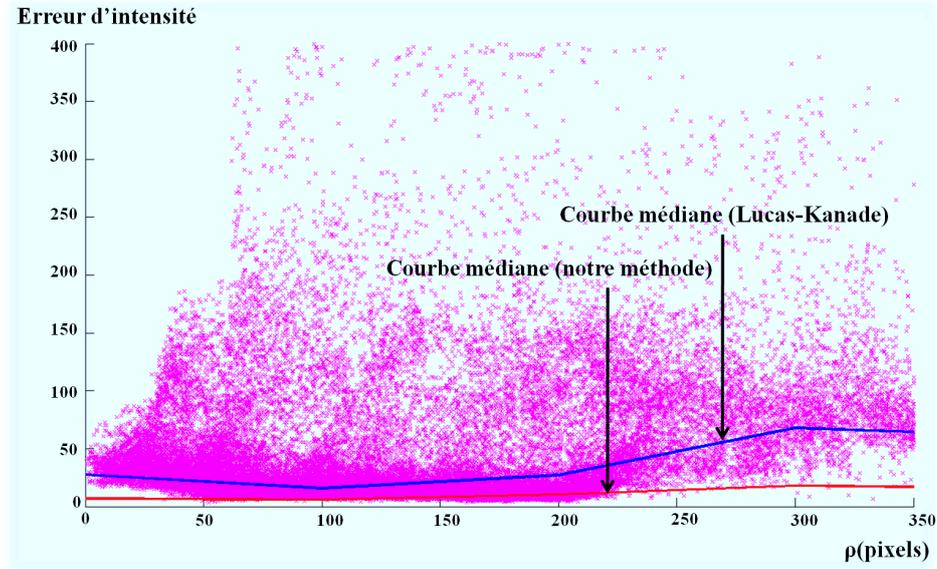


FIGURE 3.16 – Représentation de l'erreur d'intensité par la méthode de Lucas-Kanade.

calcul du flot optique :

1) Calcul du flot optique par propagation gaussienne du flot optique des points de Harris, en utilisant la première proposition de distribution gaussienne (voir § 3.3.2), Suivie des tests de domaine de recherche et de similarité (Méthode 1). Nous rappelons que cette distribution a comme moyenne la valeur moyenne du flot optique des points de Harris.

2) Calcul du flot optique par propagation gaussienne du flot optique de point de Harris, en utilisant la deuxième proposition de distribution gaussienne, (voir § 3.3.2), Suivie des tests de domaine de recherche et de similarité (Méthode 2). Nous rappelons que cette distribution est appliquée à chaque ensemble de pixels défini par un point de Harris. Un ensemble de pixels lié a un point de Harris est composé des pixels de l'objet les plus proche à ce point.

3) Calcul du flot optique avec la méthode de Lucas-Kanade (Méthode 3).

4) Calcul du flot optique avec la méthode proposée (Méthode 4).

Dans cette figure, nous remarquons que la courbe médiane obtenue par la méthode 2 (courbe rouge) est toujours en dessous de celle obtenue par la méthode 1 (courbe bleue). Ceci permet de dire que la deuxième distribution gaussienne est plus intéressante que la première. C'est donc la deuxième distribution gaussienne qui est utilisée dans notre méthode. Nous constatons aussi que la courbe médiane obtenue par la méthode proposée (courbe magenta) est toujours en dessous de toutes les autres courbes. Ceci prouve que notre méthode fournit un flot optique plus robustes même pour les pixels éloignés du centre géométrique des points de Harris. De plus, on remarque que la valeur médiane de l'erreur d'intensité augmente de ma-

nière significative lorsque les pixels sont de plus en plus loins du centre géométrique des points de Harris pour toutes les méthodes, Sauf pour notre méthode pour laquelle l'erreur d'intensité demeure quasi-constante. Les méthodes de Lucas-Kanade et les méthodes 1 et 2 sont intéressantes seulement pour le calcul du flot optique des pixels les plus proches au centre géométrique des points de Harris. Ceci permet de conclure que ces trois méthodes dépendent fortement du nombre de points de Harris extraits et de leur distribution sur l'objet suivi. Notre méthode dépend peu de cette contrainte, puisque la valeur médiane de l'erreur d'intensité reste quasi-constante pour tous les pixels qu'ils soient loins ou proches du centre géométrique des points de Harris. Cette performance est obtenue grâce au modèle de filtre de Kalman et à la procédure de rectification du flot optique. On peut constater sur cette même figure une baisse de l'erreur d'intensité pour les pixels suivis les plus éloignés du centre géométrique des points de Harris. Ceci est dû au fait que le nombre de ces pixels suivis est moins importante. En effet, la majorité des pixels les plus éloignées du centre géométrique des points de Harris n'ont pas suivi ce flot optique est considérée comme imprécise grâce aux tests de domaine de recherche et de similarité.

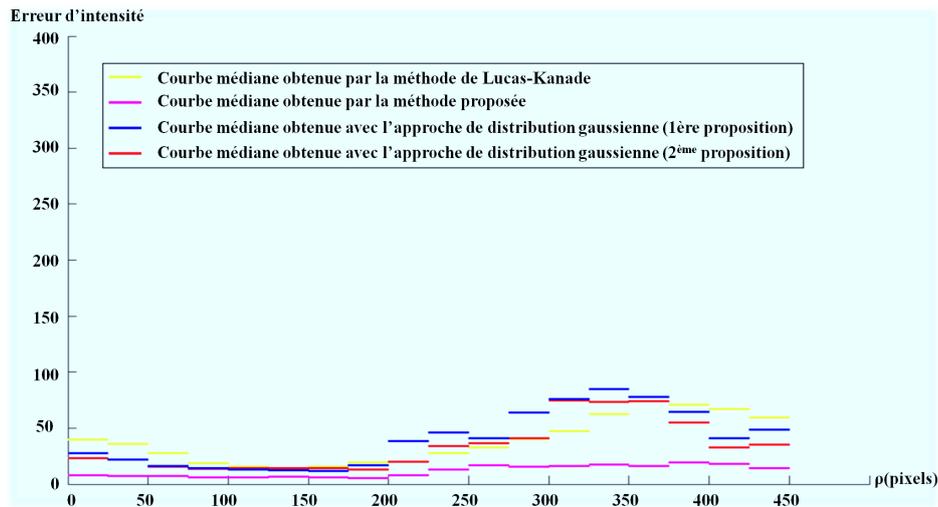


FIGURE 3.17 – Courbes médianes obtenues par les différentes méthodes.

3.5.2 Expérimentation 2 : Evaluation avec des courbes gaussiennes

L'objectif de cette expérimentation est de montrer comment est distribué le flot optique des pixels appartenant à un objet suivi par l'application de la méthode proposée. Le graphe de cette distribution est déterminé en fonction du flot optique des points de Harris et de sa moyenne, tout en considérant que ces points présentent le flot optique le plus stable dans l'objet. On cherche à montrer à travers cette expérimentation que plus la distribution du flot optique est large, plus on peut

trouver des solutions possibles pour suivre les pixels de l'objet. Nous comparons ensuite les résultats obtenus avec les deux distributions proposées (§ 3.3.2). Cette étude est effectuée sur la composante horizontale du flot optique ($ft(x)$) selon les fonctions de distribution suivantes :

$$dist1 = e^{-\frac{1}{2}\left(\frac{ft(x)-m_x}{\sigma_x}\right)^2} \tag{3.36}$$

$$dist2 = e^{-\frac{1}{2}\left(\frac{ft(x)-F_x^i}{\sigma_x}\right)^2} \tag{3.37}$$

avec m_x le moyenne du flot optique des points de Harris sur l'axe x . F_x^i est le flot optique du point de Harris numéro i le plus proche du pixel (x, y) de l'objet. σ_x est la somme de l'écart type du flot optique des points de Harris σ_{hx} sur l'axe x et l'écart type d_{σ_x} d'un bruit inconnu (eq.3.12).

3.5.2.1 1ère distribution proposée (en fonction du flot optique moyen des points de Harris)

La figure 3.18 présente la distribution initiale du flot optique des pixels de l'objet. Les valeurs du flot optique sont initialisés autour du flot optique moyen des points de Harris. On peut constater que 89% de ces valeurs ont une probabilité de distribution supérieure à 0.8. En effet grâce à la propagation gaussienne, la plupart des valeurs du flot optique sont concentrées autour de la moyenne calculée à partir du flot optique des points de Harris.

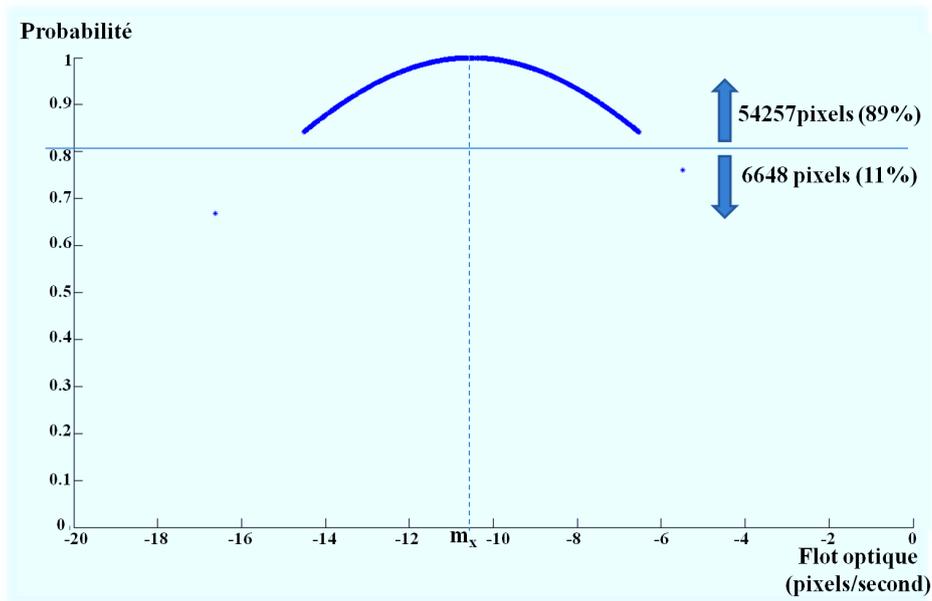


FIGURE 3.18 – Distribution initiale du flot optique.

La figure 3.19 présente les résultats de la distribution finale du flot optique des pixels de l'objet après application de la méthode proposée à base de la distribution considérée. Nous pouvons noter que 41% des valeurs du flot optique estimé ont une probabilité de distribution supérieure à 0.8. Nous pouvons dire alors que 41% du flot optique estimé, représentant des valeurs initiales des pixels les plus proches du centre géométrique des points de Harris, restent assez proche des valeurs initiales. D'autre part, 59% des valeurs du flot optique estimé, représentant le flot optique des pixels les plus loins du centre géométrique des points de Harris, ont été rectifiés de manière significative en utilisant le filtre de Kalman itératif suivi de la procédure de rectification.

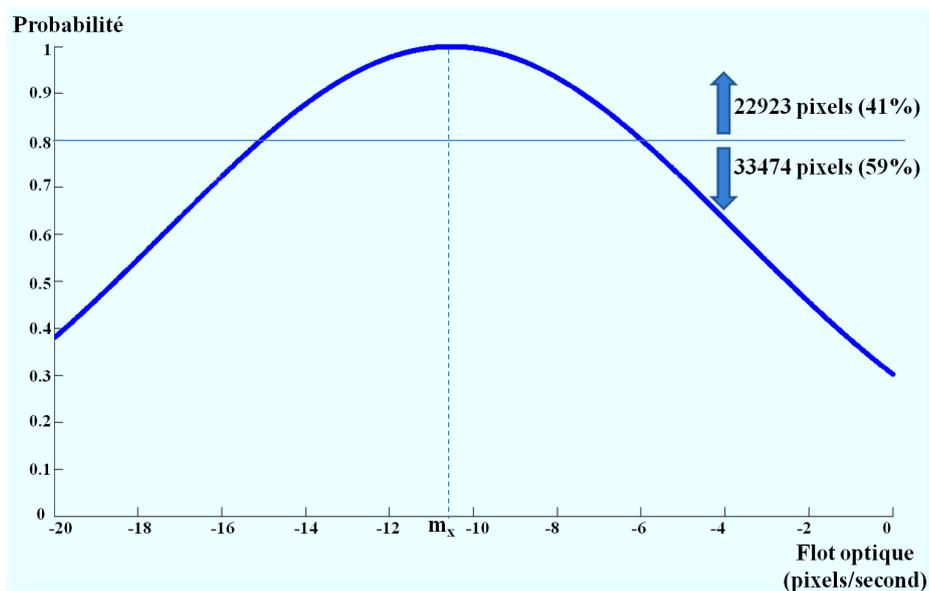


FIGURE 3.19 – Distribution du flot optique après application de la méthode proposée (avec la 1ère distribution gaussienne).

3.5.2.2 2ème distribution proposée (en fonction du flot optique de point de Harris le plus proche)

Chaque point de Harris détecté dans l'image possède un ensemble de pixels considérés comme les points les plus proches de ce point. Dans notre exemple, 5 points de Harris ont été détectés. A chaque point correspond alors un ensemble composé des valeurs du flot optique des pixels les plus proches de ce point. La figure 3.20 illustre les résultats des distributions initiale et finale de chaque ensemble de flots optiques. Initialement (avant filtrage et rectification), 15%, 8%, 16%, 1%, 7% sont les pourcentages de répartition du flot optique pour les groupes de pixels ayant une probabilité de distribution inférieure à 0.8. 62%, 14%, 49%, 0%, 29% sont les pourcentages de répartition du flot optique pour les groupes de pixels ayant une

probabilité de distribution inférieure à 0.8, après filtrage et rectification. On peut constater alors que le flot optique des pixels de chaque groupe (représentant les pixels les plus éloignés du point de Harris correspondant) a été rectifié de manière significative par rapport à sa valeur initiale (15%, 8%, 16%, 1%, 7%).

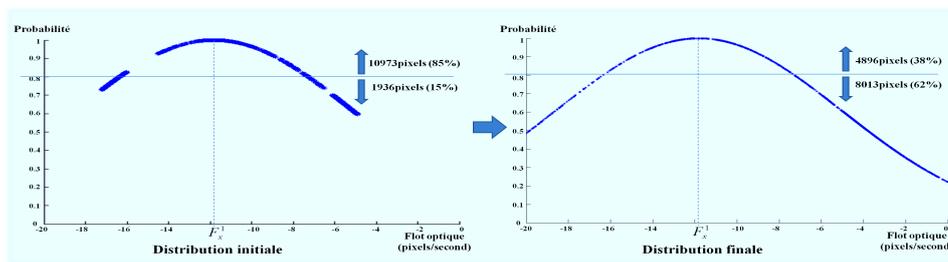
Si nous considérons tous les pixels de manière globale, nous avons 35% des valeurs du flot optique ayant une probabilité de distribution inférieure à 0.8. Les 35% sont obtenus en faisant la somme des pixels pour lesquels le flot optique a une probabilité de distribution inférieure à 0.8, dans chaque groupe, divisé par le nombre total de pixels. Pour comparer les deux distributions gaussiennes, une analyse globale (avec tous les pixels) de la répartition du flot optique ayant une probabilité de distribution inférieure à 0.8 permet de constater, qu'avec la deuxième distribution, 35% (contre 59% pour la première distribution) du flot optique a subit une correction suite au filtrage et à la rectification. Ceci montre qu'avec la deuxième distribution gaussienne, notre méthode n'a en besoin de corriger que 35% du flot optique initial (contre 59% avec la première distribution gaussienne). Cela montre que l'utilisation de la deuxième distribution gaussienne permet de fournir au préalable un flot optique correct pour un certain nombre de pixels.

En utilisant la moyenne et la variance du flot optique des points de Harris, la figure 3.21 donne la distribution du flot optique estimé avec notre méthode à base de la deuxième distribution gaussienne. Sur cette figure, nous constatons que 83% (contre 59% avec la première distribution (Fig. 3.19)) du flot optique ont une probabilité de distribution inférieure à 0.8. On peut conclure alors, qu'avec la deuxième distribution gaussienne, on obtient une meilleure distribution du flot optique sur l'objet.

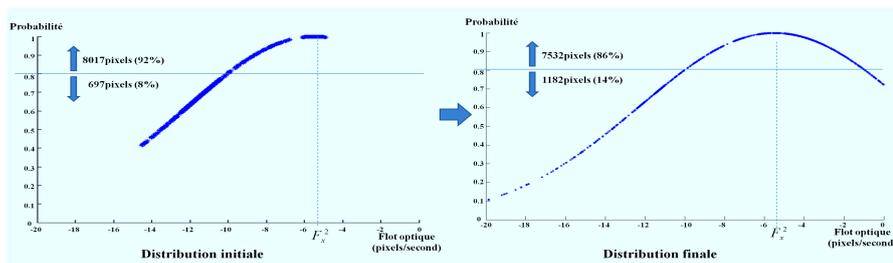
3.5.3 Expérimentation 3 : Suivi d'un objet dans une séquence d'images simulées

L'objectif de cette expérimentation est d'évaluer notre méthode de suivi en ayant des données vérité terrain. Ces données vérité terrain sont générées avec une séquence d'images simulées. Une grille de forme rectangulaire est définie par sa longueur, sa largeur, sa couleur blanche et sa position dans l'espace. Nous avons ensuite utilisé un modèle de projection avec des paramètres réels (focale, resolution, ...) d'une caméra. Enfin, nous avons projeté les noeuds de la grille sur un plan image, en imposant à cette grille un mouvement hélicoïdal dans l'espace. La figure 3.22 montre les résultats de projection de la grille sur le plan image durant son déplacement. On peut voir dans l'image que la taille de la grille diminue au cours du temps. Nous avons ajouté un bruit gaussien à chaque image de la séquence. Cette séquence d'image présente la difficulté de distinguer les pixels les uns des autres et pour les suivre correctement (même intensité, pas de changement de gradient à l'intérieur de la grille. . .). L'objectif principal de cette expérimentation est de montrer le degré d'exactitude de la méthode proposée en évaluant l'erreur de position des pixels affectés par du mouvement.

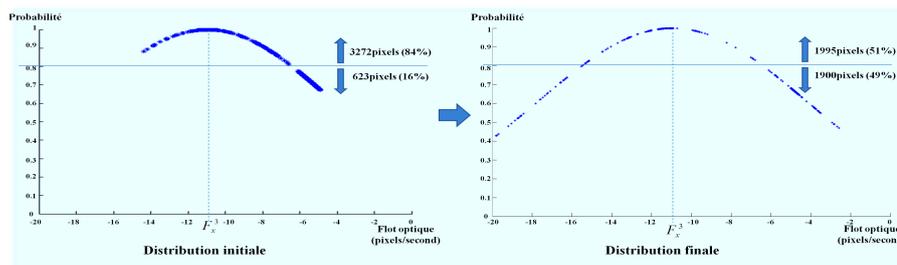
La figure 3.23 présente les pixels suivis (pixels de couleur jaune) en utilisant la



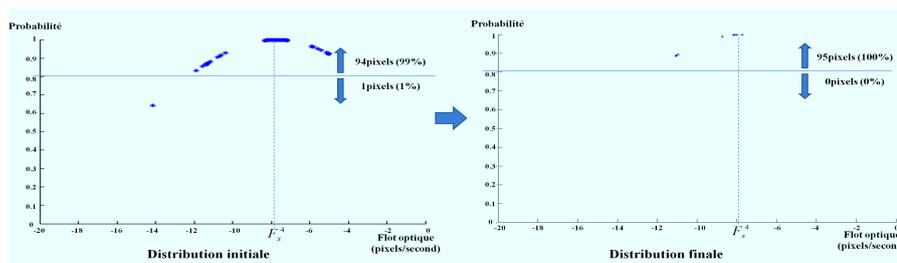
(a) Distribution du flot optique à partir du 1er point de Harris détecté



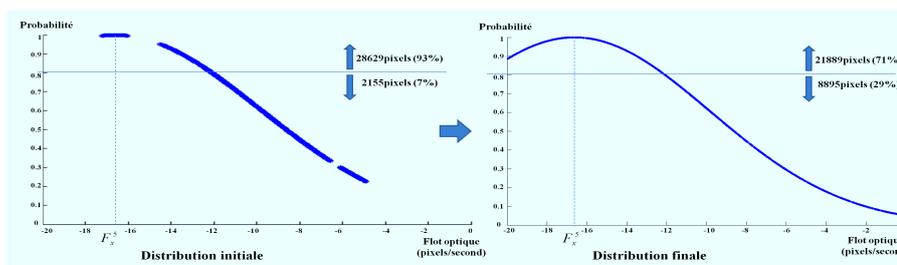
(b) Distribution du flot optique à partir du 2ème point de Harris détecté



(c) Distribution du flot optique à partir du 3ème point de Harris détecté



(d) Distribution du flot optique à partir du 4ème point de Harris détecté



(e) Distribution du flot optique à partir du 5ème point de Harris détecté

FIGURE 3.20 – Distribution du flot optique après application de la méthode proposée (avec la deuxième distribution gaussienne).

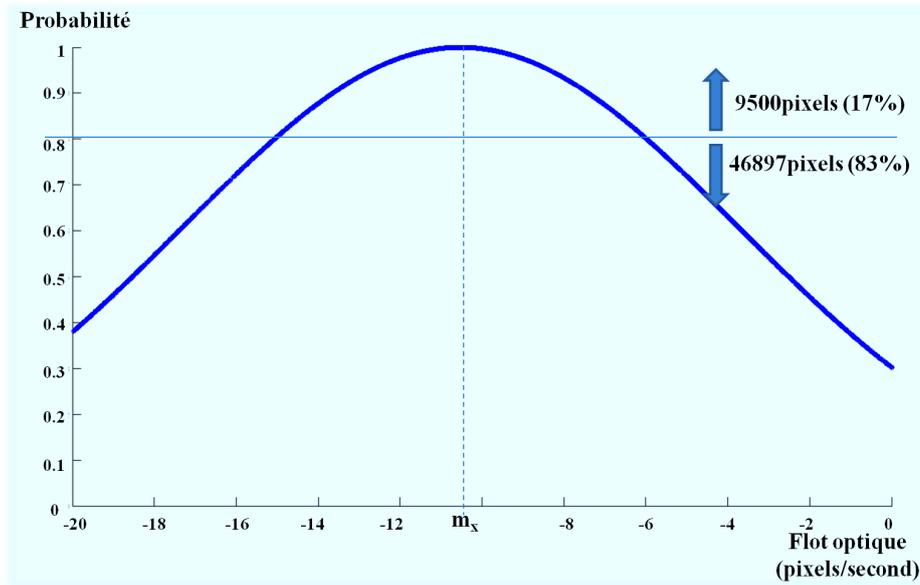


FIGURE 3.21 – Distribution du flot optique (avec la 2ème distribution gaussienne), en utilisant la moyenne et la variance du flot optique des points de Harris.

méthode proposée. Le résultat du suivi des pixels est comparé aux données fournies par la séquence d'images simulées.

Cette comparaison permet de tracer la courbe de l'erreur moyenne de position des pixels lors du déplacement de la grille (Fig. 3.24). En moyennant cette erreur sur chaque paquet de 25 images, on peut constater sur la figure figure 3.24 que l'erreur moyenne de position des pixels suivis est de l'ordre de 3.5 pixels. Cette erreur de position est considérée comme un résultat acceptable si on compare la méthode proposée avec les autres algorithmes de suivi existants [Fridberger 2004][Fuse 2000][Barron 1994] qui fait état d'une erreur de position moyenne de l'ordre de 5 pixels.

La figure 3.25 présente la distribution de l'erreur moyenne de position des pixels pour une des images de la séquence vidéo, en fonction de la distance au centre des points de Harris. Nous remarquons que la plupart des pixels suivis ont une erreur qui varie entre 2 et 3 pixels. L'erreur maximale de position des pixels est de l'ordre de 6 pixels et l'erreur minimale est de l'ordre de 0.6 pixel. La figure 3.25 montre également que l'erreur de position augmente en s'éloignant de plus en plus de centre géométrique des points de Harris. Cette erreur baisse pour les pixels suivis les plus éloignés du centre géométrique des points de Harris du fait que le nombre de ces pixels suivis est moins important.

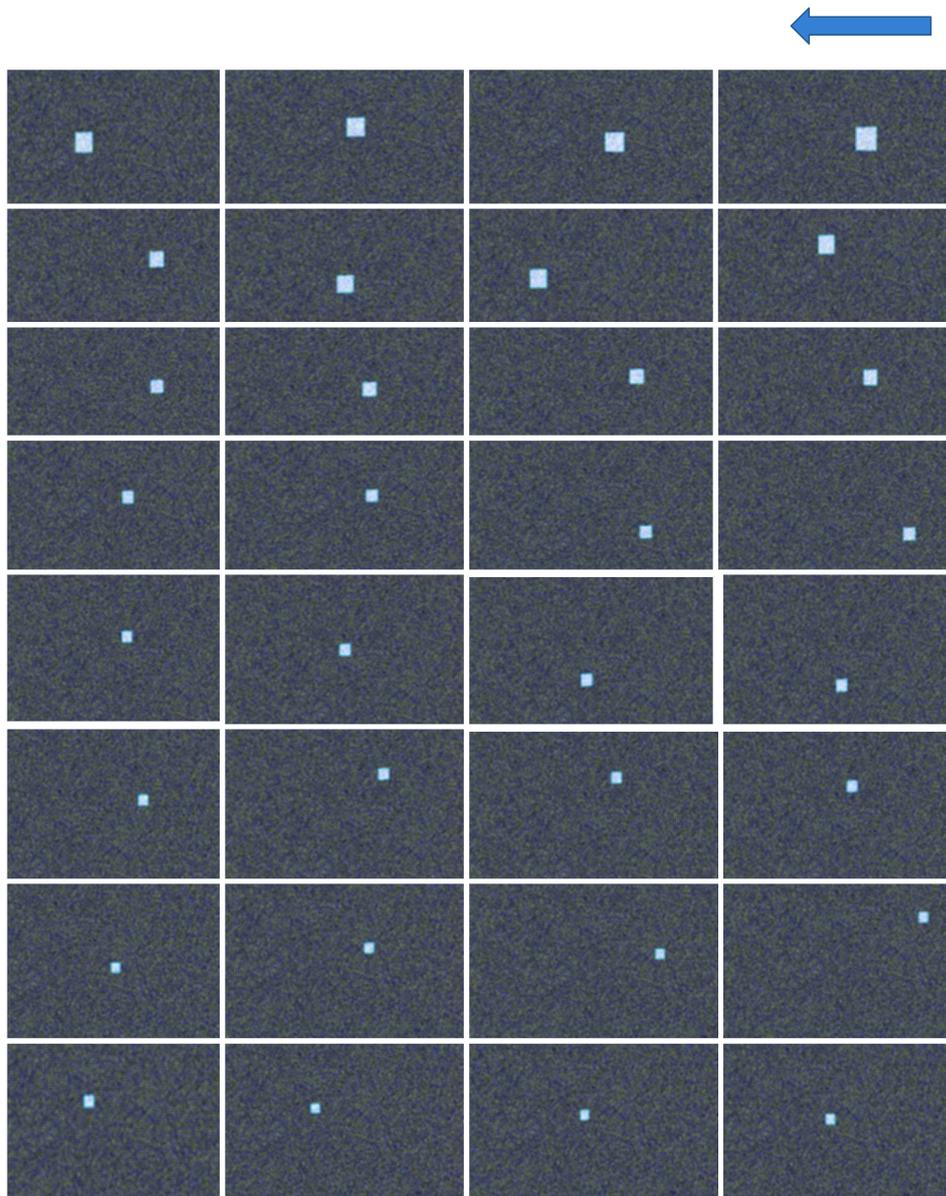


FIGURE 3.22 – Séquence d'images simulées.

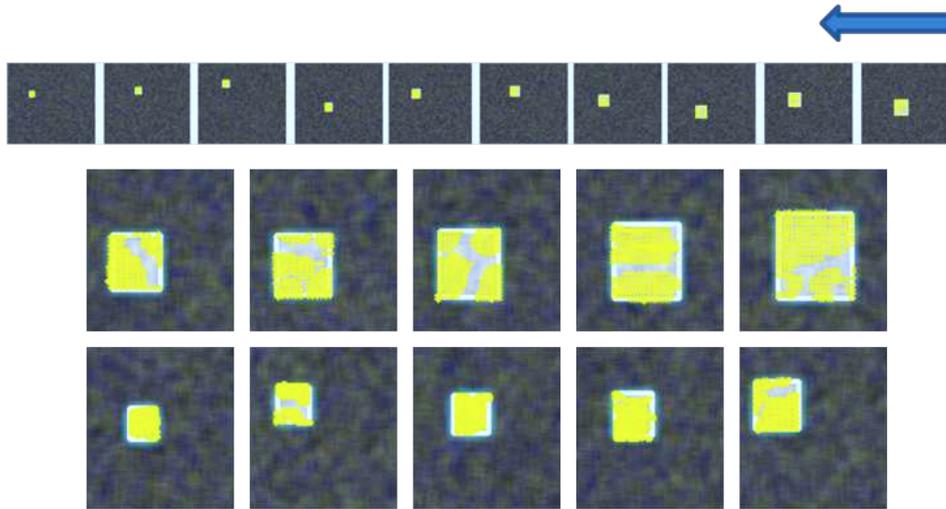


FIGURE 3.23 – Résultats du suivi de la grille.

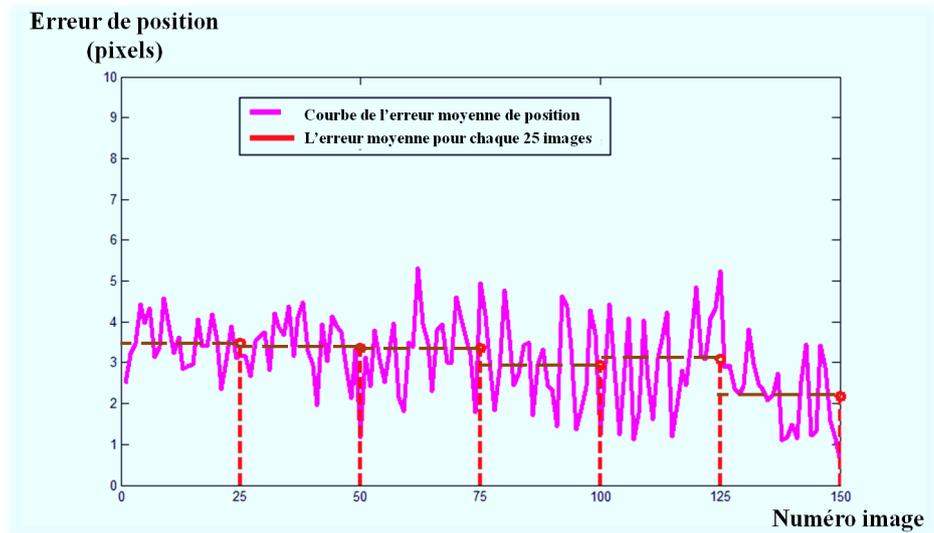


FIGURE 3.24 – Erreur moyenne de position des pixels suivis.

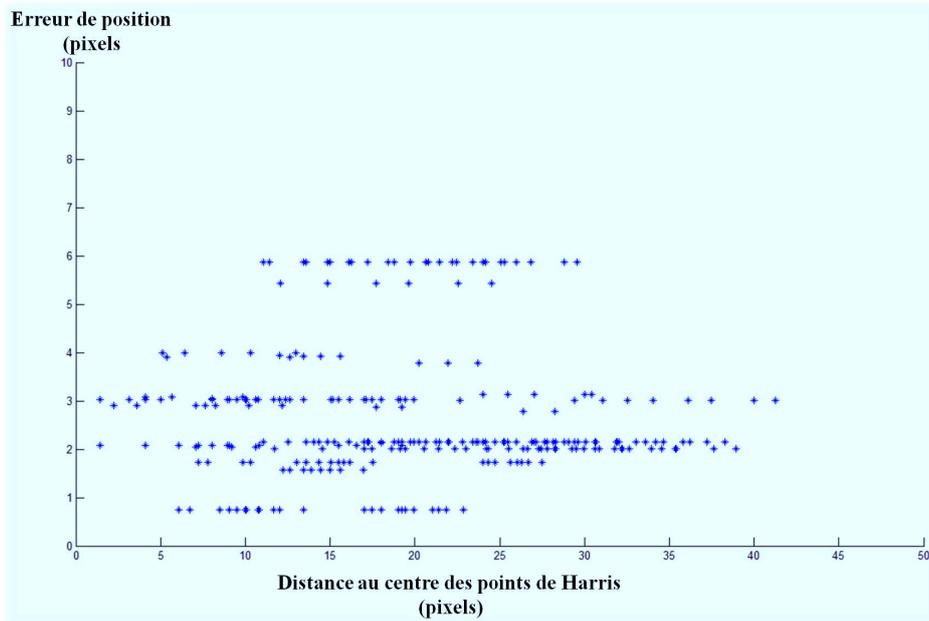


FIGURE 3.25 – distribution de l’erreur de position des pixels.

3.6 Conclusion

Dans ce chapitre, nous avons présenté dans premier temps une méthode capable de séparer les objets du fond. elle est basée sur une procédure à deux étapes : détection puis séparation. Nous avons ensuite attaqué le problème du suivi spatio-temporel des objets en utilisant une méthode de suivi local. Cette méthode est basée sur une technique de propagation du flot optique calculé sur les points de Harris, préalablement extraits des objets de la scène. Pour cela, nous avons choisi une distribution gaussienne dont les paramètres sont calculés à partir du flot optique obtenu sur les points de Harris. Le flot optique d’un point est tout d’abord initialisé par propagation du flot optique des points de Harris. Un filtre de Kalman accompagné d’un algorithme d’optimisation, est ensuite proposé pour faire face au bruit et fournir un flot optique corrigé. La méthode développée est testée sur des séquences d’images réelles et de simulation. Plusieurs expérimentations ont été effectuées pour évaluer l’efficacité de la méthode. Ces expérimentations ont montré la capacité de la méthode à suivre les pixels des objets dans des scènes relativement complexes. Dans le chapitre suivant, nous nous intéressons au problème de l’évaluation de situations potentiellement dangereuses aux passages à niveau. La résolution de ce problème exploite notre méthode de suivi d’objets afin de reconnaître la trajectoire 2D des objets en mouvement. L’analyse développée est basée sur une modélisation Markovienne et un traitement basé sur la théorie de l’évidence.

Evaluation du danger aux passages à niveau

Sommaire

4.1	Introduction : Problématique de la sécurité aux passages à niveau	85
4.2	Prédiction des trajectoires idéales d'objets mobiles	87
4.2.1	Introduction	87
4.2.2	Segmentation des objets à base du flot optique	87
4.2.3	Estimation des trajectoires idéales	88
4.3	Estimation du danger de situations potentiellement dangereuses	95
4.3.1	Introduction	95
4.3.2	Théorie de l'évidence : définition des paramètres	95
4.3.3	Modèle d'évaluation des trajectoires idéales	98
4.3.4	Définition des sources potentielles de danger	98
4.3.5	Condition d'indépendance des sources de danger	103
4.4	Résultats expérimentaux	103
4.4.1	Scénario 1 : Détection d'obstacles	104
4.4.2	Scénario 2 : Remontée de file	105
4.4.3	Scénario 3 : Passage en chicane, sans arrêt sur platelage	109
4.4.4	Scénario 4 : Passage en chicane, avec arrêt sur platelage	112
4.4.5	Scénario 5 : chute d'un piéton	115
4.5	Conclusion	121

4.1 Introduction : Problématique de la sécurité aux passages à niveau

Les passages à niveau (PN) représentent dans la plupart des cas le croisement d'une route et d'une voie de chemin de fer. Ce croisement d'infrastructures met en association ponctuelle deux autorités organisatrices, la DSCR (Direction de la Sécurité et Circulation Routières) et le RFN (Réseau Ferré National) relevant de modes de gestion très différents. La solution optimale pour réduire le risque de conflit pourrait être assurée par la séparation totale des trafics. Mais cette solution, qui

consiste donc à éliminer les PN préoccupants, pose des problèmes de coût important ainsi que des problèmes de report de trafic sur des voies routières non prévues à cet effet au départ. Une solution alternative à la suppression des PN préoccupants réside dans la possibilité d'une amélioration de leur sécurité.

Dans ce chapitre, nous proposons un système de perception de l'environnement PN afin de détecter les éléments accidentogènes et de partager cette information avec les entités empruntant le dit PN : usagers de la route ou bien un train en approche. Le système de perception, présenté dans le premier chapitre, et basé sur la vision artificielle, peut avoir deux fonctionnalités bien différentes, selon le mode considéré :

- En mode en ligne, il permettra de détecter en temps réel les situations accidentogènes pouvant se produire sur un PN.
- En mode hors ligne, il permettra de procéder à une évaluation a posteriori, du comportement des usagers aux passages à niveau.

Dans un souci constant d'amélioration de la sécurité aux passages à niveau afin de garantir la sécurité pour les usagers ainsi que le niveau de performance et de services de l'infrastructure, divers systèmes de détection d'obstacles aux passages à niveau ont été mis en place. Ces différents systèmes de détection ont pour objectif d'éviter les collisions ou au moins d'en limiter les conséquences. Ces techniques reposent sur des systèmes de balayage de la zone de passage à niveau par des faisceaux radar ou par des scanners optiques, sonores ou électromagnétiques ou sur l'analyse et le traitement de l'image vidéo.

Ce chapitre est consacré à la présentation de l'approche proposée pour l'évaluation du danger aux passages à niveau par analyse de séquences d'images vidéo. Cette analyse est basée sur les résultats du chapitre précédent concernant la détection et le suivi d'objets mobiles. De nombreux scénarii accidentogènes, pris en situation réelle, ont été considérés afin de valider l'approche proposée. Ces scénarii ont été tournés lors de campagnes d'acquisition sur le passage à niveau de Mouzon (Nord de la France), dans le cadre du projet PANsafer.

L'analyse des scénarii est basée sur des techniques d'apprentissage probabiliste telles que les modèles de Markov et les réseaux bayésiens. Dans un premier temps, nous estimons pour chaque objet détecté (voiture, camion, deux roues, piéton, ...) sa trajectoire idéale pour traverser normalement le passage à niveau. Cette trajectoire est estimée grâce à une chaîne de Markov que nous avons modélisée. Nous étudions ensuite le comportement de l'objet en mouvement de manière instantanée. Nous évaluons enfin le danger lié à l'objet durant son évolution dans l'environnement du PN. Pour cela, des paramètres caractérisant l'objet (position, taille, vitesse, nature, ...) sont fusionnés en se basant sur la théorie de l'évidence. Comme dit précédemment, le système global (acquisition et traitement) a été validé grâce à de nombreuses séquences d'images correspondant à des scénarios accidentogènes, acquises par le dispositif de prise de vues implanté sur un passage à niveau réel (site de Mouzon, France).

Le modèle de mélange gaussien [Bashir 2005], le modèle de Markov caché (MMC) [Bashir 2007] et certaines de ses extensions telles que le modèle de Markov caché hié-

rarchique (MMCH) [Nguyen 2005] et le modèle de Markov caché couplé (MMCC) [Natarajan 2007] sont généralement les méthodes les plus utilisées pour l'analyse des trajectoires d'objets associées à des scènes issues de la vidéosurveillance. Ces approches produisent d'excellents résultats dans de nombreux cas, mais ils ont besoin de bases importantes de données (ensemble de mesures) pour effectuer la phase d'apprentissage. Ainsi, le temps de calcul nécessaire constitue la limitation principale de ces méthodes. Dans ce travail de thèse, nous proposons un nouveau modèle de Markov caché applicable en temps réel. Le but de ce modèle est d'estimer les trajectoires idéales des objets en mouvement pendant une durée limitée. Une trajectoire idéale d'un objet à un instant donné est une trajectoire selon laquelle cet objet devrait traverser le PN afin d'éviter une situation potentiellement dangereuse. Pour chaque objet en mouvement détecté dans l'image courante, le flot optique est segmenté et chaque région peut être caractérisée par une trajectoire idéale. Cette approche de prédiction de trajectoires idéales d'objets tient compte en particulier de la géométrie de l'infrastructure (route/PN). Nous analysons ensuite le comportement de chaque objet en mouvement de manière instantanée. Pour cela, étant donnée une région extraite d'un objet en mouvement par segmentation du flot optique, un degré de dangerosité est attribué à la trajectoire idéale de cette région. Ce degré de dangerosité est affecté par la combinaison de plusieurs sources de danger (position, vitesse, accélération, ...), en utilisant la théorie de Dempster-Shafer [Yager 1994]. Enfin, nous retenons la valeur maximum des degrés de danger correspondant aux trajectoires des différentes régions de l'objet comme étant le degré de danger présenté par ce dernier.

4.2 Prédiction des trajectoires idéales d'objets mobiles

4.2.1 Introduction

Dans ce chapitre, nous abordons le problème de l'évaluation de scénarii à l'aide d'un système de vidéosurveillance. Nous avons développé un système d'évaluation du danger lié aux situations potentiellement dangereuses aux PN, en utilisant les résultats de suivi d'objets présentés au chapitre précédent. Pour évaluer une situation potentiellement dangereuse, nous nous basons sur l'étude des trajectoires des objets en mouvement, en considérant différentes sources de danger.

Le processus d'évaluation du danger est illustré sur la figure 4.1. Etant donné un objet en mouvement, dans un premier temps, nous réalisons une segmentation de l'objet en plusieurs régions, à base du flot optique. Nous déterminons ensuite pour chaque région extraite sa trajectoire idéale, en utilisant une chaîne de Markov. Nous évaluons enfin le danger lié à l'objet, en utilisant la théorie de l'évidence.

4.2.2 Segmentation des objets à base du flot optique

Grâce au flot optique obtenu lors de l'étape de suivi (chapitre 3), les objets en mouvement sont segmentés en régions ayant des flots optiques homogènes.

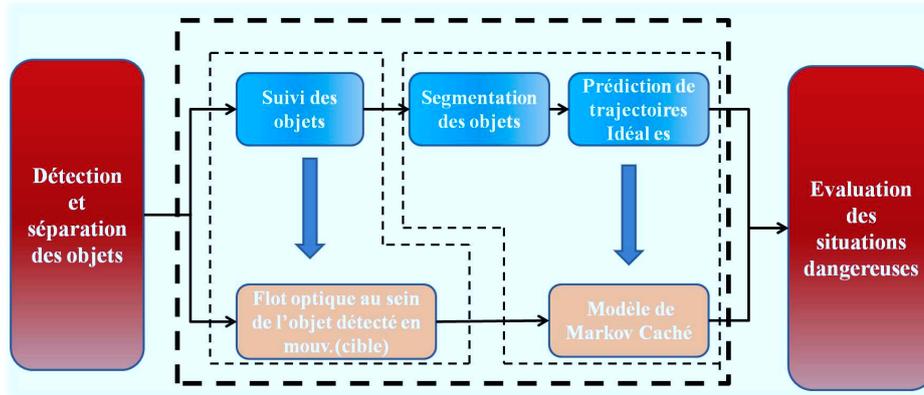


FIGURE 4.1 – Synoptique de l’approche d’évaluation du danger.

La méthode de segmentation en régions uniformes au sens du flot optique fait appel à un algorithme itératif. Etant donné un pixel, le traitement consiste à l’assigner à l’un des clusters déjà construits si la contrainte de similarité du flot optique est vérifiée. Cette contrainte, qui s’exprime sous la forme d’un seuillage, prend en compte simultanément la direction et le module du flot optique. Si la contrainte de similarité n’est pas respectée, le pixel forme un nouveau cluster. Les clusters obtenus sont ensuite représentés sous forme de régions rectangulaires (la largeur et la hauteur sont parallèles aux axes de l’image). Enfin, seules les régions de taille significative sont retenues. Les régions de très grande taille sont subdivisées afin d’obtenir globalement sur l’objet étudiée des régions de taille uniforme.

La figure 4.2 illustre un exemple de suivi d’un objet en mouvement (défini par un cadre blanc), avec les résultats de segmentation du flot optique. L’objet est partitionné en plusieurs régions, définies par des rectangles rouges. Le centre de chaque rectangle (de couleur jaune) est utilisé à l’étape suivante pour estimer la trajectoire idéale de la région correspondante.

4.2.3 Estimation des trajectoires idéales

4.2.3.1 Modèle de Markov Caché : définition des paramètres

Soit un objet en mouvement et soit une région issue de la segmentation du flot optique de cet objet. Pour mettre en oeuvre le modèle d’évaluation du danger, nous avons défini deux types de trajectoires pour le centre de la région :

- Trajectoire idéale courante : elle correspond à la trajectoire que devrait suivre le centre de la région afin d’éviter une situation potentiellement dangereuse. La trajectoire idéale courante correspond à une ligne droite du centre de la région vers la barrière du passage à niveau, comme indiqué sur la figure 4.3.

- Trajectoire idéale prédite : elle est estimée à partir du flot optique du centre de la région. Elle représente la trajectoire que devrait suivre le centre de la région pour



FIGURE 4.2 – Segmentation d'un objet en mouvement par clustering du flot optique.

revenir à la trajectoire idéale courante. Pour l'estimation des trajectoires idéales des centres des régions, nous avons développé une approche basée sur l'utilisation des HMM (Hidden Markov Model ou modèle de Markov caché).

Il s'agit à présent de définir les paramètres nécessaires à l'élaboration de notre approche d'estimation de trajectoires idéales. Ces paramètres dépendent de l'environnement à surveiller. Dans notre cas, il s'agit de l'environnement d'un passage à niveau (Fig. 4.3). Ces paramètres sont :

- La vitesse V_i et l'accélération a_i du centre de la région à étudier, avec $i \in [k, 1]$.
- La direction o_i de la " trajectoire idéale courante " du centre de la région, avec $i \in [k, 1]$.
- La distance D entre le centre de la région et la barrière du passage à niveau.
- Les intervalles de vitesses $[V_i, V_{i-1}]$, accélérations $[a_i, a_{i-1}]$ et orientations $[o_i, o_{i-1}]$ idéales définis tout au long de la trajectoire idéale courante, avec $i \in [k, 1]$.
- La position des barrières du passage à niveau (ouvert, fermé).

4.2.3.2 Prédiction des trajectoires idéales

L'objectif de notre méthode est d'estimer pour une région donnée la trajectoire idéale future permettant de retourner ou de rester sur la trajectoire idéale courante. Pour ce faire, nous utilisons un modèle de Markov où un état représente la position du centre de la région. La figure 4.4 illustre un exemple de modèle de Markov

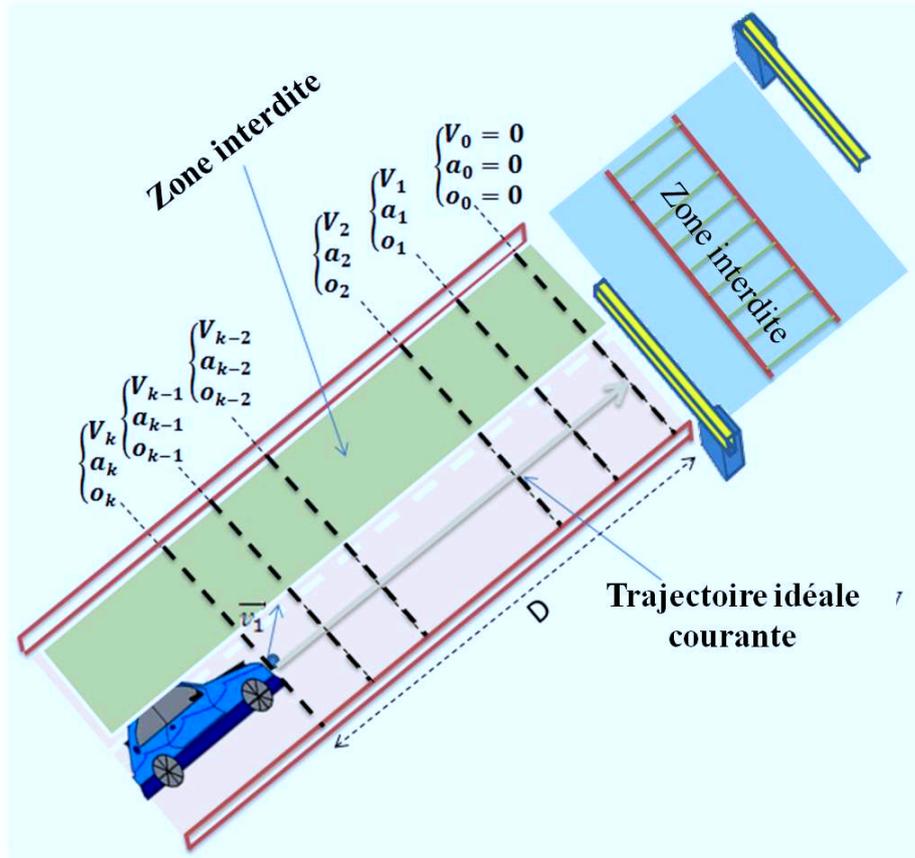


FIGURE 4.3 – Paramètres liés à l’environnement du passage à niveau nécessaires à l’estimation d’une trajectoire idéale.

mettant en oeuvre le processus d’estimation de la trajectoire idéale future à partir de la position courante de la région : q_t représente l’état courant (position du centre de la région considérée), u_t et b_t représentent respectivement l’observation et sa probabilité (définie selon une distribution normale) à l’instant t . Une observation est définie par la position, la vitesse, l’orientation et l’accélération du centre de la région. La probabilité de distribution de la variable d’état caché q_t ne dépend que de la variable d’état caché q_{t-1} à l’instant $t - 1$. De même, la distribution de l’observation u_t ne dépend que de l’état courant q_t . Le processus d’estimation de la trajectoire idéale est réalisé grâce aux algorithmes Forward-Backward, Viterbi et Baum-Welch [Baum 1968][Viterbi 1967][Baum 1970][Rabiner 1989].

La figure. 4.5 illustre l’estimation de la trajectoire idéale future à l’instant t à partir de la position et la trajectoire idéale courante du centre de la région. Étant donné le vecteur vitesse \vec{V}^t du centre de la région à l’instant t , calculé à partir du flot optique, l’état q_{t+1} est atteint, à partir de l’état q_t , avec une valeur de probabilité égale à 1. En effet, à l’instant t , le vecteur vitesse \vec{V}^t est connu, et donc, l’état

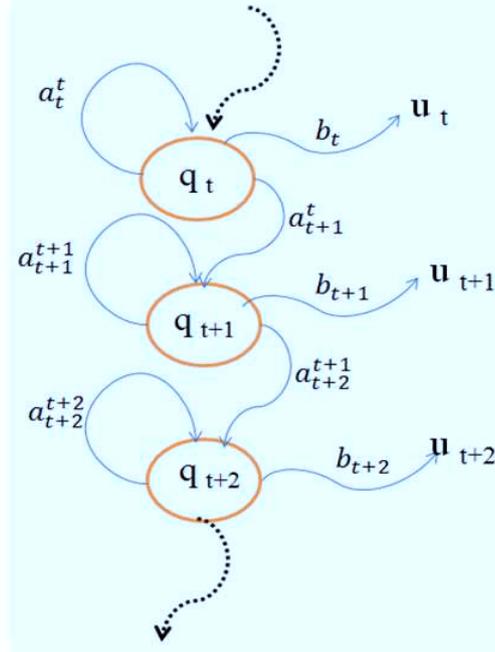


FIGURE 4.4 – Exemple de modèle de Markov pour l'estimation d'une trajectoire.

q_{t+1} est atteint de manière déterministe. Compte tenu des intervalles de vitesses, accélérations et orientations idéales, nous estimons tout d'abord le vecteur vitesse \overrightarrow{V}^{t+1} à l'instant $t+1$. Ensuite, l'état q_{t+2} est déterminé dans une zone de recherche (Fig. 4.5), définie par un cercle C de centre l'extrémité du vecteur \overrightarrow{V}^{t+1} et de rayon la valeur maximale de l'accélération calculée dans l'intervalle des accélérations idéales.

La probabilité de transition a_j^i et la probabilité d'observation $b_j(k)$ sont définies de la façon suivante :

$$a_j^i = P[q_t = S_j | q_{t-1} = S_i] \quad 1 \leq i, j \leq N \quad (4.1)$$

$$b_j(k) = P[u_t = o_k | q_t = S_j] \quad 1 \leq j \leq N \quad 1 \leq k \leq M \quad (4.2)$$

où a_j^i est la probabilité de transition de l'état S_i à l'instant $t-1$ à l'état S_j à l'instant t . S_j correspond à un état (position du centre de la région) parmi les N états possibles S_1, S_2, \dots, S_N à l'instant t . De la même manière, S_i correspond à un état parmi N états possibles S'_1, S'_2, \dots, S'_N à l'instant $t-1$. u_t représente l'observation actuelle à l'instant t . o_k correspond à une observation parmi les M observations possibles o_1, o_2, \dots, o_M à l'instant t . $b_j(k)$ est la probabilité d'observer o_k à l'instant t sachant que l'état courant q_t est égal à S_j . Nous calculons ces fonctions de probabilité à l'aide de formes gaussiennes :

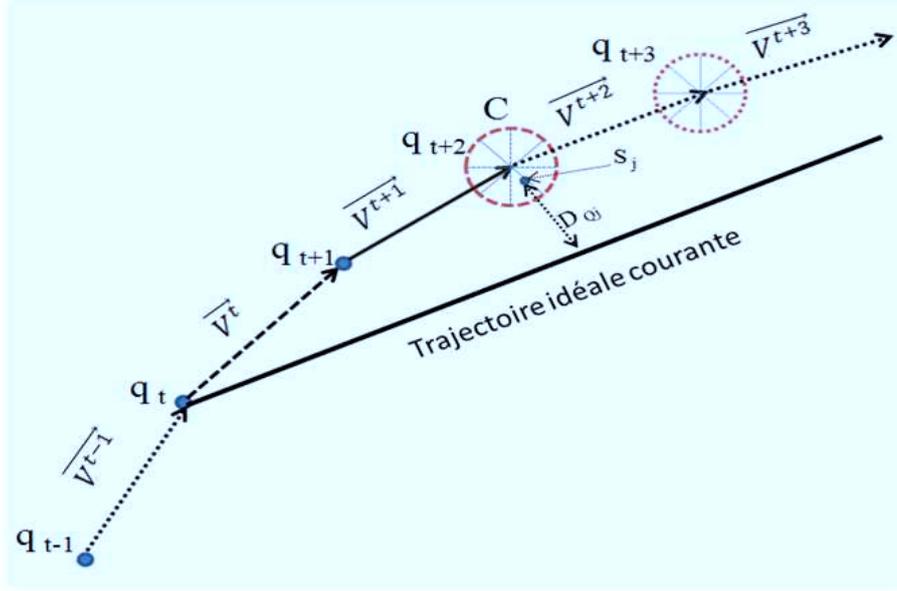


FIGURE 4.5 – Illustration de l'estimation de la trajectoire idéale à partir d'une trajectoire idéale courante.

$$a_j^i = S_Q + P_{Qj} \quad (4.3)$$

$$P_{Qj} = \int_{D_{Qj}-d_q}^{D_{Qj}} G_{m_Q, \sigma_Q}(x) dx \quad d_q = \frac{\sigma_Q}{15} \quad (4.4)$$

$$S_Q = 1 - \sum_{j=1}^N (P_{Qj}) \quad (4.5)$$

où $G_{m_Q, \sigma_Q}(x)$ est une distribution gaussienne de la variable x , représentant la distance de l'état courant q_t (position du centre de la région) dans le cercle C (Fig. 4.5) à la trajectoire idéale courante, de moyenne $m_Q = 0$ et de variance σ_Q^2 des N positions possibles de l'état courant q_t .

D_{Qj} est la distance entre la position du centre de la région (représentée par l'état $q_t = S_j$) et la trajectoire idéale courante à l'instant t (voir Fig. 4.5).

Sachant que d_q est une petite variation de distance, P_{Qj} représente la probabilité que la distance de l'état courant q_t (position du centre de la région) à la trajectoire idéale courante soit proche de D_{Qj} .

Nous rappelons que nous utilisons une observation composée de la position, de la vitesse, de l'orientation et de l'accélération du centre de la région. La probabilité d'observation $b_j(k)$ est calculée à partir des probabilités d'observation pour l'accélération (sachant que la position, la vitesse et l'accélération sont des grandeurs

liées) et l'orientation. La probabilité d'observation s'exprime alors de la manière suivante :

$$b_j(k) = \sum_{j=1}^M p_{ac}^j(k) * p_{or}^j(k) \quad (4.6)$$

$$p_{ac}^j(k) = S_a^k + P_{aj}(k) \quad p_{or}^j(k) = S_o^k + P_{oj}(k) \quad (4.7)$$

$$P_{aj}(k) = \int_{Ac_j^k - d_a}^{Ac_j^k} G_{m_a, \sigma_a}(x) dx \quad P_{oj}(k) = \int_{Oc_j^k - d_o}^{Oc_j^k} G_{m_o, \sigma_o}(x) dx \quad (4.8)$$

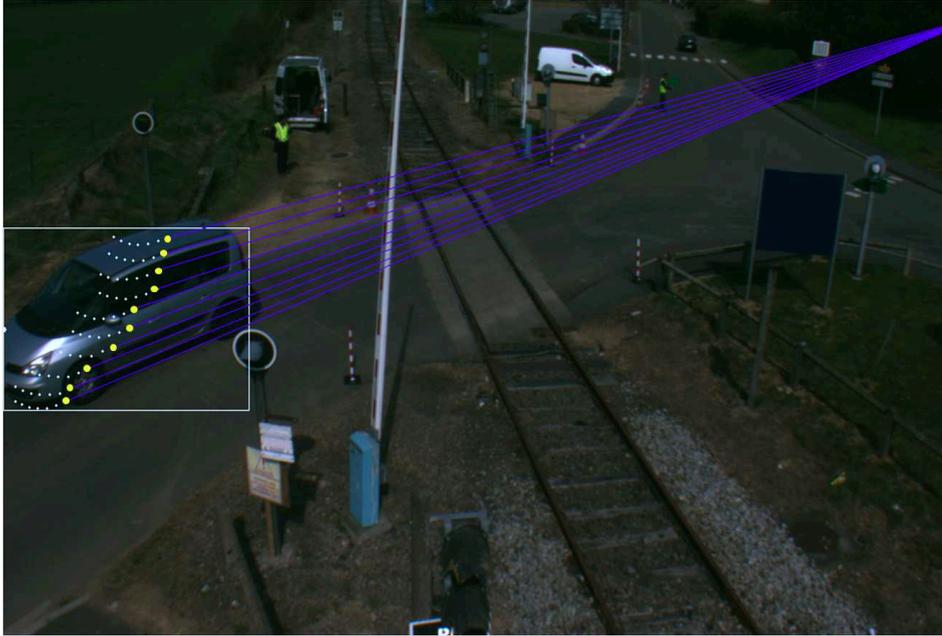
$$d_a = \frac{\sigma_a}{15} \quad d_o = \frac{\sigma_o}{15} \quad (4.9)$$

$$S_a^k = 1 - \sum_{j=1}^N (P_{aj}) \quad S_o^k = 1 - \sum_{j=1}^N (P_{oj}) \quad (4.10)$$

$G_{m_a, \sigma_a}(x)$ est une distribution gaussienne de la variable x , représentant l'accélération de l'état courant q_t (position du centre de la région) dans le cercle C (Fig. 4.5), de moyenne m_a et de variance σ_a^2 . m_a et σ_a sont respectivement la moyenne et l'écart type des M accélérations observées dans le cercle C (Fig. 4.5). Ac_j^k est l'accélération observée à l'instant t . Sachant que d_a est une petite variation d'accélération, $P_{aj}(k)$ représente la probabilité que l'accélération de l'état courant q_t (position du centre de la région) soit proche de Ac_j^k .

$G_{m_o, \sigma_o}(x)$ est une distribution gaussienne de la variable x , représentant l'orientation de la vitesse de l'état courant q_t (position du centre de la région) dans le cercle C (Fig. 4.5), de moyenne m_o et de variance σ_o^2 . m_o et σ_o sont respectivement la moyenne et l'écart type des M orientations observées dans le cercle C (Fig. 4.5). Oc_j^k est l'orientation observée à l'instant t . Sachant que d_o est une petite variation d'orientation, $P_{oj}(k)$ représente la probabilité que l'orientation de la vitesse de l'état courant q_t (position du centre de la région) soit proche de Oc_j^k .

La figure 4.6 représente la trajectoire idéale prédite du centre (points de couleur jaune) de chaque région extraite des objets en mouvement détectés. Les points blancs représentent le déplacement prédit du centre des régions. Les lignes bleues sont les lignes de fuites dans l'image passant par les centres des régions. Elles représentent les trajectoires idéales courantes. On peut observer dans cette figure la direction des trajectoires idéales prédites durant le déplacement d'une voiture s'éloignant du PN (a) et d'une voiture s'approchant du PN (b).



(a) Véhicule s'éloignant du PN



(b) Véhicule s'approchant du PN

FIGURE 4.6 – Exemples de trajectoires idéales prédites.

4.3 Estimation du danger de situations potentiellement dangereuses

4.3.1 Introduction

Pour évaluer le niveau de dangerosité d'une situation, il faut au préalable déterminer les sources pouvant en être responsables. Par la suite, il faut fusionner ces sources de danger pour évaluer le risque de l'évènement étudié. Dans ce travail, nous utilisons la théorie de l'évidence [Smets 1990][Yager 1994] pour traiter ce problème de fusion de données. Cette théorie permet d'étudier des événements non nécessairement exclusifs. Elle permet aussi de représenter explicitement l'incertitude sur un événement. Elle présente l'aptitude à représenter parfaitement ce qui est déjà connu, en prenant en compte en même temps ce qui reste inconnu [Shafer 1976].

Dans cette section, nous évaluons le degré de dangerosité lié à la trajectoire idéale prédite du centre de chacune des régions extraites de l'objet en mouvement considéré. Nous rappelons que cette extraction de régions est effectuée grâce à une segmentation du flot optique de l'objet.

4.3.2 Théorie de l'évidence : définition des paramètres

4.3.2.1 Cadre de discernement

Nous définissons au départ un ensemble d'hypothèse H , appelé cadre de discernement en théorie de l'évidence :

$$H = \{D, S\} \quad (4.11)$$

où D, S représentent respectivement les hypothèses de danger et de sécurité.

A partir du cadre de discernement H , on construit l'ensemble noté 2^H , composé de toutes les disjonctions possibles de ces hypothèses :

$$2^H = \{\emptyset, D, S, H\} \quad (4.12)$$

où l'élément \emptyset désigne une hypothèse non incluse dans le cadre de discernement H .

Le choix d'un cadre de discernement H de petite taille permet d'une part de simplifier la complexité algorithmique qui est directement liée à la taille de ce cadre et d'autre part d'éviter l'aspect conflictuel existant entre deux sources d'information (voir section 4.3.2.4).

4.3.2.2 Jeux de masses

On appelle jeu de masses une fonction $m : 2^H \rightarrow [0; 1]$ qui vérifie les propriétés suivantes :

$$m(\emptyset) = 0 \quad (4.13)$$

$$\sum_{A \subseteq H} m(A) = 1 \quad (4.14)$$

Pour chaque partie A de H , $m(A)$ représente la croyance que l'on place dans la proposition A . Cette masse unité est répartie sur les éléments de l'ensemble 2^H . Cette répartition dépend de la connaissance sur la source que l'on cherche à modéliser. Dans notre cas, soit $m^i(A)$ la masse de croyance d'un élément donné A de l'ensemble 2^H , liée à la source de danger ayant le numéro i . On considère alors que chaque jeu de masses engage une croyance sur deux hypothèses seulement :

- d'une part l'hypothèse de danger D (singleton). Soit $m^i(D) = x$, qui signifie que je crois l'existence d'un danger D sur la source numéro i avec un degré x .

- d'autre part l'hypothèse H ($H = D \cup S$). Soit $m^i(H) = 1 - x$, qui signifie que je ne crois pas du tout à l'hypothèse de sécurité $\bar{D} = S$ ($m^i(S) = 0$). La croyance restante ($1 - x$) est appelée incertitude.

4.3.2.3 Fonction de crédibilité

La fonction de crédibilité $Cr^i(A)$ permet de prendre en compte toute la croyance placée sur A . $Cr^i(A)$ est calculée de la manière suivante :

$$Cr^i(A) = \sum_{B \subseteq A} m^i(B) \quad (4.15)$$

Dans notre cas, les crédibilités des différentes hypothèses seront alors :

$$Cr^i(D) = x \quad (4.16)$$

$$Cr^i(S) = 0 \quad (4.17)$$

$$Cr^i(H) = 1 - x \quad (4.18)$$

Ceci montre que $Cr^i(A) = m^i(A)$. Cela signifie que toutes les informations données par la source numéro i soutient uniquement l'hypothèse de danger D ($Cr^i(\bar{D}) = Cr^i(S) = 0$).

4.3.2.4 Combinaison de sources de danger (Loi de combinaison de Dempster)

L'étape suivante consiste à fusionner M ($M = 5$) sources d'information distinctes définies sur le même cadre de discernement $H = \{D, S\}$. Chaque source S_i ($i = 1 \dots M$) émet son propre jugement sur le problème donné pour évaluer le risque de l'évènement étudié.

Les sources de danger que nous avons identifiées sont liées à : la position, la vitesse (en module et en direction), l'accélération et la distance entre le centre de la région considérée et la trajectoire idéale courante. Ensuite, nous combinons les

4.3. Estimation du danger de situations potentiellement dangereuses 97

masses élémentaires correspondantes à l'aide de la règle de Dempster pour calculer une masse de croyance globale correspondant au degré de dangerosité de la région étudiée.

L'intérêt de la fusion multi-sources est qu'elle permet d'accéder à un jugement plus fiable. Cette fusion repose justement sur la possibilité de construire une masse de croyance unique $m(A)$ résultant de la combinaison de toutes les masses de croyance $m^i(A)$ issues de plusieurs sources de danger distinctes. Considérons dans un premier temps, le cas de deux sources de danger distinctes S_1 et S_2 . Soient $m^1(A)$ et $m^2(A)$ les masses de croyance correspondantes. La combinaison est illustrée par la figure 4.7.

		masses assignées à $D \cap S$			masses assignées à D			masses assignées à S	
$A_1 \rightarrow$	$m^1(D)$	$m^1(D).m^2(D)$	0	$m^1(D).(1 - m^2(D))$					
$A_2 \rightarrow$	$m^1(S) = 0$	0	0	0					
$A_3 \rightarrow$	$m^1(H) = 1 - m^1(D)$	$m^2(D).(1 - m^1(D))$	0	$(1 - m^1(D)).(1 - m^2(D))$					
		$m^2(D)$	$m^2(S) = 0$	$m^2(H) = 1 - m^2(D)$					
		\uparrow B_1	\uparrow B_2	\uparrow B_3					

FIGURE 4.7 – Représentation graphique de la fusion de deux sources de danger distinctes selon la loi de combinaison de Dempster.

On distribue les masses de la source S_1 verticalement à gauche du tableau (cases brunes) et les masses de la source S_2 horizontalement en bas du tableau (cases bleues). Chaque case du tableau est définie alors par une masse $m^1(A_i)$ ($i = 1, 2, 3$) et une masse $m^2(B_j)$ ($j = 1, 2, 3$). A_i et B_j représentent respectivement une proposition (D , S ou H) de S_1 et S_2 . La masse correspondant à chaque case du tableau est alors égale à $m^1(A_i).m^2(B_j)$ et elle est assignée à la proposition $A = A_i \cap B_j$. La masse totale affectée à la proposition A est alors égale à :

$$m(A) = \sum_{A_i \cap B_j = A} m^1(A_i).m^2(B_j) \quad (4.19)$$

Il faut bien noter que, du fait que les masses correspondant à $D \cap S$ sont nulles sur la figure 4.7, il n'y a pas de conflit entre nos hypothèses durant la combinaison. Le jeu de masse résultant de cette combinaison est :

$$m(D) = m^1(D).m^2(D) + m^1(D).(1 - m^2(D)) + m^2(D).(1 - m^1(D)) \quad (4.20)$$

$$m(S) = 0 \quad (4.21)$$

$$m(H) = (1 - m^1(D)) * (1 - m^2(D)) \quad (4.22)$$

Les résultats de combinaison obtenus sont alors généralisés pour un ensemble de M masses de croyance, en appliquant itérativement l'équation 4.19.

Les étapes décrites plus haut sont appliquées pour estimer le degré de dangerosité de chacune des régions composant l'objet. Nous attribuons ensuite à l'objet le degré de dangerosité maximum des degrés calculés pour les régions qui le composent.

4.3.3 Modèle d'évaluation des trajectoires idéales

Le problème provient essentiellement de la façon d'analyser la dangerosité de la trajectoire idéale prédite du centre de la région considérée. Comme indiqué précédemment, cette trajectoire est estimée en utilisant un processus (section 4.2) basé sur un modèle de Markov caché. À chaque instant $(t, t + 1, \dots, t + t_f)$ de la trajectoire idéale prédite (Fig. 4.8), nous prenons en compte plusieurs paramètres pour évaluer la situation : la vitesse (v_t, \dots, v_{t+t_f}) , l'accélération (a_t, \dots, a_{t+t_f}) , l'orientation (o_t, \dots, o_{t+t_f}) , la position (p_t, \dots, p_{t+t_f}) et la distance (D_t, \dots, D_{t+t_f}) entre le centre de la région et la trajectoire idéale courante.

4.3.4 Définition des sources potentielles de danger

La première source de danger provient d'une vitesse pratiquée dépassant une vitesse nominale. La vitesse nominale est calculée par rapport à l'environnement du passage à niveau, de la manière suivante :

$$V_N = \frac{D_{max}}{2 * n * T_{min}} \quad (4.23)$$

où D_{max} (en pixels) est la distance maximale parcourue par un objet dans l'image, T_{min} (en secondes) est le temps minimum pour parcourir la distance D_{max} et n (images/seconde) est le nombre d'images acquises par seconde par le système de recueil.

On peut alors écrire la masse de croyance $m_v^1(D)$ résultant de cette source de danger, comme suit :

$$m_v^1(D) = \begin{cases} 0.01 & si & (V_c^{t_f} - V_N) \leq 0 \\ \frac{P_v - 0.5}{0.5} & si & (V_c^{t_f} - V_N) > 0 \end{cases} \quad (4.24)$$

$$P_v = \int_{-\infty}^{V_c^{t_f}} G_{V_N, \sigma_v}(x) dx \quad \sigma_v = \frac{V_N}{4} \quad (4.25)$$

$G_{V_N, \sigma_v}(x)$ est une distribution gaussienne de la vitesse x du centre de la région, de

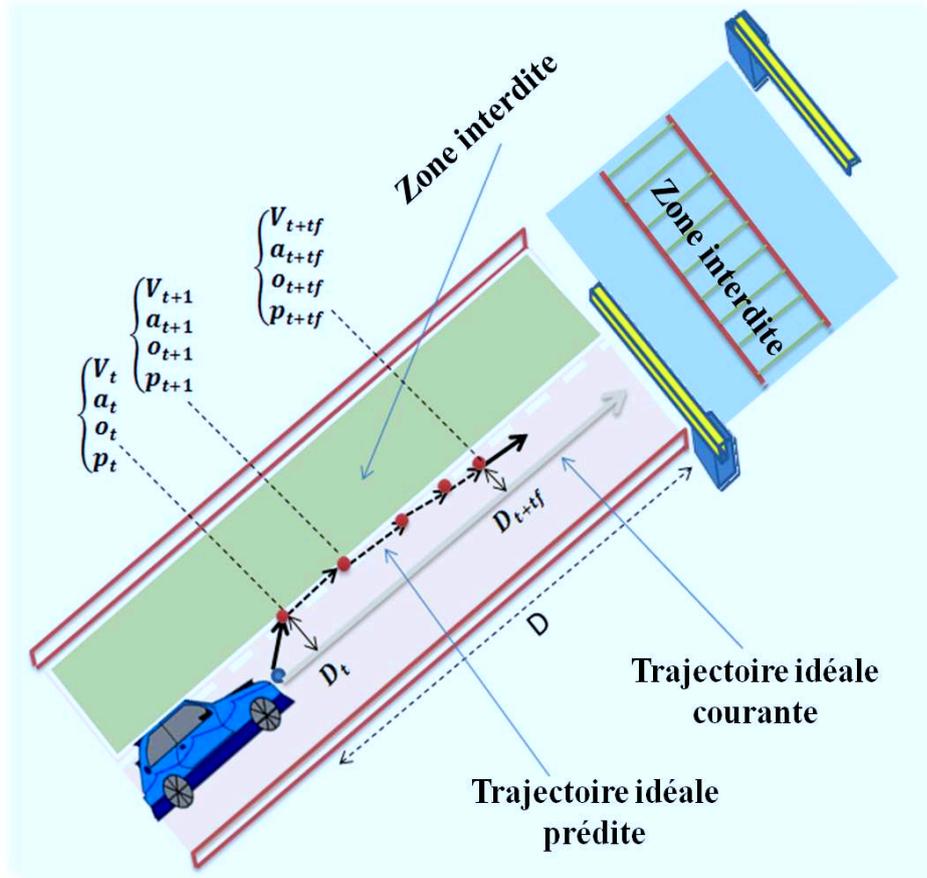


FIGURE 4.8 – Analyse de la trajectoire idéale prédite.

moyenne V_N et de variance σ_v^2 . $V_c^{t_f}$ est la vitesse du centre de la région prédite à l'instant t_f (dernière instant de la trajectoire idéale prédite).

$$t_f = t_i + T_{mise\,jour} \quad (4.26)$$

où t_i exprime le temps de début de calcul de la trajectoire idéale prédite. $T_{mise\,jour}$ est la durée durant laquelle l'objet devrait parcourir sa trajectoire idéale prédite, elle représente également la période de mise à jour de la trajectoire idéale durant le processus d'analyse du mouvement de l'objet.

L'écart type σ_v de la gaussienne est calculée de manière expérimentale. La valeur de l'écart type $\sigma_v = \frac{V_N}{4}$ est directement proportionnelle à la valeur de la moyenne V_N de la gaussienne. Ce choix indique que la masse de croyance $m_v^1(D)$ s'approche de sa valeur maximale 1 (c'est-à-dire le degré de danger provenant de cette source est presque égal à 100%) si la vitesse $V_c^{t_f}$ est plus grande que la moyenne V_N au minimum d'un écart égal à $3\sigma_v = \frac{3V_N}{4}$.

La lecture de la masse de croyance $m_v^1(D)$ se fait de la manière suivante : si la vitesse prédite $V_c^{t_f}$ à l'instant t_f (au bout de la trajectoire idéale prédite) est inférieure à la vitesse nominale V_N , une faible probabilité constante est assignée à la masse $m_v^1(D)$ (danger provenant de la vitesse est moins probable); sinon plus la vitesse $V_c^{t_f}$ est élevée plus la masse de croyance $m_v^1(D)$ est grande (danger plus probable).

Une fois la période $T_{misejour}$ est écoulée, une nouvelle trajectoire idéale est prédite et le système d'évaluation est réinitialisé, avec :

$$\begin{cases} t_i = t_f \\ t_f = t_f + T_{misejour} \end{cases} \quad (4.27)$$

La deuxième source de danger concerne l'accélération de l'objet le long de la trajectoire parcourue. La masse de croyance $m_a^2(D)$ résultant de cette source de danger est construite de la façon suivante :

$$m_a^2(D) = \begin{cases} 0.01 & \text{si } (a_c^{t_f} - a_N) \leq 0 \\ \frac{P_a - 0.5}{0.5} & \text{si } (a_c^{t_f} - a_N) > 0 \end{cases} \quad (4.28)$$

$$P_a = \int_{-\infty}^{a_c^{t_f}} G_{a_N, \sigma_a}(x) dx \quad \sigma_a = \frac{a_N}{4} \quad (4.29)$$

$$a_c^{t_f} = \frac{V_c^{t_f} - V_c^{t_i}}{n * T_{misejour}} \quad (4.30)$$

$$a_N = -N_{misejour} * \frac{V_c^*}{n * T_{min}} \quad (4.31)$$

$G_{a_N, \sigma_a}(x)$ est une distribution gaussienne de l'accélération x du centre de la région, de moyenne a_N et de variance σ_a^2 . $a_c^{t_f}$ est l'accélération du centre de la région prédite à l'instant t_f . V_c^* est la vitesse du centre de la région calculée à l'instant $t_i = t_f$ (c'est-à-dire lors de la réinitialisation du système d'évaluation pour prédire une nouvelle trajectoire idéale). $N_{misejour}$ est un nombre (sans unité) représentant le nombre d'images acquises par le système de recueil durant la durée $T_{misejour}$.

L'écart type σ_a de la gaussienne est calculée de manière expérimentale. La valeur de l'écart type $\sigma_a = \frac{a_N}{4}$ est directement proportionnelle à la valeur de la moyenne a_N de la gaussienne. Ce choix indique que la masse de croyance $m_a^2(D)$ s'approche de sa valeur maximale 1 (c'est-à-dire le degré de danger provenant de cette source est presque égal à 100%) si l'accélération $a_c^{t_f}$ est plus grande que la moyenne a_N au minimum d'un écart égal à $3\sigma_a = \frac{3a_N}{4}$.

La lecture de la masse de croyance $m_a^2(D)$ se fait de la manière suivante : si l'accélération prédite $a_c^{t_f}$ à l'instant t_f (au bout de la trajectoire idéale prédite)

4.3. Estimation du danger de situations potentiellement dangereuses 101

est inférieure à l'accélération a_N , une faible probabilité constante est assignée à la masse $m_a^2(D)$ (danger provenant de l'accélération est moins probable); sinon plus l'accélération $a_c^{t_f}$ est élevée plus la masse de croyance $m_a^2(D)$ est grande (danger plus probable).

La troisième source de danger provient du changement d'orientation de la vitesse du centre de la région étudiée par rapport à l'orientation de la trajectoire idéale courante. La masse de croyance $m_o^3(D)$ résultant de cette source de danger est exprimée de la manière suivante :

$$m_o^3(D) = \frac{|P_o - 0.5|}{0.5} \quad (4.32)$$

$$P_o = \int_{-\infty}^{o_c^{t_f}} G_{o_N, \sigma_o}(x) dx \quad \sigma_o = \frac{2 * \pi}{7} \quad (4.33)$$

$G_{o_N, \sigma_o}(x)$ est une distribution gaussienne de l'orientation x de la vitesse du centre de la région, de moyenne o_N (direction de la trajectoire idéale courante) et de variance σ_o^2 . $o_c^{t_f}$ est l'orientation de la vitesse du centre de la région à l'instant t_f .

La valeur de l'écart type $\sigma_o = \frac{2*\pi}{7}$ est déterminée de manière expérimentale. Elle indique que la masse de croyance $m_o^3(D)$ s'approche de sa valeur maximale 1 (c'est-à-dire le degré de danger provenant de cette source est presque égal à 100%) si l'orientation $o_c^{t_f}$ est plus grande que l'orientation moyenne o_N au minimum d'un écart égal à $3\sigma_o = \frac{6*\pi}{7}$. En réalité, cette valeur maximale n'est jamais atteinte puisque dans notre cas la différence d'orientation ne peut excéder $\frac{\pi}{2}$. Par conséquent, le degré de danger maximum (100%) associé à cette source n'est jamais atteint.

La masse de croyance $m_o^3(D)$ exprime le fait que plus l'orientation de la vitesse du centre de la région considérée s'écarte de celle de la trajectoire idéale courante, plus le danger provenant de cette source est probable.

La quatrième source de danger est liée à la position du centre de la région considérée dans la zone du passage à niveau. La masse de croyance $m_p^4(D)$ représentant le degré de dangerosité de cette source, dépend de la distance entre le centre de la région et la barrière du passage à niveau. Elle est définie de la façon suivante :

$$m_p^4(D) = \frac{|P_d - 0.5|}{0.5} \quad (4.34)$$

$$P_d = \int_{-\infty}^{d_c^{t_f}} G_{d_N, \sigma_d}(x) dx \quad \sigma_d = \sqrt{(D_{max})} \quad d_N = 0 \quad (4.35)$$

$G_{d_N, \sigma_d}(x)$ est une distribution gaussienne de la variable x représentant une grandeur composée des fonctions suivantes :

- la position du centre de la région, en terme de présence dans une zone interdite (à l'extérieur du PN) ou non.

- la distance du centre de la région par rapport à la barrière du PN.
- la vitesse du centre de la région.

$d_N = 0$ et σ_d^2 sont respectivement la moyenne et la variance de la distribution $G_{d_N, \sigma_d}(x)$. D_{max} est la distance maximale parcourue par un objet dans l'image. d_c^{tf} est une fonction liée aux différentes informations décrite ci-dessus. Le calcul de cette fonction est donné en annexe (B). Plus d_c^{tf} prend des valeurs élevées plus le danger augmente.

La valeur de l'écart type $\sigma_d = \sqrt{(D_{max})}$ est déterminée de manière expérimentale. Ce choix indique que la masse de croyance $m_p^4(D)$ s'approche de sa valeur maximale 1 (c'est-à-dire le degré de danger provenant de cette source est presque égal à 100%) si la fonction d_c^{tf} est plus grande que la moyenne d_N au minimum d'un écart égal à $3\sigma_d = 3\sqrt{(D_{max})}$.

La cinquième source de danger provient de la distance entre le centre de la région étudiée et la trajectoire idéale courante. La masse de croyance $m_D^5(D)$ liée à cette source de danger est définie de la façon suivante :

$$m_D^5(D) = \frac{|P_D - 0.5|}{0.5} \quad (4.36)$$

$$P_D = \int_{-\infty}^{D^{tf}} G_{D_N, \sigma_D}(x) dx \quad \sigma_D = 2V_N \quad D_N = 0 \quad (4.37)$$

$G_{D_N, \sigma_D}(x)$ est une distribution gaussienne de la distance x du centre de la région par rapport à la trajectoire idéale courante, de moyenne D_N et de variance σ_D^2 . D^{tf} est la distance du centre de la région par rapport à la trajectoire idéale courante à l'instant t_f .

La valeur de l'écart type $\sigma_D = 2V_N$ est déterminée de manière expérimentale. Ce choix indique que la masse de croyance $m_D^5(D)$ s'approche de sa valeur maximale 1 (c'est-à-dire le degré de danger provenant de cette source est presque égal à 100%) si la distance D^{tf} est plus grande que la moyenne D_N au minimum d'un écart égal à $3\sigma_D = 6V_N$.

Plus la distance du centre de la région par rapport à la trajectoire idéale courante est élevée plus la masse de croyance $m_D^5(D)$ est grande (le danger lié à cette source devient de plus en plus probable).

Enfin, en se basant sur la combinaison de Dempster [Yager 1994] (voir section 4.3.2.4), on construit une masse de croyance unique par combinaison de toutes les masses élémentaires représentant les sources de danger :

$$Danger = Dempster - Shafer \left(m_v^1(D), m_a^2(D), m_o^3(D), m_p^4(D), m_D^5(D) \right) \quad (4.38)$$

4.3.5 Condition d'indépendance des sources de danger

Pour pouvoir combiner les sources, celles-ci doivent être indépendantes les unes des autres. Bien que cette notion d'indépendance reste dans la pratique une notion ambiguë dans le cadre de la théorie de l'évidence, des sources sont considérées comme étant indépendantes si chacune donne son avis sur le problème traité sans avoir connaissance de l'avis des autres. Dans notre cas, chaque source évalue le danger par rapport à ses propres paramètres liés à l'environnement du passage à niveau. Par exemple, pour la source de danger lié à la vitesse V_c^{tf} (V_c^{tf} est obtenue par le modèle de Markov proposé), le danger est évalué par comparaison de la vitesse V_c^{tf} et la vitesse nominale V_N (équ.4.23). V_N étant la vitesse maximale que peut atteindre un objet sur l'image. Prenant maintenant, la source de danger concernant l'accélération a_c^{tf} (a_c^{tf} est obtenue par le modèle de Markov proposé). Le danger lié à cette source est évalué par comparaison de l'accélération a_c^{tf} et l'accélération nominale a_N calculée indépendamment de la vitesse nominale V_N . a_N étant l'accélération du centre de la région prédite lors de la réinitialisation du système d'évaluation pour prédire une nouvelle trajectoire idéale. Ces deux sources peuvent être considérées comme indépendantes puisque l'évaluation du danger qui leur est associé est basée sur des paramètres indépendants. Le même raisonnement est adopté pour justifier l'indépendance des autres sources.

4.4 Résultats expérimentaux

L'objectif premier du système de perception de l'environnement PN, développé dans le cadre du projet PANsafer, est la détection de situations potentiellement dangereuses telles que la présence d'un obstacle sur le platelage, la détection d'une formation de file avec des véhicules sur la zone de danger ou le passage en chicane. Le processus de détection est basé sur l'analyse du " comportement " des objets évoluant dans l'environnement PN. Les résultats présentés ici sont issus de tests effectués sur le PN de Mouzon (nord de la France). Après l'installation et les opérations de calibrage du système de prise de vues, différents scénarios ont été joués et enregistrés.

Cinq scénarios sont considérés et traités par notre système de perception. Ils correspondent aux principales situations reconnues comme dangereuses pour un PN : présence d'obstacles, remontée de file, passage en chicane sans arrêt sur platelage, passage en chicane avec arrêt sur platelage, chute de piétons sur platelage. Dans ce qui suit, nous présentons les résultats de détection et d'évaluation du danger selon la procédure décrite dans la section 4.3, sur les différents scénarios joués sur site.

Plusieurs acquisitions vidéo ont été effectuées (pour évaluer les algorithmes proposés) sur le site du PN à Mouzon dans les Ardennes (Fig. 4.9). Un système de perception stéréoscopique à deux caméras, placé à une hauteur maximale de 6 mètres, a été mis en place. La disposition géométrique du système stéréoscopique offre une bonne vue de la zone de PN, grâce à une vue plongeante des champs de vue des caméras (Fig. 4.9). Le système d'acquisition permet d'effectuer un enregistrement

non compressé sur disque dur, à une cadence proche de la cadence vidéo standard de 25 images/seconde. Les images ont été prises dans des conditions de prises de vues variable (en termes de luminosité, soleil, présence de vent...).

Bien que le système de prise de vue soit constitué de deux caméras, nous rappelons que notre approche est basée sur une analyse monoculaire ; seules les séquences issues d'une seule caméra sont utilisées dans toutes les étapes du processus de détection et d'évaluation du danger aux passages à niveau. Bien évidemment, l'approche proposée pourrait être étendue dans le cadre d'une analyse stéréoscopique.



FIGURE 4.9 – Installation du système de perception sur site (Mouzon-France).

4.4.1 Scénario 1 : Détection d'obstacles

Dans le scénario de la figure 4.10, un véhicule s'approche du PN, les barrières étant ouvertes. Durant la traversée du PN, le véhicule s'arrête sur le platelage, et devient par conséquent un obstacle. Le véhicule redémarre ensuite, puis quitte le PN. La figure 4.10 illustre les résultats de détection et d'évaluation du danger sur une suite d'images acquises par le système de prise de vues, installé sur le PN. Le véhicule est détecté et suivi durant toute la séquence. L'évaluation de son comportement montre que le danger augmente au fur et à mesure qu'il s'approche du PN. Lorsqu'il entre dans la zone de danger (entre les deux barrières), le danger devient maximum, avec une valeur de 46%. Notons que dans le cas de barrières ouvertes, le danger maximum dans la zone de danger est de 50% lorsque le véhicule est détecté en mouvement. Lorsque le véhicule devient immobile dans la zone de

danger, sa stationnarité est détectée. Le danger passe par conséquent à 100% de manière arbitraire pour pouvoir distinguer le danger associé à un objet fixe du danger associé à un objet mobile à l'intérieur de la zone du PN. Lorsque le véhicule redémarre, le danger diminue au fur et à mesure qu'il quitte le PN. Notons également la détection de la présence d'un piéton en mouvement en bas -à gauche dans les différentes images. Son comportement est évalué en fonction de sa position par rapport à la zone de danger.

La figure 4.11 montre l'évolution du danger lié à chacun des objets (véhicule et piéton) tout au long de la séquence d'images, avec la détection de la situation potentiellement dangereuse " obstacle fixe dans la zone PN" lorsque le véhicule s'arrête sur le platelage. Le piéton, détecté en mouvement dans la zone PN, peut être considéré également comme étant une source potentielle de danger.

4.4.2 Scénario 2 : Remontée de file

Dans le scénario 2, illustré sur la figure 4.12, un premier véhicule traverse le PN, puis s'arrête juste après la zone de danger, les barrières étant ouvertes. Ensuite, un deuxième véhicule s'engage dans la zone de danger, puis se bloque derrière le premier véhicule arrêté. Un troisième véhicule effectue la même manoeuvre, et se retrouve bloqué derrière la file formée par les deux premiers véhicules. Cette situation est très courante et se présente lorsqu'un véhicule se retrouve involontairement arrêté sur la zone de danger.

La figure 4.12 présente les résultats de détection et d'évaluation du danger sur une série d'images acquises par le système de prise de vues, installé sur le PN. Les trois véhicules sont détectés et suivis durant toute la séquence. Le danger induit par le premier véhicule augmente au fur et à mesure qu'il s'approche du PN. Lorsqu'il traverse la zone de danger, le danger estimé est de 46%. Comme précisé dans le scénario 1, le danger maximum est de 50% dans le cas de barrières ouvertes pour des objets en mouvement dans la zone de danger. Au fur et à mesure que le véhicule quitte le PN, le danger diminue jusqu'à ce qu'il devienne nul après sa sortie de la zone de danger. Lorsque le véhicule s'arrête, sa stationnarité est détectée. La détection de cette situation provoque alors une alerte d'un risque de remontée de file au PN. De la même manière, le danger induit par le deuxième véhicule augmente au fur et à mesure qu'il s'approche du PN. Lorsqu'il se bloque derrière le premier véhicule, sa stationnarité est détectée dans la zone de danger. Le danger passe par conséquent à 100% de manière arbitraire comme expliqué précédemment. La détection de cette situation provoque alors une alarme de présence d'une remontée de file au PN, avec un obstacle fixe dans la zone de danger. L'analyse du comportement du troisième véhicule s'effectue de la même manière. On détecte alors un deuxième obstacle dans la zone de danger, en maintenant l'alarme de la présence d'une remontée de file au PN. Lorsque le premier véhicule redémarre, la situation se débloque et les véhicules obstacles se remettent en mouvement, et le danger diminue au fur et à mesure qu'ils quittent le PN. Notons également la détection de la présence d'un piéton en mouvement en bas à gauche des images. Son comportement est évalué en fonction de sa



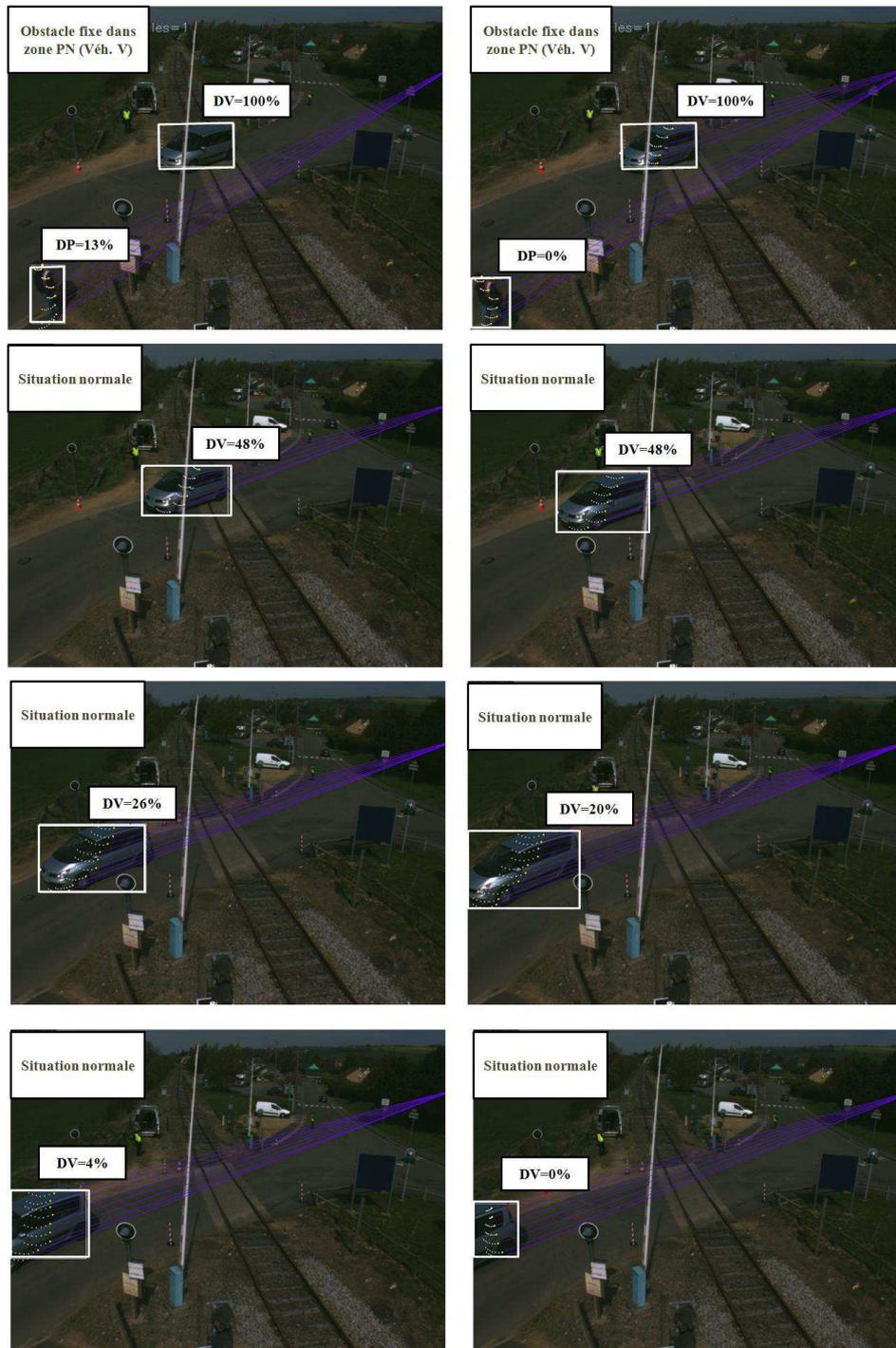


FIGURE 4.10 – Détection d'obstacles - DV : danger lié au véhicule ; DP : danger lié au piéton. Sens de lecture de la séquence : haut-bas, gauche-droite.

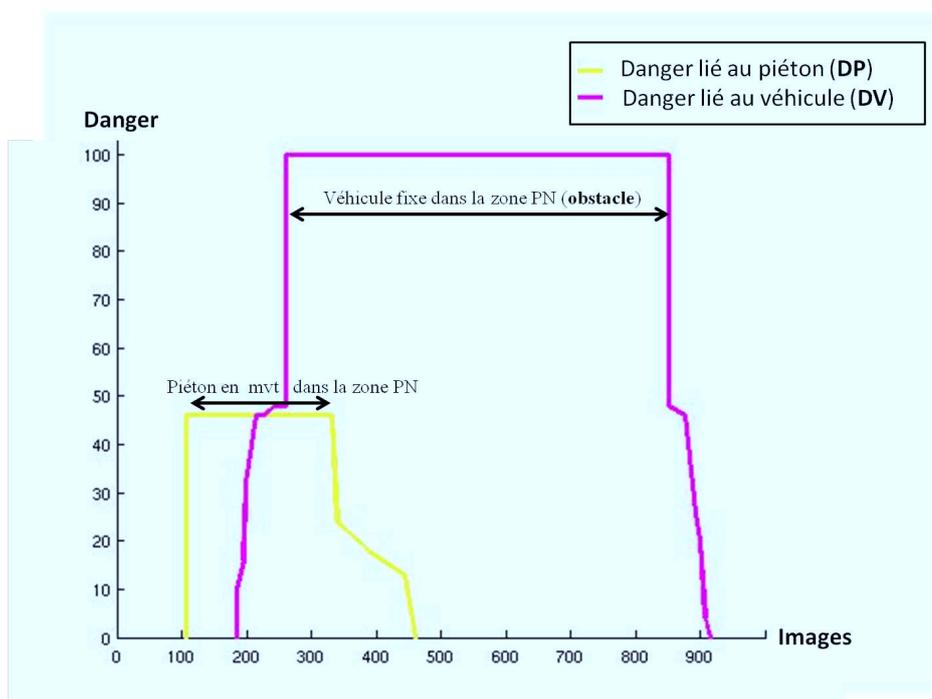


FIGURE 4.11 – Détection d'obstacles - Evolution du danger lié aux différents objets au cours du temps.

position par rapport à la zone de danger.

La figure 4.13 montre l'évolution du danger lié à chacun des objets du scénario 2 (véhicules et piéton) tout au long de la séquence d'images. Cette figure illustre aussi la détection de trois situations potentiellement dangereuses. La première, " risque d'une remontée de file ", est détectée lorsque le véhicule V1 s'arrête juste après la zone de danger. La deuxième situation, " remontée de file avec un obstacle dans la zone PN ", est détectée lorsque le véhicule V2 devient immobile derrière le véhicule V1. La troisième situation, " remontée de file avec deux obstacles dans la zone PN " est détectée lorsque le véhicule V3 s'arrête à son tour derrière le véhicule V3.

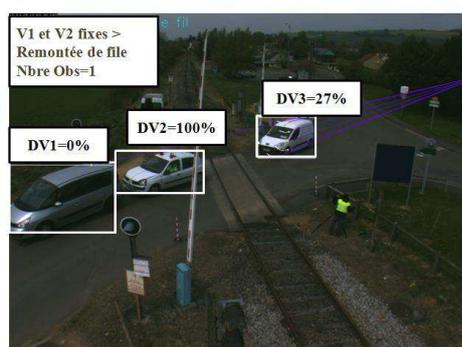
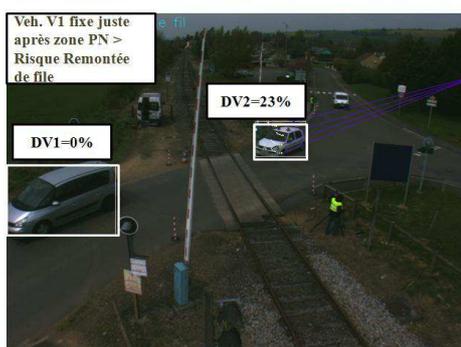
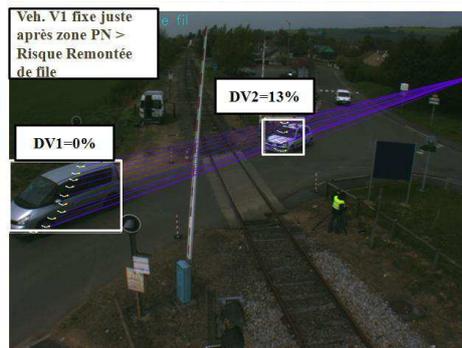
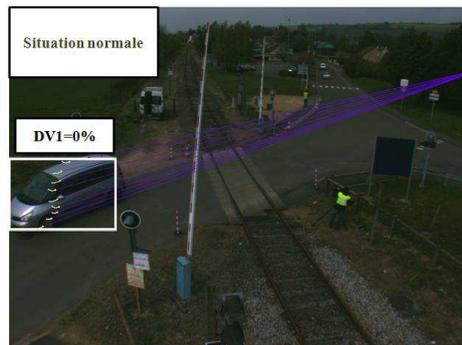
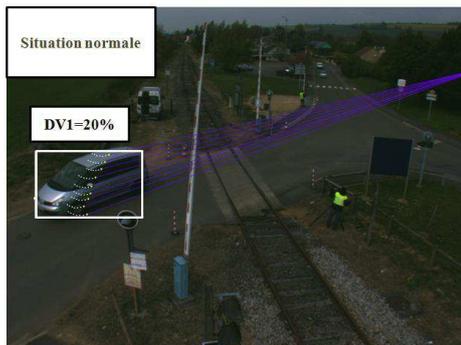
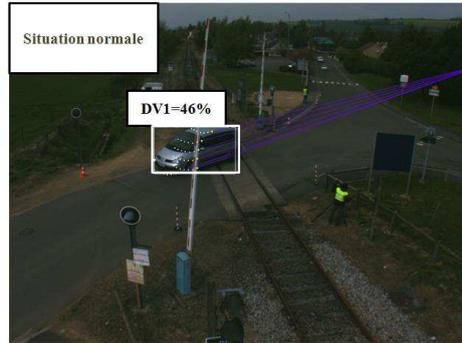
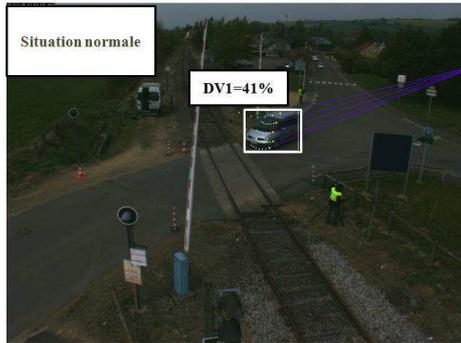
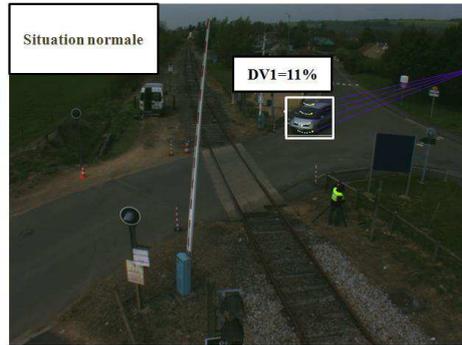
La figure 4.14 présente un autre scénario de remontée de file. Le premier véhicule traverse le PN, s'arrête un moment dans la zone de danger, redémarre, puis s'arrête juste après la barrière. Deux autres véhicules s'approchent du PN (l'un derrière l'autre), s'engagent dans la zone de danger, puis se retrouvent bloqués derrière le premier véhicule. Dans la séquence d'images traitée, les deux derniers véhicules sont détectés et suivis en un seul objet, à cause de la faible distance qui les sépare durant la manoeuvre. L'analyse de leur comportement est réduite à l'analyse du comportement de l'objet global qui les représente.

Les résultats sont identiques à ceux obtenus pour le premier scénario de remontée de file, sauf que dans le cas présent, la remontée de file est détectée avec un seul obstacle constitué des deux véhicules pour la raison évoquée ci-dessus. Lorsque le premier véhicule redémarre, la situation se débloque, et les deux véhicules constituant le même obstacle se mettent en mouvement. Leur inter-distance devient favorable à une détection séparée. Les deux véhicules sont alors suivis séparément, et l'analyse du comportement s'effectue pour chacun d'entre eux. Notons aussi sur cette séquence la détection successive de trois véhicules qui s'approchent puis s'éloignent du PN, en changeant de direction.

La figure 4.15 montre pour ce deuxième scénario de remontée de file, l'évolution du danger lié à chacun des objets (véhicules) tout au long de la séquence d'images. Cette figure illustre aussi la détection de trois situations potentiellement dangereuses. La première, " obstacle fixe dans la zone PN ", est détectée lorsque le véhicule V1 s'arrête un petit moment dans la zone de danger. La deuxième situation, " risque d'une remontée de file ", est détectée lorsque le véhicule V1 s'arrête juste après la barrière. La troisième situation, " remontée de file avec un obstacle dans la zone PN ", est détectée lorsque l'objet, composé des véhicules V2 et V3, s'arrête derrière le véhicule V1.

4.4.3 Scénario 3 : Passage en chicane, sans arrêt sur platelage

Dans le scénario de passage en chicane, illustré sur la figure 4.16, un véhicule s'approche du PN alors que les barrières sont fermées, puis le traverse en effectuant un passage en chicane. La figure 4.16 montre les résultats de détection et d'évaluation du danger correspondants sur une série d'images acquises par le système de prise de vues, installé sur le PN. Le comportement du véhicule est analysé dès son approche au PN. En effet, sa vitesse et sa trajectoire estimées anormales provoquent



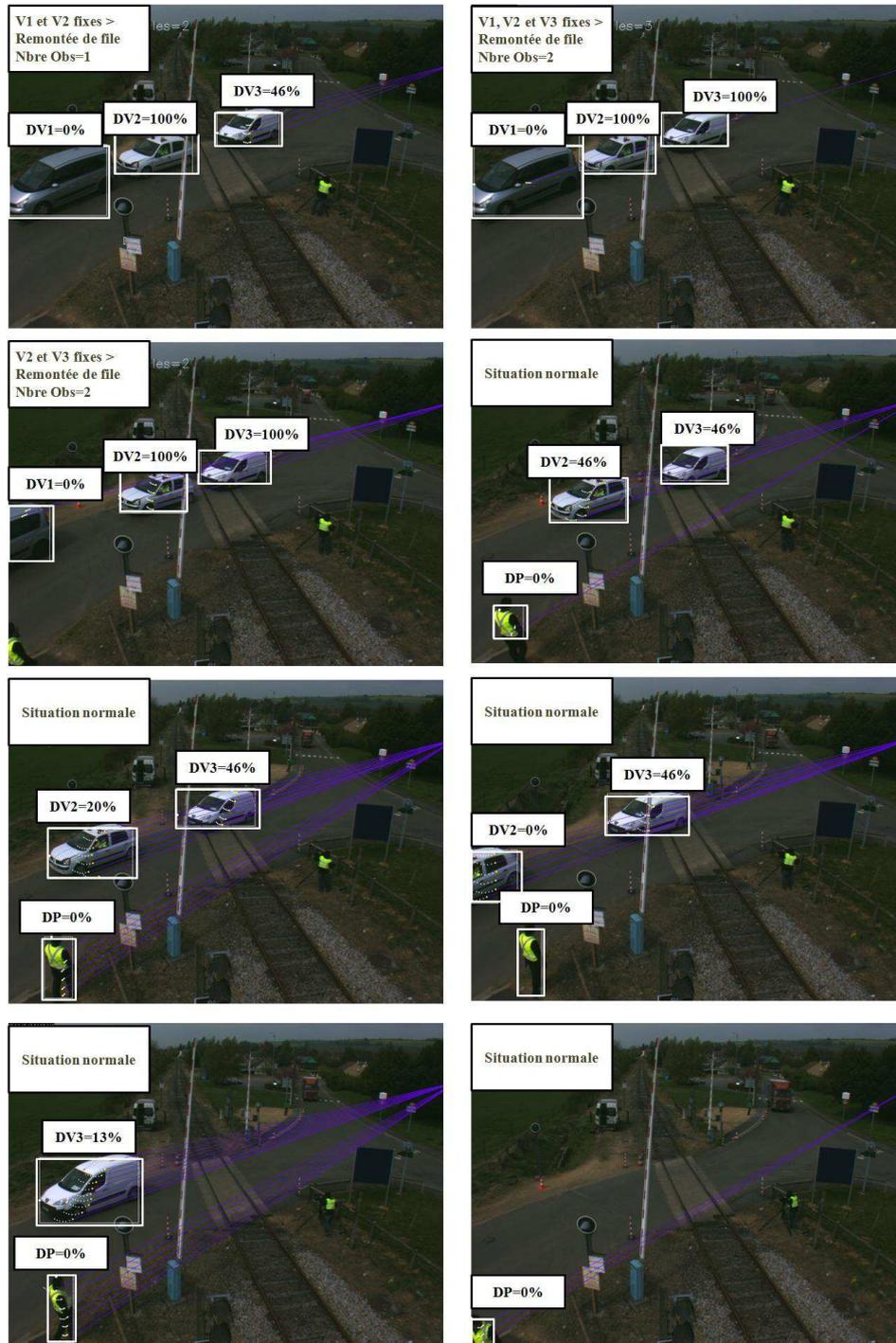


FIGURE 4.12 – Remontée de file 1 - DV_i : danger liée au véhicule V_i; DP : danger lié au piéton. Sens de lecture de la séquence : haut-bas, gauche-droite.

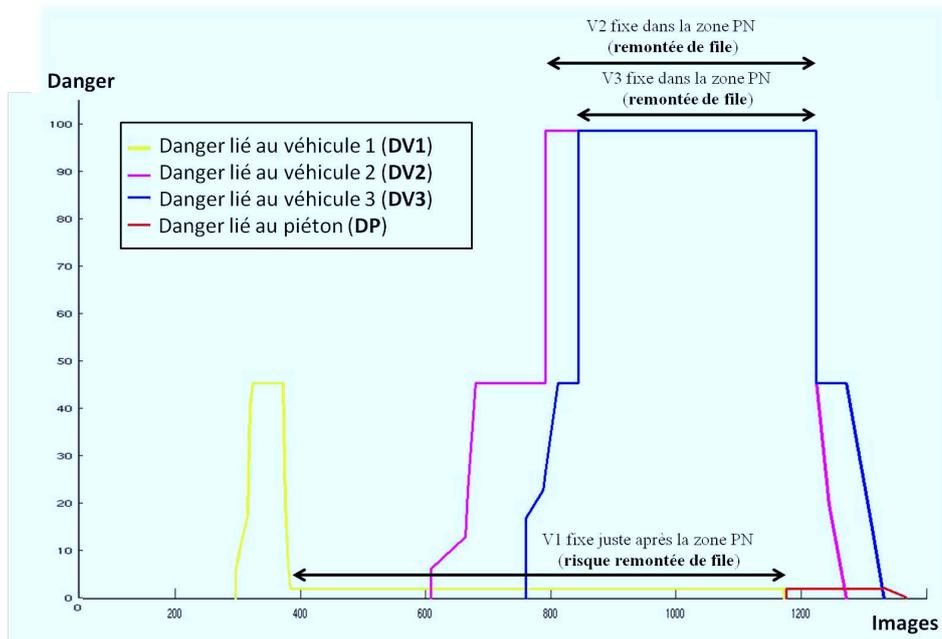


FIGURE 4.13 – Remontée de file 1 - Evolution du danger lié aux différents objets.

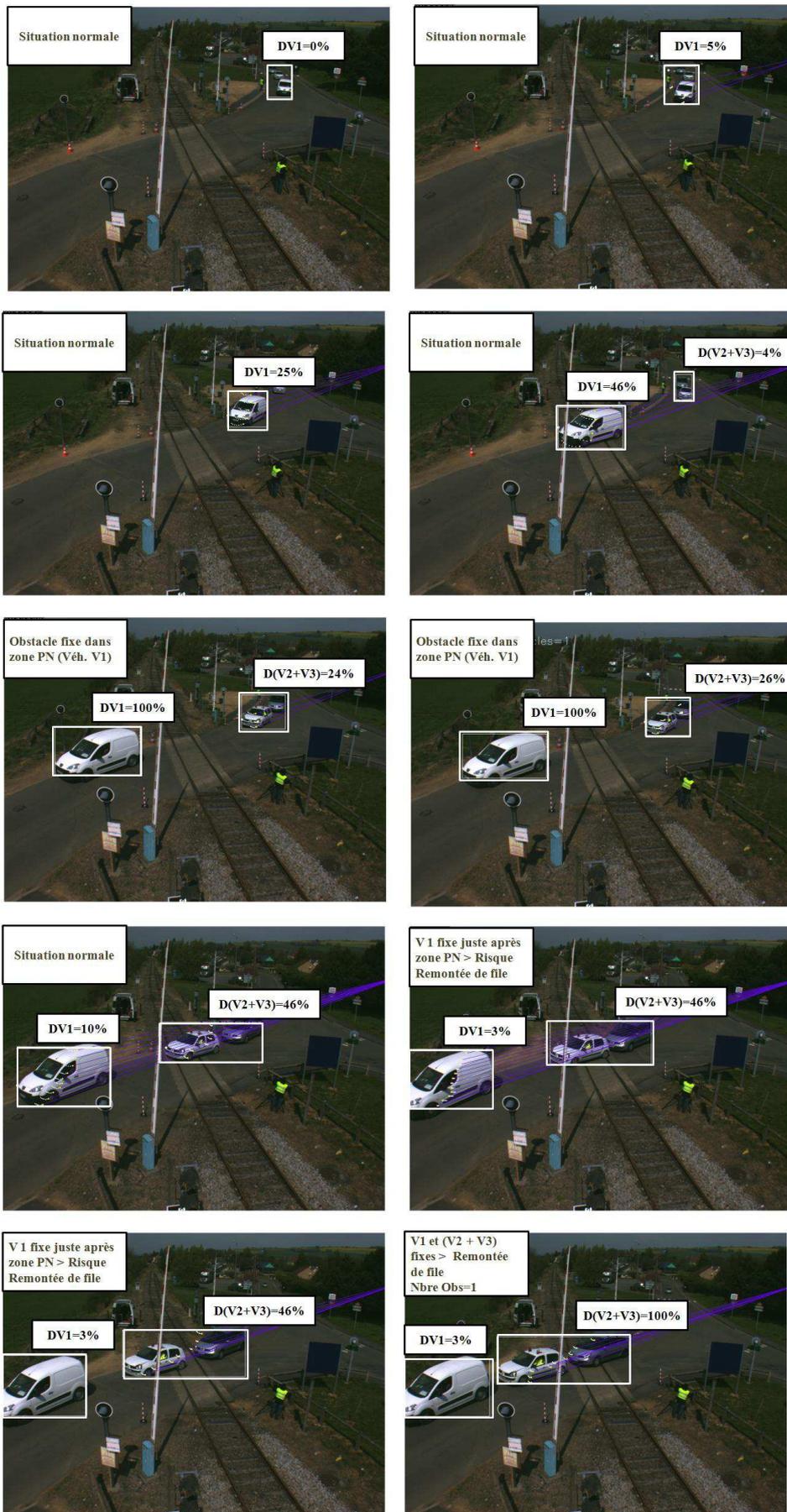
une augmentation du danger au fur et à mesure qu'il s'approche du PN. Le danger continue à augmenter lorsqu'il entre dans la zone de danger dans laquelle il atteint sa valeur de dangerosité maximale (100%). Aucune stationnarité n'est détectée dans la zone de danger, et le véhicule est considéré comme un obstacle mobile. Le danger diminue ensuite au fur et à mesure que le véhicule quitte le PN.

La figure 4.17 montre l'évolution du danger lié au véhicule tout au long de la séquence d'images.

4.4.4 Scénario 4 : Passage en chicane, avec arrêt sur platelage

Dans le scénario passage en chicane avec arrêt sur platelage, illustré sur la figure 4.18, un véhicule s'approche du PN, les barrières étant fermées. Il traverse ensuite le PN en effectuant un passage en chicane, puis s'arrête dans la zone de danger. Après un moment, le véhicule redémarre et quitte le PN.

La figure 4.18 montre les résultats de détection et d'évaluation du danger sur une série d'images acquises par le système de prise de vues, installé sur le PN. Comme dans le scénario 3, le comportement anormal du véhicule (en termes de vitesse et de trajectoire) est détecté dès son approche du PN, ce qui conduit à une augmentation du danger. Ce dernier continue à augmenter durant la traversée du PN où le véhicule est considéré comme étant un obstacle mobile. Lorsque le véhicule s'arrête, la stationnarité est détectée. Par conséquent, le véhicule devient un obstacle immobile avec un degré de danger fixé à 120% de manière arbitraire pour pouvoir distinguer le danger associé à un obstacle fixe du danger associé à



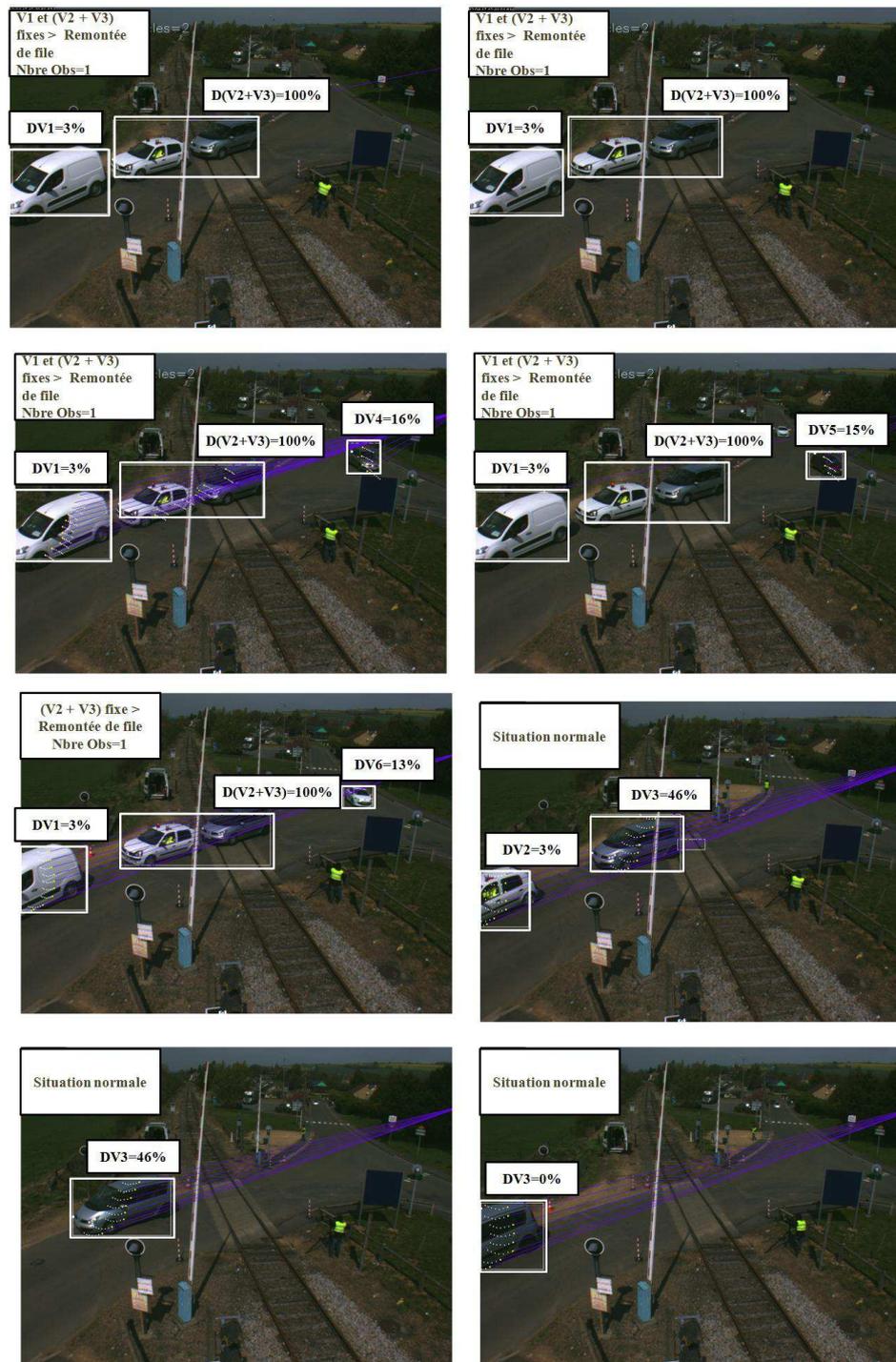


FIGURE 4.14 – Deuxième scénario de remontée de file. Sens de lecture de la séquence : haut-bas, gauche-droite.

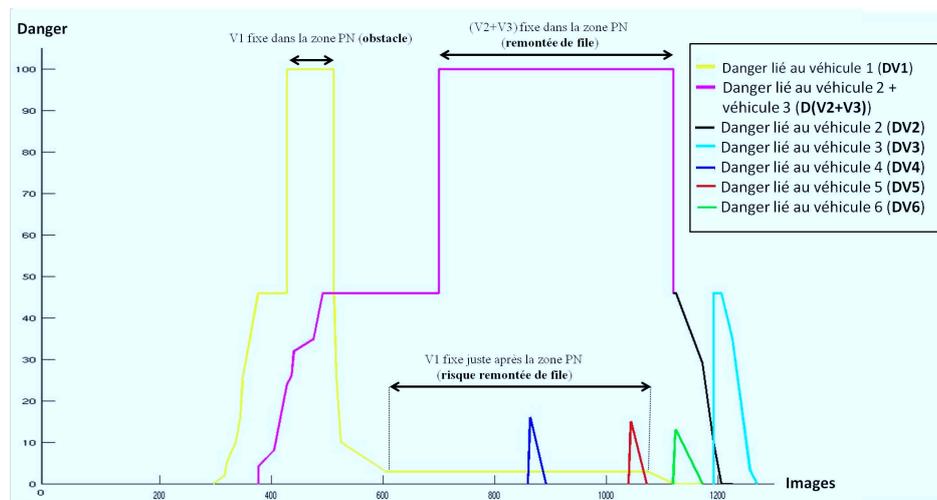


FIGURE 4.15 – Remontée de file 2 - Evolution du danger lié aux différents objets.

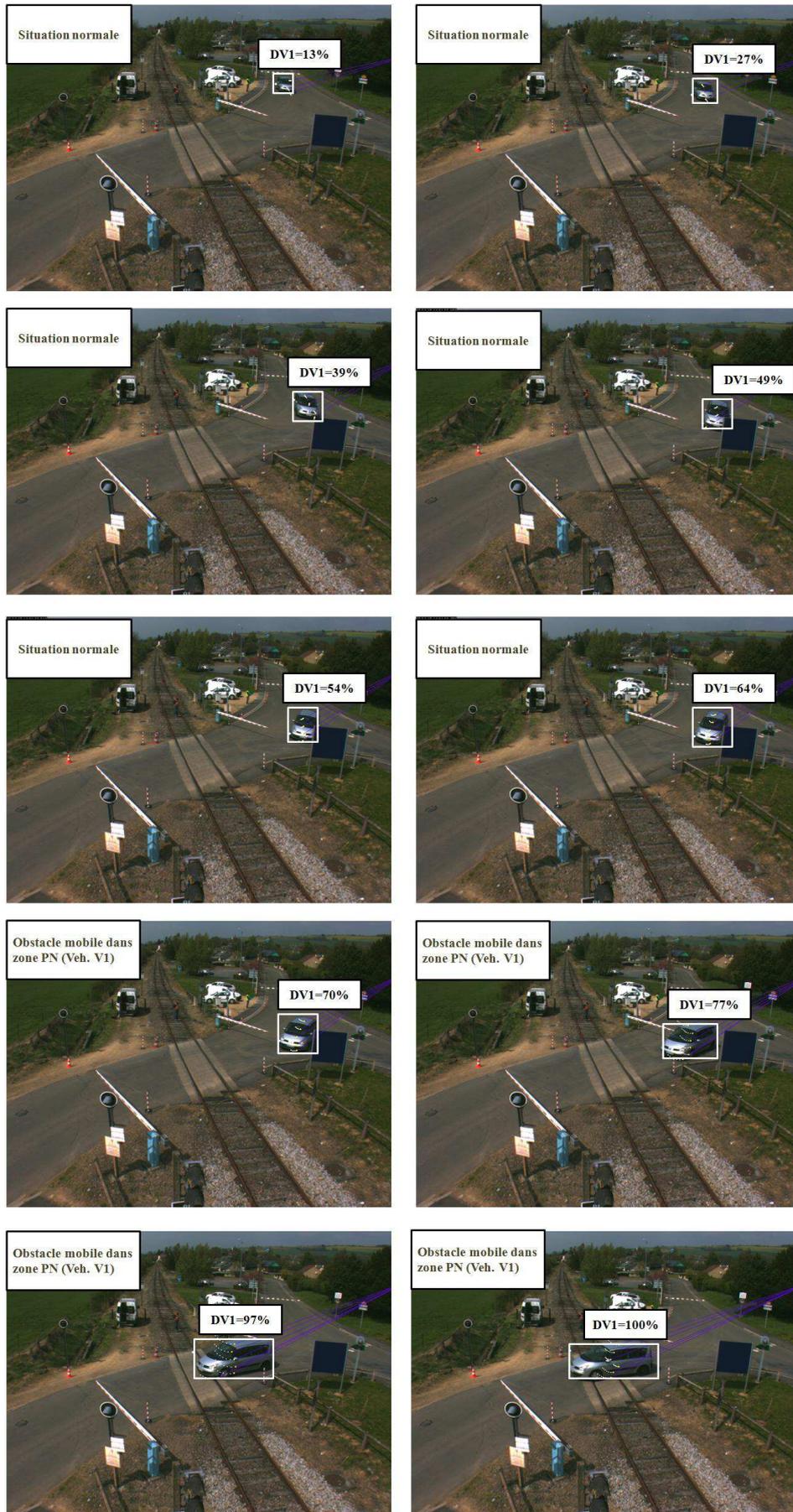
un obstacle mobile à l'intérieur de la zone du PN (ici les barrières sont fermées). Lorsque le véhicule redémarre, le danger diminue au fur et à mesure qu'il quitte le PN.

La figure 4.19 montre l'évolution du danger lié au véhicule tout au long de la séquence d'images. Cette figure illustre également la détection de trois situations potentiellement dangereuses sur le PN, les barrières étant fermées. La première, " obstacle mobile ", est détectée lorsque le véhicule en mouvement est présent dans la zone PN (danger supérieur ou égal à 70%). La deuxième situation " obstacle fixe " est détectée lorsque le véhicule s'arrête dans la zone PN (danger égal à 120%). La dernière situation potentiellement dangereuse, " obstacle mobile ", est détectée lorsque le véhicule redémarre pour quitter le PN, et ceci tant que le danger qui lui est associé est supérieur ou égal à 70%.

4.4.5 Scénario 5 : chute d'un piéton

Dans le scénario chute d'un piéton illustré sur la figure 4.20, trois piétons (P1, P2 et P3) évoluent dans l'environnement PN avec barrières fermées. Le piéton P3, en mouvement durant un bref moment au début de la séquence, devient immobile dans la zone de danger durant le reste de la séquence. Le piéton P2 s'approche du PN, puis le traverse en effectuant un passage en chicane. Durant la traversée du PN, le piéton P2 chute sur le platelage en s'immobilisant. En même temps, le piéton P1 se dirige vers le piéton P2 pour le secourir. Dans ce scénario, le degré de danger maximum estimé par la fonction de fusion de Dempster est de 100%. Le passage à un degré de 120% est choisi de manière arbitraire pour pouvoir distinguer les deux cas de stationnarité au nom du piéton.

La figure 4.20 montre les résultats de détection et d'évaluation du danger sur



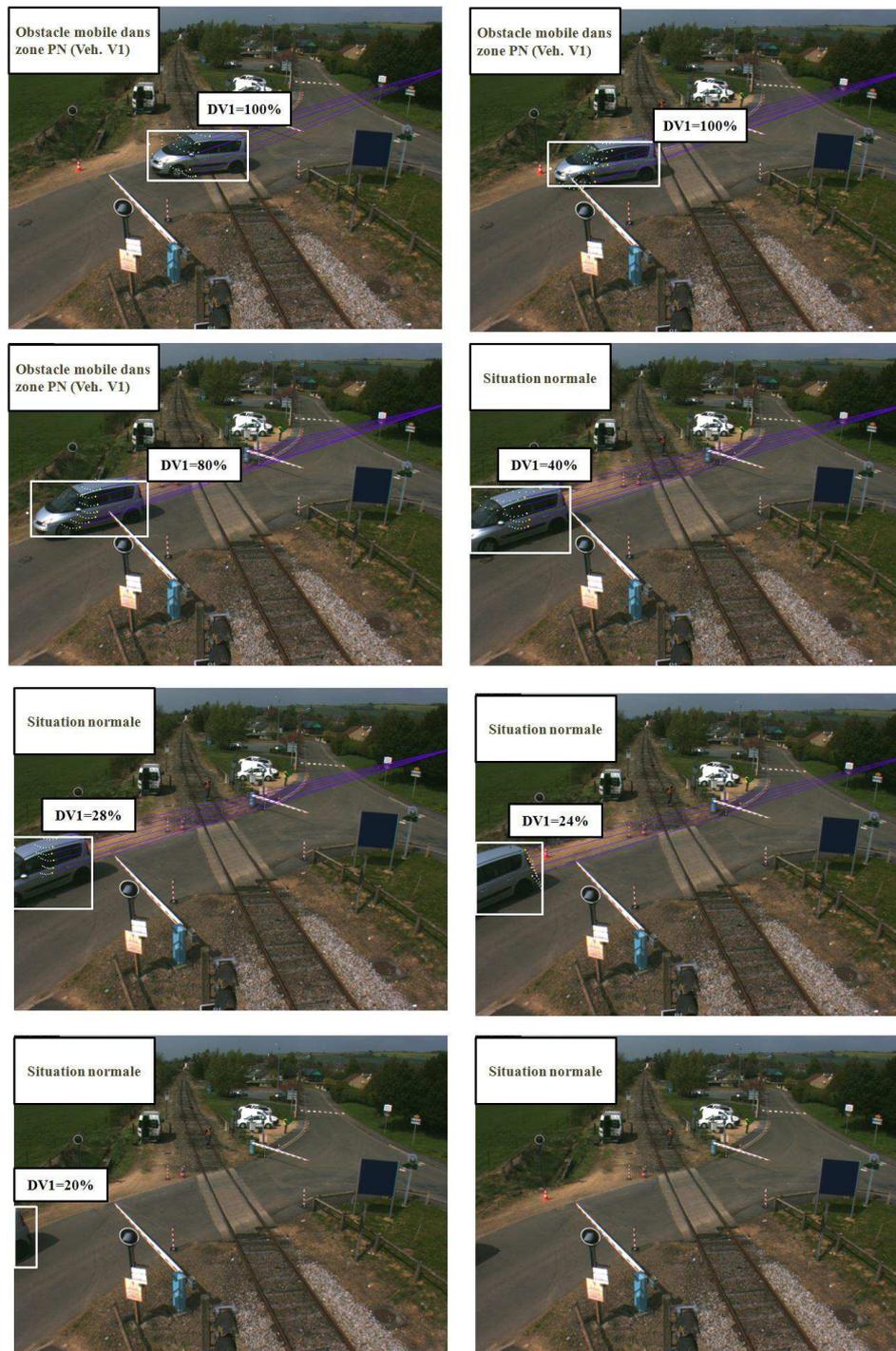


FIGURE 4.16 – Passage en chicane sans arrêt sur platelage. Sens de lecture de la séquence : haut-bas, gauche-droite.

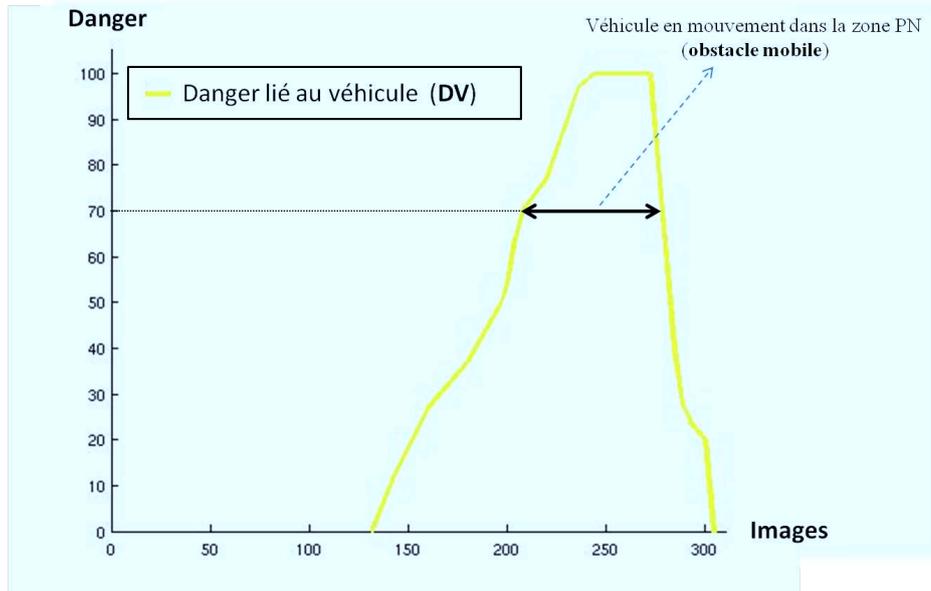


FIGURE 4.17 – Passage en chicane sans arrêt sur platelage - Evolution du danger lié au véhicule.

une série d'images acquise par le système de prise de vues, installé sur le PN. Après un bref mouvement, la stationnarité du piéton P3 est détectée avec un danger de 120% dans la zone de danger. Il est considéré alors comme un obstacle fixe. Le piéton P2 est détecté et suivi durant l'approche et la traversée du PN, le danger induit par son comportement augmente. Lorsqu'il traverse le PN, il est considéré comme étant un obstacle mobile. Lorsqu'il s'immobilise juste après sa chute sur le platelage, il est détecté comme étant un obstacle fixe avec un danger de 120%. Le piéton P1 est détecté avec un danger qui augmente au fur et à mesure qu'il s'approche du PN. Durant sa traversée du PN, il est considéré comme un obstacle mobile. Lorsqu'il arrive au niveau du piéton P2, les deux piétons deviennent très proches l'un par rapport à l'autre. Par conséquent, un seul objet stationnaire est détecté sur le platelage, avec un danger de 120%. Cet objet est considéré alors comme étant un obstacle fixe. Les deux piétons P1 et P2 se relèvent puis se séparent. Leurs détections séparées deviennent alors possibles. Le piéton P1 se dirige vers le piéton P3, et lorsqu'ils deviennent proches l'un par rapport à l'autre, un seul objet stationnaire est détecté, avec un danger de 120%. Cet objet est considéré alors comme étant un obstacle fixe. Le piéton P2 sort de la zone du PN avec un danger qui diminue au fur et à mesure qu'il s'éloigne. Par ailleurs, un véhicule est détecté avec l'estimation de son danger lors de son approche du PN. Ce véhicule s'éloigne ensuite du PN en empruntant une autre direction.



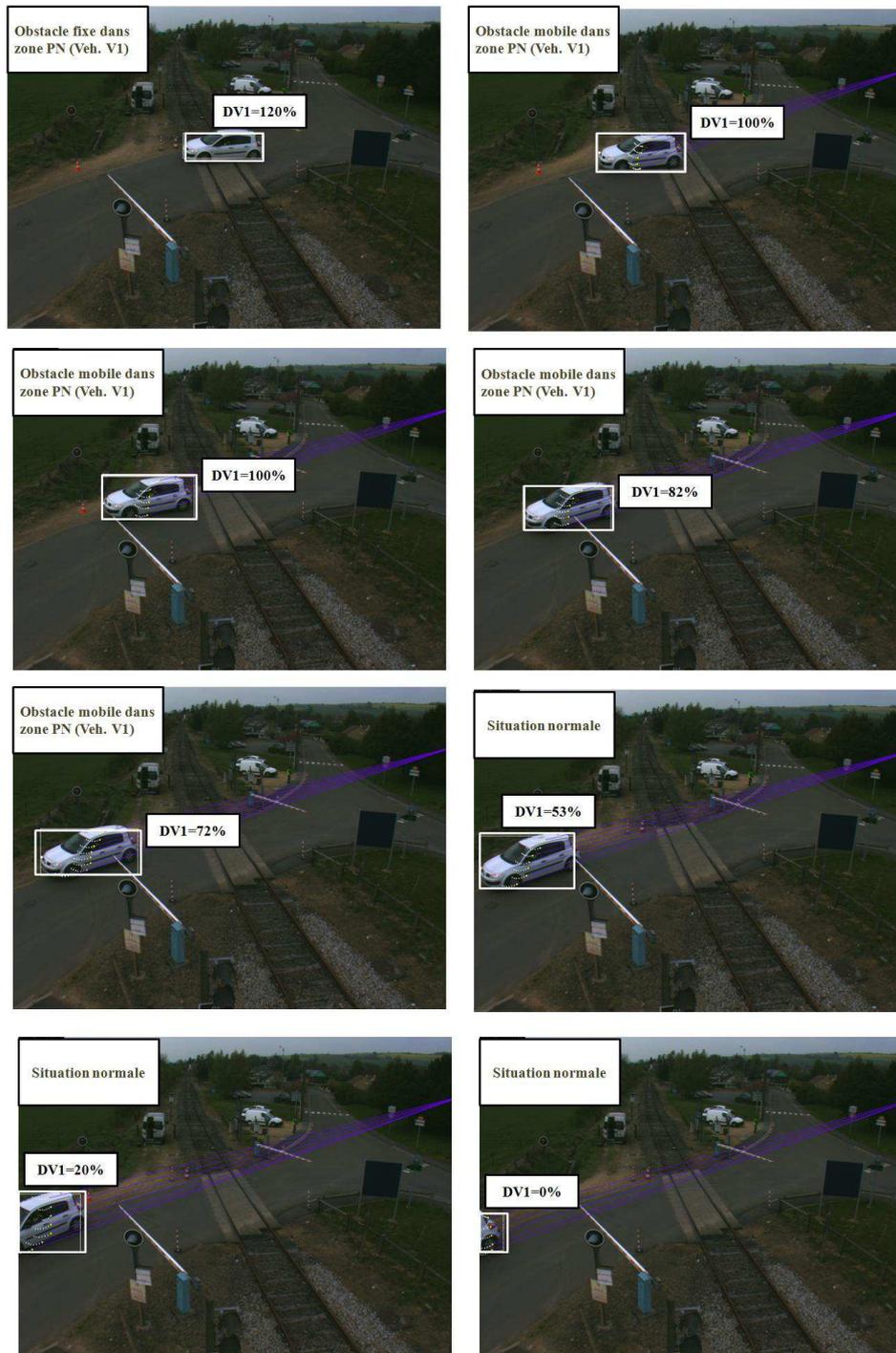


FIGURE 4.18 – Passage en chicane avec arrêt sur platelage. Sens de lecture de la séquence : haut-bas, gauche-droite.

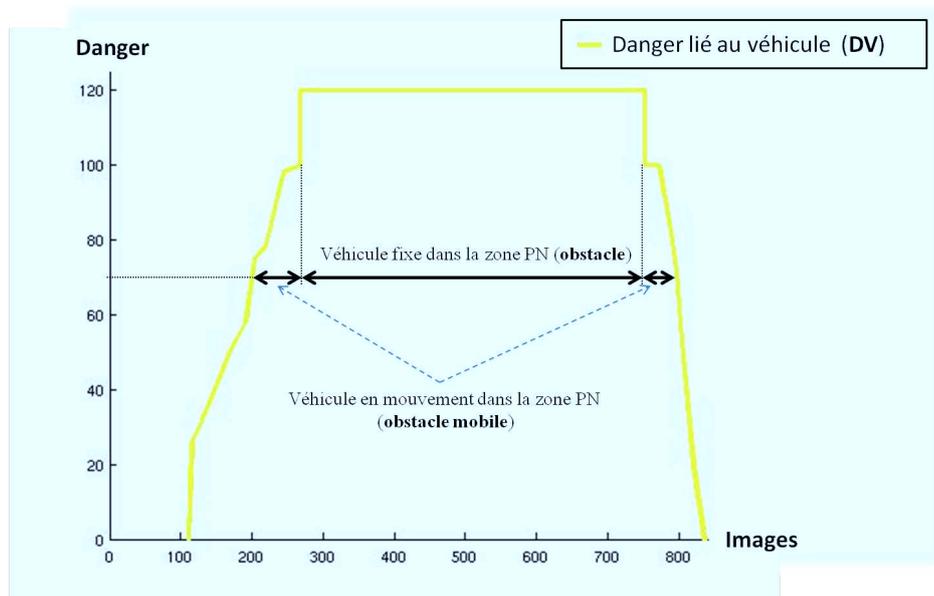
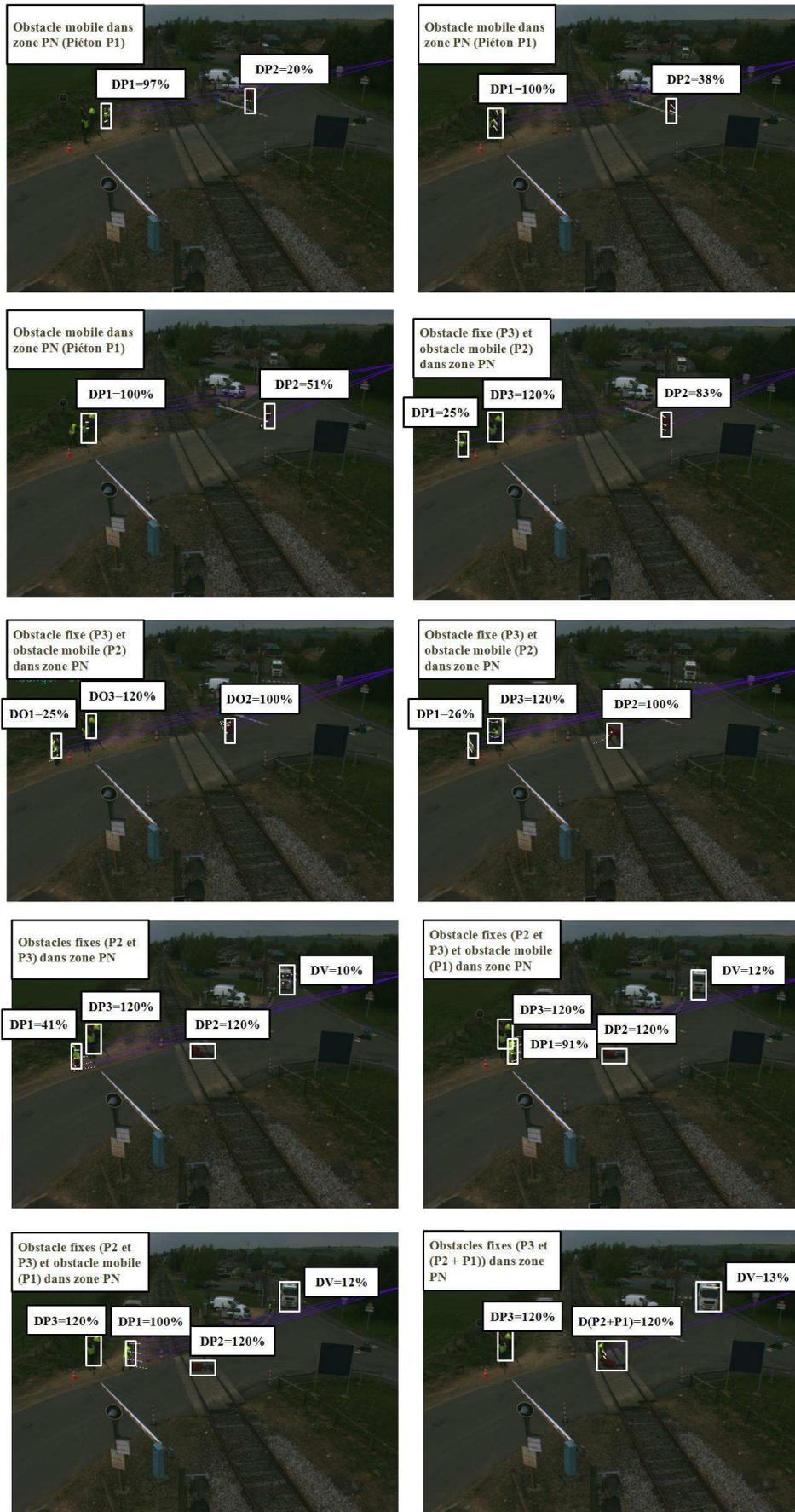


FIGURE 4.19 – Passage en chicane avec arrêt sur platelage - Evolution du danger lié au véhicule.

4.5 Conclusion

Dans ce chapitre, nous avons présenté une méthode d'évaluation de danger de scénarios d'accident autour d'un passage au niveau (PN), en utilisant une modélisation Markovienne et un modèle de calcul de risque basé sur la théorie de l'évidence. Le processus d'évaluation utilise les notions de trajectoire idéale courante et trajectoire idéale prédite. Dans un premier temps, nous commençons par estimer, pour chaque objet en mouvement et à chaque instant, sa trajectoire idéale lui permettant de revenir sur la trajectoire idéale courante qui assure une approche normale au PN. Le processus d'estimation de la trajectoire idéale est mis en oeuvre grâce à une technique basée sur une modélisation par chaîne de Markov cachée. Nous procédons ensuite à l'analyse du comportement de l'objet en mouvement en lui associant un degré de danger à chaque instant. Cette analyse est menée par modélisation et combinaison de plusieurs sources de danger (position, vitesse, orientation, accélération), en utilisant la théorie de l'évidence.

La méthode proposée a été testée sur de nombreuses séquences d'images acquises sur site (Mouzon - Ardennes), lors de campagnes d'acquisition dans le cadre du projet PANsafer. Les scénarios mis en oeuvre représentent les principales situations potentiellement dangereuses aux PN (obstacles sur platelage, passages en chicane, remontées de file). Les résultats obtenus sont satisfaisants et intéressants. En effet, à chaque instant, le système d'évaluation associe un degré de danger à chaque objet en mouvement. Le système peut également identifier la situation liée à chaque objet de la scène (obstacles fixes ou en mouvement, remontées de file, ...).



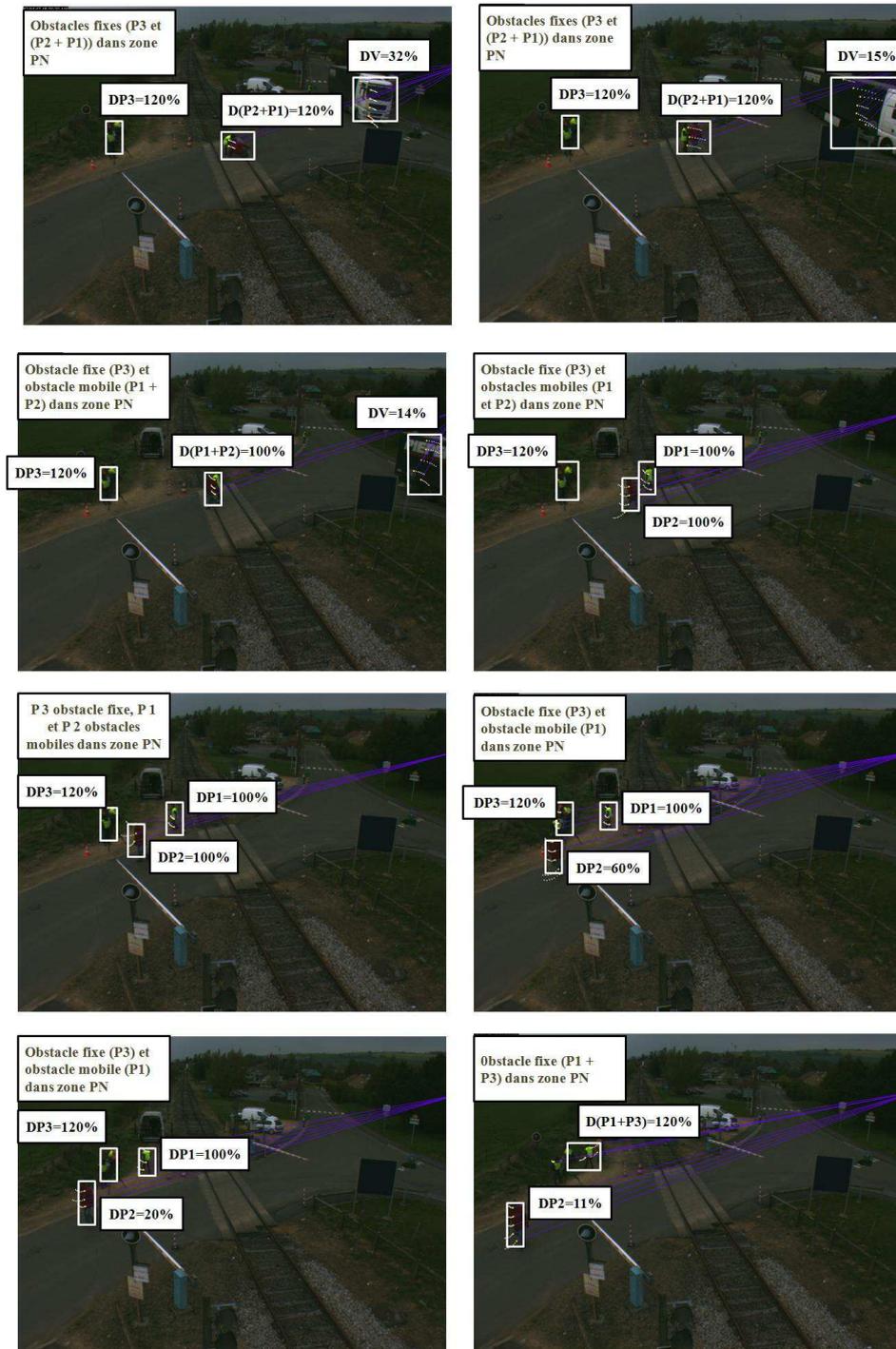


FIGURE 4.20 – Chute d'un piéton sur platelage. Sens de lecture de la séquence : haut-bas, gauche-droite.

Avec une fréquence de traitement de l'ordre de 3 à 5 images par seconde (sur un PC intel core i5 - 2,67 GHz / 3,7 GB), le système d'évaluation présente un coût calculatoire qui ne permet pas de répondre à la contrainte temps réel. Cette limitation nous a empêché de faire interagir notre système de perception et le système de communication, développé également dans le cadre du projet PANsafer, afin d'évaluer la chaîne " détection de situations / restitution de l'information aux usagers " de bout en bout. En plus de l'optimisation algorithmique et logicielle, une implémentation sur architecture spécialisée peut être envisagée pour répondre à cette contrainte.

Le système d'évaluation de danger aux PN est basé uniquement sur un traitement monoculaire. Une amélioration peut être apportée avec l'exploitation de cartes de profondeur issues d'un traitement stéréoscopique des images acquises par les caméras du système de prise de vues.

Conclusion générale

Le croisement particulier entre une ligne ferroviaire et une voie routière ou piétonne, donnant lieu à un passage à niveau (PN), possède un caractère potentiellement dangereux. Les PN constituent un thème d'étude et d'intérêt fort pour de nombreuses commissions de sécurité et équipes de recherche dans de nombreux états en Europe et au-delà. L'un des objectifs de ces différentes communautés est de développer des techniques et/ou méthodologies permettant d'améliorer la sécurité aux passages à niveau.

Les travaux de cette thèse ont été développés dans le cadre du projet ANR (PANsafer) dont le but a été la contribution active à la réduction des accidents aux passages à niveau par : 1/ L'analyse très fine des bases de données d'accidents existants survenus sur une base de passages à niveau les plus préoccupants pour en déterminer les éléments explicatifs techniques, humains, organisationnels. 2/ La mise en évidence des principaux facteurs d'accident par l'identification fonctionnelle des scénarios de ces derniers. 3/ L'analyse des comportements induits par l'infrastructure et ses modalités d'exploitation. 4/ La détection, reconnaissance et évaluation des situations dangereuses ou potentiellement dangereuses et établissement d'un niveau de criticité. 5/ L'exploration des possibilités de solutions techniques, organisationnelles . . . , en particulier, les nouvelles technologies de perception de l'environnement des interactions route/rail, de télécommunication et d'échanges d'informations. Les activités du projet PANsafer doivent directement permettre l'amélioration et le développement de la collaboration inter-modale route/rail.

Le sujet de la thèse a porté sur le développement d'un système de perception pour la détection, la reconnaissance et l'évaluation de situations potentiellement dangereuses autour de l'environnement d'un passage à niveau. Bien que l'analyse a été effectuée sur des séquences d'images monoculaires, le système d'acquisition est composé de deux caméras constituant ainsi un capteur stéréoscopique.

Au début de la thèse, l'un des premiers problèmes qui s'était posée est celui de la mise en place et l'installation sur site du système d'acquisition de données, au vu des contraintes sécuritaires, économiques et logistiques liées à la présence humaine aux abords d'un passage à niveau pour une longue durée. Nous avons donc eu besoin de définir et développer des outils permettant d'aider au bon déroulement des travaux de la thèse, notamment en termes d'acquisition de données. C'est dans ce cadre que le simulateur PANsafer, à base de réalité virtuelle, a été développé. Ce simulateur a permis de générer des séquences d'images virtuelles à l'aide de capteurs vidéo placés dans des environnements virtuels. Le placement de ces capteurs peut se faire

de manière manuelle ou automatique. Ainsi, Dans le premier chapitre, nous avons développé une méthode d'optimisation permettant de trouver la position et l'orientation optimales de caméras, étant donné l'environnement ou la zone de passage à niveau à percevoir ou à surveiller. Basée sur la géométrie descriptive, la méthode proposée procède à l'optimisation de la position et la surface de plans de projection du champ de vue des caméras, modélisé avec une représentation pyramidale.

Bien que le système de prise de vues ait été conçu avec deux caméras pour constituer un capteur stéréoscopique, l'approche développée pour le système d'évaluation du danger aux passages à niveau est basée sur une analyse monoculaire. Ce choix a été adopté afin d'éviter les problèmes liées aux procédures de mise en correspondance, notamment pour des images prises dans des conditions météorologiques difficiles (soleil, nuages, vent, changement de luminosité, ...).

Le système d'évaluation du danger aux PN est composé principalement de deux modules. Le premier concerne la détection et le suivi d'objets en mouvement. Le deuxième module est dédié à l'analyse de la trajectoire des objets afin de quantifier et qualifier le danger de situations potentiellement dangereuses aux PN.

Dans le chapitre 3, qui présente le premier module, nous avons présenté une méthode de suivi d'objets en mouvement à partir d'une séquence d'images monoculaire. Cette méthode commence tout d'abord par détecter puis séparer les objets en mouvement avant de les suivre. Le processus de suivi est basé sur le calcul du flot optique, avec une technique de propagation gaussienne du flot optique des points de Harris. Afin d'obtenir un flot optique robuste pour le plus grand nombre possible de pixels, une procédure de rectification de flot optique est mise en oeuvre à l'aide d'un modèle de filtrage de Kalman puis un algorithme d'optimisation de la différence d'intensité entre deux images successives, tout en utilisant un ensemble de contraintes de similarité et d'espace de recherche.

Dans le chapitre 4, qui présente le second module, nous avons décrit notre méthode d'évaluation du danger aux passages à niveau, en analysant le comportement des objets à travers leur suivi spatio-temporels dans les séquences d'images. La méthode proposée évalue à chaque instant le danger d'un comportement en définissant deux notions : trajectoire idéale courante, trajectoire que devrait prendre le mobile pour éviter une situation potentiellement dangereuse, et trajectoire idéale prédite, trajectoire que devrait prendre le mobile afin de rester ou revenir sur la trajectoire idéale courante. Pour prédire une trajectoire idéale du mobile à un instant donné, nous avons proposé une technique à base de modèle de Markov caché. Etant donnée une trajectoire idéale prédite à un instant donné, l'évaluation du danger lié au mobile est estimé en modélisant et combinant un ensemble de sources de danger à l'aide de la théorie de l'évidence. Le système d'évaluation de danger aux PN a été testé avec de nombreuses séquences d'images acquises sur site, en mettant en oeuvre différents scénarios de situations potentiellement dangereuses (obstacles sur platelage, passages en chicane, remontées de file . . .) lors de campagnes d'acquisition sur un PN (Mouzon - Ardennes), mis à disposition par RFF dans le cadre du projet PANsafer. Les résultats obtenus sont satisfaisants et intéressants dans la mesure où le système d'évaluation fournit, à chaque instant, un indice sur la dangerosité d'une

situation. A travers cet indice, le système peut également identifier la nature de la situation instantanée (présence d'obstacles fixes ou mobiles sur platelage, remontée de file ...) tout au long de l'évolution du mobile dans l'environnement du passage à niveau. En revanche, le système ne peut pas aujourd'hui opéré en temps réel, à cause de la quantité des traitements et la complexité des algorithmes mis en oeuvre. Cette contrainte ne nous a pas permis de tester de manière globale notre système de perception avec le système de communication, développé également dans le cadre du projet PANsafer.

Perspectives :

Les travaux de cette thèse présentent de nombreuses perspectives. Tout d'abord, sur la méthode d'optimisation pour le positionnement optimal des capteurs vidéo dans l'environnement PN à surveiller, nous proposons d'étendre cette méthode à plusieurs caméras (la méthode a été développée en considérant uniquement deux caméras). Pour des résultats plus fins, nous envisageons également d'utiliser un modèle de représentation conique pour le champ de vue des caméras, au lieu d'une représentation pyramidale.

Concernant le système d'évaluation de dangers au PN, nous proposons d'étendre notre approche avec un système de perception stéréoscopique, en exploitant l'information de profondeur qui permettrait d'améliorer les résultats des différents processus de traitement (détection et séparation d'objets en mouvement, suivi d'objets à base de flot optique 3D, évaluation de danger avec la prise en compte de mesures réelles 3D et en s'affranchissant de la perspective dans le cas monoculaire ...). La méthode d'évaluation de danger pourrait être améliorée en enrichissant le système avec des informations " type d'objets " (voiture, moto, piéton), qui permettront d'intégrer des connaissances a priori sur les objets étudiés.

Concernant toujours le système d'évaluation de dangers au PN, une contrainte qu'il faudra respecter à terme est la contrainte temps réel. Pour y parvenir, nous envisageons tout d'abord de réduire la complexité de nos algorithmes pour pouvoir accélérer la cadence de traitement des séquences d'images. Une implémentation sur architecture spécialisée pourra certainement contribuer à cet objectif. Répondre à la contrainte temps réel, permettra de tester le système global PANsafer, composé de notre système de perception de danger et le système de communication (transmission d'une information/alerte aux acteurs de PN) développé également dans le même cadre.

Aujourd'hui, le système développé est testé sur une dizaine de séquences d'images. Les résultats sont satisfaisants et intéressants. Il faudra, toutefois, élargir les tests avec une base importante de séquences d'images prises dans différentes conditions pour pouvoir effectuer une évaluation quantitative sur le taux de bonnes et mauvaises reconnaissances. Ceci permettrait de valider le système proposé pour une éventuelle industrialisation dans le cadre d'un projet FUI par exemple.

Le système de perception développé dans le cadre de cette thèse a été conçu pour une utilisation en ligne, avec un système de communication PANsafer. Ce système

pourrait être étendu pour une utilisation hors ligne, avec l'objectif d'une analyse automatique des interactions usagers/infrastructure sur des bases de données enregistrées plus importantes. Cette extension ne peut être possible que dans le cadre d'un projet partenarial du moment où de nombreuses disciplines sont mises en jeu.

Annexe

A.1 Algorithme de Levenberg-Marquardt

Entrée :- une fonction $f : R^m \rightarrow R^n$ avec $n \geq m$.

- Un vecteur de mesure $x \in R^n$.

- Une estimation des paramètres initiaux $p_0 \in R^m$.

minimiser la fonction :- $\|x - f(p)\|^2$.

Sortie :- un vecteur $p_{final} \in R^m$.

Algorithme :

$K = 0; \nu = 2; p = p_0;$

$Hessien = J^T J; \varepsilon = x - f(p); grad = J - T\varepsilon; // J$:est la matrice jacobienne.

$trouver = (\|grad\| \leq \varepsilon_1); \mu = \tau \times \max_{i=1,\dots,m}(H_{ii}); // \varepsilon_1 = 10^{-8}, \mu = 0.01.$

while (non trouver) && ($K < K_{max}$) // $K_{max} = 40$.

$K = K + 1;$

repeat

$solve(Hessien + \mu I)\delta_p = grad;$

if($\|\delta_p\| \leq \varepsilon_2 \|p\|$) // $\varepsilon_2 = 10^{-10}$

$trouver = juste;$

else

$p_{new} = p + \delta_p;$

$\rho = (\|\varepsilon\|^2 - \|x - f(p_{new})\|^2) / ((\delta_p)^T (\mu \delta_p + grad));$

if $\rho > 0$

$p = p_{new};$

$Hessien = J^T J; \varepsilon = x - f(p); grad = J - T\varepsilon;$

$trouver = (\|grad\| \leq \varepsilon_1);$

$\mu = \mu * \max(1/3, 1 - (2 * \rho - 1)^3); \nu = 2;$

else

$\mu = \mu \times \nu; \nu = 2 \times \nu$

endif

endif

until($\rho > 0$)or($trouver = juste$)

endwhile

$p_{final} = p$

Annexe

B.1 Calcul de la fonction d_c^{tf} liée à la masse de croyance $m_p^4(D)$ de la source de danger relatif à la position du centre de la région considérée (voir chapitre 4) :

La masse de croyance élémentaire $m_p^4(D)$, représentant le degré de dangerosité de la source de danger relatif à la position, dépend de la valeur de la fonction d_c^{tf} . plus d_c^{tf} prend des valeurs élevées plus masse de croyance $m_p^4(D)$ augmente. On définit d_c^{tf} de la manière suivante :

$$d_c^{tf} = W_1 + W_2 \quad (\text{B.1})$$

avec W_1 est une fonction qui dépend de :

- la position du centre de la région considérée, en terme de présence dans une zone interdite (à l'extérieur du PN) ou non.
- la distance du centre de la région par rapport à la barrière du PN.

W_2 est une fonction qui dépend de :

- la vitesse du centre de la région considérée.
- la distance du centre de la région par rapport à la barrière du PN.

Calcul de la fonction W_1 :

$$W_1 = C_o * \sqrt{D_s} \quad (\text{B.2})$$

$$C_o = D_s / (D_p + D_i) \quad D_s = D_{max} / 5 \quad (\text{B.3})$$

avec D_p est le module du vecteur $\vec{D}_p = [D_{px}, D_{py}]^T$, distance du centre de la région par rapport à la barrière du PN. D_p est inversement proportionnel à W_1 .

D_{max} est la distance maximale parcourue par un objet dans l'image.

D_i est une fonction à double valeur et inversement proportionnel à W_1 . Elle est définie de la manière suivante : si le centre de la région se trouve dans une position interdite (à l'extérieur du PN), D_i est nulle ($D_i = 0$). Sinon D_i est égale à D_{max} ($D_i = D_{max}$).

La lecture de la fonction W_1 se fait de la manière suivante : plus la position du centre de la région est proche de la barrière, plus la fonction W_1 va influencer la fonction d_c^{tf} . Si le centre de la région se trouve aussi dans une position interdite (à l'extérieur du PN), la fonction W_1 va influencer encore plus la fonction d_c^{tf} .

Calcul de la fonction W_2 :

$$W_2 = \begin{cases} 0 & \text{si } \overrightarrow{V_c^{t_f}} \cdot \overrightarrow{D_p} \leq 0 \\ \overrightarrow{V_c^{t_f}} \cdot \overrightarrow{D_p}^* & \text{si } \overrightarrow{V_c^{t_f}} \cdot \overrightarrow{D_p} > 0 \end{cases} \quad (\text{B.4})$$

avec $\overrightarrow{V_c^{t_f}} = [V_{x_c^{t_f}}, V_{y_c^{t_f}}]^T$ est le vecteur vitesse prédite du centre de la région à l'instant t_f . $\overrightarrow{D_p}^* = [1/D_{p_x}, 1/D_{p_y}]^T$ est un vecteur de distance de même sens que le vecteur $\overrightarrow{D_p} = [D_{p_x}, D_{p_y}]^T$. $\overrightarrow{V_c^{t_f}} \cdot \overrightarrow{D_p}^*$ est le produit scalaire de ces deux vecteurs. Notons que les vecteurs $\overrightarrow{V_c^{t_f}}$ et $\overrightarrow{D_p}^*$ sont directement proportionnels à W_2 .

La lecture de la fonction de W_2 se fait de la manière suivante : plus la position du centre de la région est proche de la barrière, plus la vitesse $\overrightarrow{V_c^{t_f}}$ de ce centre est grande et plus cette vitesse $\overrightarrow{V_c^{t_f}}$ est dirigée vers l'intérieur du PN, plus la valeur de W_2 va influencer la fonction $d_c^{t_f}$. Par contre, si $\overrightarrow{V_c^{t_f}}$ est dirigée vers l'extérieur du PN, W_2 est nulle et n'a aucun effet sur la fonction $d_c^{t_f}$.

Liste des publications

Sommaire

C.1 Conférences internationales	133
C.2 Conférences nationales	134
C.3 Journal	134

C.1 Conférences internationales

Houssam SALMANE, Yassine RUICHEK, and Louahdi KHOUDOUR. "***A Descriptive Geometry Based Method for Total and Common Cameras Fields of View Optimization***" In Proc. of SPIE Optical Metrology, Conference on Videometrics, Range Imaging and Applications, Munich, Germany, 2011.

Houssam SALMANE, Yassine RUICHEK, and Louahdi KHOUDOUR. "***Object Tracking Using Harris Corner Points Based Optical Flow Propagation and Kalman Filter***" In Proc. of 14th IEEE Intelligent Transportation Systems Conference (ITSC'2011), pp. 67-73, Washington D.C., USA, 2011.

Houssam SALMANE, Yassine RUICHEK, and Louahdi KHOUDOUR. "***Gaussian propagation model based dense optical flow for objects tracking***" In Proc. of International Conference on Image Analysis and Recognition (ICIAR'2012), LNCS, pp. 234-244, Aveiro, Portugal 2012.

Houssam SALMANE, Yassine RUICHEK, and Louahdi KHOUDOUR. "***Using Hidden Markov Model and Dempster-Shafer Theory for Evaluating and Detecting Dangerous Situations in Level Crossing Environments***" In Proc. of Mexican International Conference on Artificial Intelligence (MICAI'2012), LNCS, San Luis Potosi, Mexico, 2012.

Houssam SALMANE, Louahdi KHOUDOUR, and Yassine RUICHEK. "***Improving safety of level crossings by detecting hazard situations using video based processing***" In IEEE ICIRT International Conference on Intelligent Transportation., Beijing, China, 2013.

C.2 Conférences nationales

Houssam SALMANE, Yassine RUICHEK, and Louahdi KHOUDOUR. "***Suivi D'Objets En Mouvement Dans Une Séquence D'Images En Utilisant Un Filtre Basé Sur Le Flot Optique***" In Proc. of Journées des doctorants SPI-STIC de l'IFSTTAR 2011, 2011.

Houssam SALMANE, Yassine RUICHEK, and Louahdi KHOUDOUR. "***Perception De L'Environnement D'Un Passage à Niveau Par Un Système Vidéo Multi-Vues***" In Proc. of Journées des doctorants SPI-STIC de l'INRETS 2010, INRETS, 2010.

C.3 Journal

Houssam SALMANE, Yassine RUICHEK, and Louahdi KHOUDOUR. "***A novel evidence based model for detecting dangerous situations in level crossing environments***" In Proc. Expert Systems with Applications (2013), <http://dx.doi.org/10.1016/j.eswa>

Bibliographie

- [Agency 2012] European Railway Agency. *Level crossing safety in the European Union*. "www.era.europa.eu/Document-Register/Documents/Level_crossing_safety_EU_2012.pdf", 2012. (Cité en pages ix, 1, 6 et 7.)
- [Aubert 2003] Jean Aubert. Cours de dessin d'architecture : à partir de la géométrie descriptive : à l'usage des élèves de première année des écoles d'architecture. Éd. de la Villette, 2003. (Cité en pages 18 et 28.)
- [Barron 1994] J.L. Barron, D.J. Fleet et S.S. Beauchemin. *Performance of optical flow techniques*. International Journal of Computer Vision, vol. 12, no. 1, pages 43–77, 1994. (Cité en page 80.)
- [Bashir 2005] F. Bashir, A. Khokhar et D. Schonfeld. *Automatic Object Trajectory-Based Motion Recognition Using Gaussian Mixture Models*. In Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on, pages 1532–1535, 2005. (Cité en page 86.)
- [Bashir 2007] F.I. Bashir, A.A. Khokhar et D. Schonfeld. *Object Trajectory-Based Activity Classification and Recognition Using Hidden Markov Models*. Image Processing, IEEE Transactions on, vol. 16, no. 7, pages 1912–1919, 2007. (Cité en page 86.)
- [Baum 1968] L. E. Baum et G. R. Sell. *Growth Functions for Transformations on Manifolds*. Pac. J. Math., vol. 27, pages 211 – 227, 1968. (Cité en page 90.)
- [Baum 1970] Leonard E Baum, Ted Petrie, George Soules et Norman Weiss. *A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains*. The annals of mathematical statistics, vol. 41, no. 1, pages 164–171, 1970. (Cité en page 90.)
- [Bea 2012] *Rapport d'enquête technique sur la collision entre un TER et un ensemble routier*. "www.bea-tt.equipement.gouv.fr/IMG/pdf/Rapport_BEATT_2011_011.pdf", mai 2012. (Cité en page 1.)
- [Blake 1998] Andrew Blake et Michael Isard. *Active contours*, 1998. (Cité en page 52.)
- [Bouguet 2002] Jean-Yves Bouguet. *Pyramidal implementation of the affine Lucas Kanade feature tracker description of the algorithm*. Intel Corporation, 2002. (Cité en pages 53, 56 et 59.)
- [Boukour 2012] F. Boukour, J.L. Bruyelle, J. Dereux, A. Flancquart, J.P. Ghys, M. Heddebautl, L. Khoudour, J.F. Mendès, Y. Ruichek, C. Meurie et H. Salmane. *Rapport détaillé sur la réalisation et l'évaluation en laboratoire du système de perception de l'environnement PN et de communication avec les usagers route et fer PANsafer*. In Livrable D4.2, 2012. (Cité en pages 19 et 21.)

- [Boykov 2001] Y. Boykov, O. Veksler et R. Zabih. *Fast approximate energy minimization via graph cuts*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 23, no. 11, pages 1222–1239, 2001. (Cité en page 52.)
- [Boykov 2004] Y. Boykov et V. Kolmogorov. *An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 26, no. 9, pages 1124–1137, 2004. (Cité en page 52.)
- [Burke 1996] Michael W Burke. Image acquisition, volume 1. Chapman & Hall New York, 1996. (Cité en page 24.)
- [BUSSEREAU 2009] Dominique BUSSEREAU. *Bilan du plan BUSSEREAU en faveur de la sécurité des passages à niveau*. DOSSIER DE PRESSE.<http://www.developpement-durable.gouv.fr/>, 2009. (Cité en page 8.)
- [Chen 2005] Trista P Chen, Horst Haussecker, Alexander Bovyrin, Roman Belenov, Konstantin Rodyushkin, Alexander Kuranov et Victor Eruhimov. *Computer vision workload analysis : Case study of video surveillance systems*. Intel Technology Journal, vol. 9, no. 2, pages 109–118, 2005. (Cité en page 17.)
- [Cheng 1995] Yizong Cheng. *Mean shift, mode seeking, and clustering*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 17, no. 8, pages 790–799, 1995. (Cité en page 56.)
- [Ciblac 2004] Thierry Ciblac, Pierre Macé et Louis-Paul Untersteller. *Recherche de l'Intersection de Polyèdres dans un seul Dessin en Perspective*. Proceedings of Journées AFIG, 2004. (Cité en page 26.)
- [Cousin 1981] Marie-Elisabeth Cousin. *Les conventions internationales conclues par la France et publiées au Journal Officiel de la République française en 1981*. Annuaire français de droit international, vol. 27, no. 1, pages 781–799, 1981. (Cité en page 1.)
- [Derpanis 2004] Konstantinos G Derpanis. *The Harris corner detector*. York University, 2004. (Cité en page 59.)
- [Dubois-Arber 2006] F Dubois-Arber, A Jeannin et G Meystre-Agustoni. *Un système de surveillance de deuxième génération pour améliorer la surveillance du VIH/sida en Suisse*. Bulletin OFSP, vol. 15, pages 277–281, 2006. (Cité en page 47.)
- [ERADIS] ERADIS. *ERADIS-European Railway Agency Database of Interoperability and Safety*. "http://pdb.era.europa.eu/safety_docs/csi/default.aspx". (Cité en page 7.)
- [Fakhfakh 2011a] N Fakhfakh. *Détection et localisation tridimensionnelle par stéréovision d'objets en mouvement dans des environnements complexes : application aux passages à niveau*. PhD thesis, Ecole Centrale de Lille, 2011. (Cité en page 51.)
- [Fakhfakh 2011b] N Fakhfakh, L Khoudour, EM El-Kourssi, J Jacot et A Dufaux. *A video-based object detection system for improving safety at level crossings*. Open Transportation Journal, vol. 5, pages 45–59, 2011. (Cité en page 18.)

- [Foresti 1998] G.L. Foresti. *A real-time system for video surveillance of unattended outdoor environments*. Circuits and Systems for Video Technology, IEEE Transactions on, vol. 8, no. 6, pages 697–704, 1998. (Cité en pages 18 et 48.)
- [Francois 2004] Alexandre R Francois. *Real-time multi-resolution blob tracking*. Rapport technique, University of Southern California, Los Angeles, April 2004. (Cité en page 52.)
- [Fridberger 2004] Anders Fridberger, Jerker Widengren et Jacques Boutet De Monvel. *Measuring hearing organ vibration patterns with confocal microscopy and optical flow*. Biophysical journal, vol. 86, no. 1, page 535, 2004. (Cité en page 80.)
- [Fuse 2000] Takashi Fuse, Eihan Shimizu et Morito Tsutsumi. *A comparative study on gradient-based approaches for optical flow estimation*. International Archives of Photogrammetry and Remote Sensing, vol. 33, no. B5/1 ; PART 5, pages 269–276, 2000. (Cité en page 80.)
- [Gevers 2003] Th. Gevers et H.M.G. Stokman. *Robust Photometric Invariant Region Detection in Multispectral Images*. International Journal of Computer Vision, vol. 53, no. 2, pages 135–151, 2003. (Cité en page 52.)
- [Gordon 1999] G. Gordon, T. Darrell, M. Harville et J. Woodfill. *Background estimation and removal based on range and color*. In Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., pages 464–464 Vol. 2, 1999. (Cité en page 50.)
- [Got 2012] Claude Got. *Les accidents mortels de passages à niveau*. "www.securite-routiere.org/docacrobat/passagesaniveauGot2012.pdf", September 2012. (Cité en page 3.)
- [Greenhill 2004] Stewart Greenhill, Svetha Venkatesh et Geoff West. *Adaptive Model for Foreground Extraction in Adverse Lighting Conditions*. In Chengqi Zhang, Hans W. Guesgen et Wai-Kiang Yeap, editeurs, PRICAI 2004 : Trends in Artificial Intelligence, volume 3157 of *Lecture Notes in Computer Science*, pages 805–811. Springer Berlin Heidelberg, 2004. (Cité en page 49.)
- [Grünbaum 2003] Branko Grünbaum. *Convex polytopes, volume 221 of Graduate Texts in Mathematics*, 2003. (Cité en page 28.)
- [Harris 1988] Chris Harris et Mike Stephens. *A combined corner and edge detector*. In Alvey vision conference, volume 15, page 50. Manchester, UK, 1988. (Cité en page 59.)
- [Harville 2001] M. Harville, G. Gordon et J. Woodfill. *Foreground segmentation using adaptive mixture models in color and depth*. In Detection and Recognition of Events in Video, 2001. Proceedings. IEEE Workshop on, pages 3–11, 2001. (Cité en page 50.)
- [Herault 1986] Jeanny Herault et Christian Jutten. *Space or time adaptive signal processing by neural network models*. In AIP conference proceedings, volume 151, page 206, 1986. (Cité en page 51.)

- [Ilyas 2010] Atif Ilyas, Mihaela Scuturici et Serge Miguet. *A combined motion and appearance model for human tracking in multiple cameras environment*. In Emerging Technologies (ICET), 2010 6th International Conference on, pages 198–203. IEEE, 2010. (Cité en page 18.)
- [Jain 2009] Anil K. Jain et Stan Z. Li. *Encyclopedia of biometrics*,. New York, NY Springer, 2009. (Cité en page 24.)
- [Kettner 1999] V. Kettner et R. Zabih. *Bayesian multi-camera surveillance*. In Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., volume 2, pages –259, 1999. (Cité en page 17.)
- [Khoudour 2010] L. Khoudour, P. Feltz, D. Bertrand, M. Ghazel, M. Heddebaut, S. Collart-Dutilleul, Y. Ruichek et A. Flancquart. *PANsafer Project : Towards a safer level crossing*. In 11th Level Crossing Symposium, Tokyo, 2010. (Cité en page 10.)
- [Kim 2005] Kyungnam Kim, Thanarat H Chalidabhongse, David Harwood et Larry Davis. *Real-time foreground-background segmentation using codebook model*. Real-time imaging, vol. 11, no. 3, pages 172–185, 2005. (Cité en page 49.)
- [Lamotte 2010] Olivier Lamotte, Cyril Meurie et Yassine Ruichek. *A 3D virtual universe based simulation platform : Application to intelligent video surveillance of level crossings*. In 11th Level Crossing Symposium, Tokyo, Tokyo, Japanese, 2010. (Cité en page 19.)
- [Leibe 2008] Bastian Leibe, Konrad Schindler, Nico Cornelis et Luc Van Gool. *Coupled object detection and tracking from static cameras and moving vehicles*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 30, no. 10, pages 1683–1698, 2008. (Cité en page 49.)
- [Li 2010] Xiaohu Li, Taiyi Zhang, Xiaodong Shen et Jiancheng Sun. *Object tracking using an adaptive Kalman filter combined with mean shift*. Optical Engineering, vol. 49, no. 2, pages 020503–020503, 2010. (Cité en page 52.)
- [Lin 2004] Lin Lin, Yaakov Bar-Shalom et Thiagalingam Kirubarajan. *Data association combined with the probability hypothesis density filter for multitarget tracking*. In Defense and Security, pages 464–475. International Society for Optics and Photonics, August 2004. (Cité en page 52.)
- [Lourakis 2005] Manolis IA Lourakis. *A brief description of the Levenberg-Marquardt algorithm implemented by levmar*. Institute of Computer Science, Foundation for Research and Technology, vol. 11, 2005. (Cité en page 33.)
- [MacCormick 1999] John MacCormick et Andrew Blake. *A probabilistic exclusion principle for tracking multiple objects*. In Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference, volume 1, pages 572–578. IEEE, 1999. (Cité en page 52.)
- [Min 2011] *Bilan du plan de sécurisation des passages à niveau*. Ministère de l'écologie, du développement durable, des transports et du logement, juin 2011. (Cité en page 1.)

- [Monroe 2003] David A Monroe. *Ground based security surveillance system for aircraft and other commercial vehicles*, Avril 8 2003. US Patent 6,545,601. (Cité en page 47.)
- [Nan 2011] Dongji Nan, Soichiro Morishita et Hiroshi Yokoi. *A tracking method of multiple moving objects for animal watch sensor systems-Measures for random mobility and the problem with connection and separation*. In SICE Annual Conference (SICE), 2011 Proceedings of, pages 1991–1994. IEEE, 2011. (Cité en page 49.)
- [Natarajan 2007] P. Natarajan et R. Nevatia. *Coupled Hidden Semi Markov Models for Activity Recognition*. In Motion and Video Computing, 2007. WMVC '07. IEEE Workshop on, pages 10–10, 2007. (Cité en page 87.)
- [Nguyen 2005] N.T. Nguyen, D.Q. Phung, S. Venkatesh et H. Bui. *Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model*. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 2, pages 955–960 vol. 2, 2005. (Cité en page 87.)
- [Ohta 2005] Masaru Ohta. *Level Crossings Obstacle Detection System Using Stereo Cameras*. Quarterly Report of RTRI, vol. 46, no. 2, pages 110–117, 2005. (Cité en page 18.)
- [Oja 2000] E Oja et A Hyvarinen. *Independent component analysis : algorithms and applications*. Neural Networks, vol. 13, no. 4-5, pages 411–430, 2000. (Cité en page 51.)
- [Panahi 2008] Sorayya Panahi, Samira Sheikhi, Shahrzad Hadadan et Niloofar Gheissari. *Evaluation of background subtraction methods*. In Computing : Techniques and Applications, 2008. DICTA'08. Digital Image, pages 357–364. IEEE, 2008. (Cité en page 51.)
- [Peng 2009] Yan Peng, Liu Xue-bin, Hu Bing-liang et Shao Xiu-juan. *Application of Multispectral Image in the Detection of Small Target*. In Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on, pages 1–4, 2009. (Cité en page 52.)
- [Rabiner 1989] L. Rabiner. *A tutorial on hidden Markov models and selected applications in speech recognition*. Proceedings of the IEEE, vol. 77, no. 2, pages 257–286, 1989. (Cité en page 90.)
- [Rittscher 2000] J. Rittscher, J. Kato, S. Joga et A. Blake. *A Probabilistic Background Model for Tracking*. In David Vernon, editeur, Computer Vision - ECCV 2000, volume 1843 of *Lecture Notes in Computer Science*, pages 336–350. Springer Berlin Heidelberg, 2000. (Cité en page 50.)
- [Sacchi 2000] Claudio Sacchi et Carlo S Regazzoni. *A distributed surveillance system for detection of abandoned objects in unmanned railway environments*. Vehicular Technology, IEEE Transactions on, vol. 49, no. 5, pages 2013–2026, 2000. (Cité en page 48.)

- [Safety] Road Safety. *International Level Crossing Awareness Day (ILCAD)*. European Commission "www.ilcad.org/ILCAD-2013.html". (Cit  en pages 7 et 8.)
- [Schmid 2000] Cordelia Schmid, Roger Mohr et Christian Bauckhage. *Evaluation of interest point detectors*. International Journal of computer vision, vol. 37, no. 2, pages 151–172, 2000. (Cit  en page 59.)
- [Schnieder 2008] E. Schnieder, R. Slovak, E. M. El Koursi, L. Tordai et M. Woods. *A European Contribution to Level Crossing Safety*. In Proceedings of FOVUS, Stuttgart, Germany, September 2008. (Cit  en page 10.)
- [SET 1996] *Situation et enjeux de s curit  aux passages   niveau*. SETRA - CSTR / SNCF, D cembre 1996. (Cit  en page 3.)
- [Shafer 1976] Glenn Shafer. *A mathematical theory of evidence*, volume 1. Princeton university press Princeton, 1976. (Cit  en page 95.)
- [Shi 1994] Jianbo Shi et Carlo Tomasi. *Good features to track*. In Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on, pages 593–600. IEEE, 1994. (Cit  en page 58.)
- [Shigeru 1988] Haga Shigeru. *Prevention of Accidents at Road-Rail Level Crossings Protected with Automatic Barriers*. In Proceedings of the Human Factors Society 32nd Annual Meeting, volume 2 of *Safety : Transportation Safety*, pages 933–937, 1988. (Cit  en page 4.)
- [Silmon 2010] Joe Silmon et Clive Roberts. *Using functional analysis to determine the requirements for changes to critical systems : Railway level crossing case study*. Rel. Eng. & Sys. Safety, vol. 95, no. 3, pages 216–225, 2010. (Cit  en page 2.)
- [Singh 2010] Sanjay Singh, Srinivasa Murali Dunga, AS Mandal, Chandra Shekhar et Anil Vohra. *Moving Object Tracking Using Object Segmentation*. In VinuV Das et R. Vijaykumar,  diteurs, Information and Communication Technologies, volume 101 of *Communications in Computer and Information Science*, pages 691–694. Springer Berlin Heidelberg, 2010. (Cit  en page 51.)
- [Slovak 2006] R. Slovak, E. M. El Koursi, L. Tordai, M. Woods et E. Schnieder. *a New European Project for Safer Level Crossings*. In 9th International Level Crossing Safety and Trespass Prevention Symposium, Montr al, Canada, September 2006. (Cit  en page 10.)
- [Smets 1990] Philippe Smets. *The combination of evidence in the transferable belief model*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 12, no. 5, pages 447–458, 1990. (Cit  en page 95.)
- [Smith 2004] Kevin Smith et Daniel Gatica-Perez. *Order matters : a distributed sampling method for multi-object tracking*. In British Machine Vision Conference, pages 1–10, 2004. (Cit  en page 52.)
- [Stauffer 1999] Chris Stauffer et W.E.L. Grimson. *Adaptive background mixture models for real-time tracking*. In Computer Vision and Pattern Recognition,

1999. IEEE Computer Society Conference on., pages –252 Vol. 2, 1999. (Cité en page 50.)
- [Stauffer 2000] Chris Stauffer et W.E.L. Grimson. *Learning patterns of activity using real-time tracking*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 22, no. 8, pages 747–757, 2000. (Cité en page 50.)
- [Stenger 2001] B. Stenger, V. Ramesh, N. Paragios, F. Coetzee et J.M. Buhmann. *Topology free hidden Markov models : application to background modeling*. In Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on, pages 294–301 vol.1, 2001. (Cité en page 51.)
- [Suzuki 2003] Kenji Suzuki, Isao Horiba et Noboru Sugie. *Linear-time connected-component labeling based on sequential local operations*. Computer Vision and Image Understanding, vol. 89, no. 1, pages 1–23, 2003. (Cité en page 52.)
- [Tamgade 2009] Sukeshtni N Tamgade et Vibha R Bora. *Motion vector estimation of video image by Pyramidal implementation of Lucas Kanade Optical flow*. In Emerging Trends in Engineering and Technology (ICETET), 2009 2nd International Conference on, pages 914–917. IEEE, 2009. (Cité en page 59.)
- [Tey 2009] Li-Sian Tey, Luis Ferreira et Hussein Dia. *Evaluating cost-effective railway level crossing protection systems*. In 32nd Australasian Transport Research Forum, 2009. (Cité en page 2.)
- [Tey 2012] Li-Sian Tey, Inhi Kim et Luis Ferreira. *Evaluating safety at railway level crossings with microsimulation modeling*. Transportation Research Record : Journal of the Transportation Research Board, vol. 2298, no. 1, pages 70–77, 2012. (Cité en page 2.)
- [TREN 2003] DG TREN. *La sécurité aux passages à niveau*. Commission européenne "www.cete-ouest.developpement-durable.gouv.fr/IMG/pdf/Rapport_Europeen_securite_PN_cle554d5d.pdf", décembre 2003. (Cité en page 2.)
- [Tsai 2009] Du-Ming Tsai et Shia-Chih Lai. *Independent Component Analysis-Based Background Subtraction for Indoor Surveillance*. Image Processing, IEEE Transactions on, vol. 18, no. 1, pages 158–167, 2009. (Cité en page 51.)
- [UIC] UIC. *Union internationale des chemins de fer*. "www.uic.org/spip.php?rubrique1774". (Cité en page 7.)
- [UIC 2012] UIC et RFF. *Tous ensemble pour une meilleure sécurité routière aux passages à niveau*. "www.rff.fr/public/force_download.php?dir=IMG&url=DP_7_juin_V6.pdf", Juin 2012. (Cité en pages ix et 9.)
- [Vezzani 2010] R. Vezzani et R. Cucchiara. *Event driven software architecture for multi-camera and distributed surveillance research systems*. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, pages 1–8, 2010. (Cité en page 17.)
- [Vic 2012] *Les passages à niveau : signalisation et règles de sécurité*. "www.interieur.gouv.fr/content/download/36716/277540/file/2012-passages-niveau.pdf", 2012. (Cité en pages ix et 4.)

- [Viterbi 1967] A.J. Viterbi. *Error bounds for convolutional codes and an asymptotically optimum decoding algorithm*. Information Theory, IEEE Transactions on, vol. 13, no. 2, pages 260–269, 1967. (Cité en page 90.)
- [Wang 2007] Na Wang et Guo-Yu Wang. *Shape descriptor with morphology method for color-based tracking*. International Journal of Automation and Computing, vol. 4, no. 1, pages 101–108, 2007. (Cité en page 18.)
- [Wren 1997] C.R. Wren, A. Azarbayejani, T. Darrell et A.P. Pentland. *Pfinder : real-time tracking of the human body*. Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 19, no. 7, pages 780–785, 1997. (Cité en page 50.)
- [Xu 2005] M. Xu, J. Orwell, L. Lowey et D. Thirde. *Architecture and algorithms for tracking football players with multiple cameras*. Vision, Image and Signal Processing, IEE Proceedings -, vol. 152, no. 2, pages 232–241, 2005. (Cité en page 17.)
- [Yager 1994] Ronald R. Yager, Janusz Kacprzyk et Mario Fedrizzi, éditeurs. *Advances in the dempster-shafer theory of evidence*. John Wiley & Sons, Inc., New York, NY, USA, 1994. (Cité en pages 87, 95 et 102.)
- [Yang 2005] Changjiang Yang, Ramani Duraiswami et Larry Davis. *Fast multiple object tracking via a hierarchical particle filter*. In Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, volume 1, pages 212–219. IEEE, 2005. (Cité en page 52.)
- [Yilmaz 2011] Alper Yilmaz. *Kernel-based object tracking using asymmetric kernels with adaptive scale and orientation selection*. Machine Vision and Applications, vol. 22, no. 2, pages 255–268, 2011. (Cité en page 52.)
- [Zhang 2006] Xiao-Ping Zhang et Zhenhe Chen. *An automated video object extraction system based on spatiotemporal independent component analysis and multiscale segmentation*. EURASIP J. Appl. Signal Process., vol. 2006, pages 184–184, Janvier 2006. (Cité en page 51.)
- [Zhang 2008] Li Zhang, Yuan Li et R. Nevatia. *Global data association for multi-object tracking using network flows*. In Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, pages 1–8, 2008. (Cité en page 52.)