HAL
open science

Estimation du Temps à Collision en Vision Catadioptrique

Fatima Zahra Benamar

To cite this version:

Fatima Zahra Benamar. Estimation du Temps à Collision en Vision Catadioptrique. Informatique [cs]. L'Université de Picardie Jules Verne, Amiens, France; L'université Mohammed V de rabat, Maroc, 2014. Français. NNT : . tel-01133661v2

HAL Id: tel-01133661

https://theses.hal.science/tel-01133661v2

Submitted on 24 Mar 2015

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

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

Thèse de Doctorat

Spécialité Sciences d'ingénieur

présentée à

L'Université de Picardie Jules Verne

par

Fatima Zahra BENAMAR

pour obtenir le grade de Docteur de l'Université de Picardie Jules Verne

Estimation du Temps à Collision en Vision Catadioptrique

Soutenue le 18 octobre 2014, après avis des rapporteurs, devant le jury d'examen :

Mr. Patrick BOUTHEMY, Directeur CNRS	**Rapporteur**
Mr. Abdelmadjid BADRI, Professeur	**Rapporteur**
Mr. Rachid OUAL HAJ THAMI, Professeur	**Examinateur**
Mr. EL Mustapha MOUADDIB, Professeur	**Examinateur**
Mr. Driss ABOUTAJDINE, Professeur	**Examinateur**
MR. Cédric DEMONCEAUX, Professeur	**Examinateur**
MR. Claude PEGARD, Professeur	**Examinateur**
Mme Sanaa EL FKIHI, Professeur HDR	**Examinateur**

Avant Propos

Les travaux présentés dans ce mémoire ont été effectués au Laboratoire de Recherche en Informatique et Télécommunications (LRIT), à la Faculté des Sciences de Rabat (FSR) sous la direction du Professeur Driss ABOUTAJDINE et le co-encadrement de Madame Sanaa EL FKIHI et au laboratoire Modélisation Information et Système (MIS) de l'université de Picardie Jules Verne (UPJV) sous la direction du Professeur EL Mustapha MOUADDIB et le co-encadrement du Professeur Cédric DEMONCEAUX dans le cadre d'une thèse en cotutelle pour l'obtention du grade de docteur au Maroc et de l'université de Picardie Jules Verne.

Au cours de cette thèse, j'ai bénéficié d'une bourse d'excellence octroyée par le centre national pour la recherche scientifique et technique (CNRST) dans le cadre du programme des bourses de recherche initié par le ministère de l'éducation national de l'enseignement supérieur, de la recherche scientifique et de la formation des cadres. Pour la cotutelle, j'ai bénéficié d'une bourse dans le cadre du programme de coopération franco-marocain STIC.

C'est avec la plus profonde gratitude que je souhaite remercier mon directeur de thèse Mr. Driss ABOUTAJDINE, Professeur d'enseignement supérieur à l'université Mohammed V de Rabat et directeur du laboratoire LRIT, je lui exprime ma profonde gratitude de m'avoir accepté au sien du laboratoire LRIT et de m'avoir choisi une bonne équipe de travail, je lui remercie aussi pour ses encouragements et ses conseils éclairants. Durant ces années de préparation, il a su me guider pour à la fin aboutir à l'accomplissement de

cette thèse. Merci aussi à mon co-encadrante Mme. Sanaa El FKIHI, Professeur habilité à l'ENSIAS à Rabat. Je tiens à la remercier pour ces années de soutien, pour ses précieux conseils scientifiques et pour son aide dans les tâches administratives ainsi que pour sa capacité à simplifier les problèmes rencontrés dans le cadre du travail.
Je veux également exprimer toute ma reconnaissance au Professeur d'enseignement supérieur de l'université de Picardie Jules Verne le professeur El Mustapha MOUADDIB, mon deuxième directeur de thèse et le directeur du laboratoire MIS. Je le remercie de m'avoir intégrée au sein de l'équipe *perception robotique* (PR) et de m'avoir consacré tout ce temps et toute cette énergie malgré son emploi du temps très chargé. J'ai ainsi largement pu profiter de sa grande acuité scientifique et de son enthousiasme indéfectible et communicatif pour le travail de ses doctorants. Je lui suis donc redevable de m'avoir permis de poursuivre cette thèse dans de bonnes conditions. Enfin je remercie Mr. Cédric DEMONCEAUX Professeur à l'université de Bourgogne au Creusot en France et mon deuxième co-encadrant, tout d'abord de m'avoir proposé l'idée de ce sujet de thèse ensuite de m'avoir fait profiter de ses compétences. La quantité d'idées, de suggestions et de conseils, qu'il m'a prodiguée pendant toutes ces années continuent encore à m'impressionner. De plus, sa modestie son optimisme et son enthousiasme resteront toujours un exemple à suivre pour moi.
Je tiens à remercier Mr. Rachid OULAD HAJ THAMI, Professeur d'enseignement supérieur à l'ENSIAS de Rabat, d'avoir accepté de présider le jury de ma thèse.
Mes remerciements vont également à Mr. Abdelmajid BADRI, Professeur de l'enseignement supérieur à l'université Hassan II Mohammedia à Casablanca, d'avoir accepté de rapporter ce travail. Je tiens également à remercier Mr. Patrick BOUTHEMY, Directeur de recherche de l' INRIA à Rennes en France, d'avoir accepté de rapporter ce travail. Ses remarques et ses consignes, lors de notre discussion téléphonique, m'ont été d'une grande utilité pour améliorer mon manuscrit de thèse. Je tiens aussi à remercié Mr. Claude PEGARD Professeur d'enseignement supérieur de l'université de Picardie Jules Verne d'avoir accepté d'examiner ce travail.
Je termine par les personnes que je ne saurais jamais remercier assez comme elles méritent. Merci à mes parents, mes frères, ma soeur et ma nièce Serine de m'avoir toujours soutenu moralement et financièrement et sans eux je n'en serais pas là.

Résumé

Cette thèse s'intéresse à l'estimation du temps à collision d'un robot mobile muni d'une caméra catadioptrique. Ce type de caméra est très utile en robotique car il permet d'obtenir un champ de vue panoramique à chaque instant. Le temps à collision a été largement étudié dans le cas des caméras perspectives. Cependant, ces méthodes ne sont pas directement applicables et nécessitent d'être adaptées, à cause des distorsions des images obtenues par les caméras omnidirectionnelles. Dans ce travail, nous proposons d'exploiter explicitement et implicitement le flot optique calculé sur les images omnidirectionnelles pour en déduire le temps à collision (TAC) entre le robot et l'obstacle. Nous verrons que la double projection d'un point 3D sur le miroir puis sur le plan caméra aboutit à des nouvelles formulations du TAC pour les caméras catadioptriques. La première formulation est appelée globale basée sur les gradients d'image, elle estime le TAC en exprimant le mouvement apparent en fonction du TAC et l'équation de la surface plane en fonction des coordonnées images et à partir des paramètres de son vecteur normal. Ces deux outils sont intégrés dans l'équation du flot optique (ECMA) afin d'en déduire le TAC. Cette méthode a l'avantage d'être simple rapide et fournit une information supplémentaire sur l'inclinaison de la surface plane. Néanmoins, la méthode globale basée sur les gradients est valable seulement pour les capteurs para-catadioptriques et elle peut être appliquée seulement pour les surfaces planes. La seconde formulation, appelée locale basée sur le flot optique, estime le TAC en utilisant explicitement le mouvement apparent. Cette formulation nous permet de connaître à chaque instant et pour chaque pixel de l'image le TAC

à partir du flot optique calculé en ce point. Le calcul du TAC en chaque pixel permet d'obtenir une carte des temps à collision. C'est une méthode plus générale car elle est valable pour tous les capteurs à PVU et elle peut être utiliser pour n'importe quelle forme géométrique d'obstacle.
Les deux approches sont validées sur des séquences vidéos de synthèse et réelles.
Mots clés : Temps à collision, TAC, Flot optique, Evitement d'obstacles.

Abstract

Time to contact or time to collision (TTC) is of utmost importance information for animals as well as for mobile robots because it enables them to avoid obstacles ; it is a convenient way to analyze the surrounding environment. The problem of TTC estimation is largely discussed in perspective images. Although a lot of works have shown the interest of omnidirectional camera for robotic applications such as localization, motion, monitoring, few works use omnidirectional images to compute the TTC. In this thesis, we show that TTC can be also estimated on catadioptric images. We present two approaches for TTC estimation using directly or indirectly the optical flow based on de-rotation strategy. The first, called "gradient based TTC", is simple, fast and it does not need an explicit estimation of the optical flow. Nevertheless, this method cannot provide a TTC on each pixel, valid only for para-catadioptric sensors and requires an initial segmentation of the obstacle.The second method, called "TTC map estimation based on optical flow", estimates TTC on each point on the image and provides the depth map of the environment for any obstacle in any direction and is valid for all central catadioptric sensors. Some results and comparisons in synthetic and real images will be given.

Keywords : Time to collision, TTC, optical flow, Obstacle avoidance.

Table des matières

1 Contexte général et problématique 21

1.1 Contexte général 21

1.2 Définition générale 23

1.3 Problématique 24

1.4 Contributions 24

1.5 Plan du mémoire 26

2 Caméras omnidirectionnelles : principes et modélisation 27

2.1 Introduction 28

2.2 Vision omnidirectionnelle : principe et conceptions 29

2.2.1 Images multiples pour l'obtention d'une image omnidirectionnelle . 29

2.2.2 Objectifs grand-angle (fish-eye) 30

2.2.3 Capteurs catadioptriques 32

2.3 Modélisation géométrique des capteurs à point de vue unique 33

2.3.1 Modélisation des caméras perspectives 33

2.3.2 Modélisation des caméras para-catadioptriques 35

2.3.3 Modélisation des caméras catadioptriques avec miroir hyperboloïde 36

2.3.4 Modélisation des caméras catadioptriques : modèle générique 38

2.4 Conclusion 39

3 Estimation du temps à collision : Etat de l'art . . . 41
3.1 Introduction . . . 42
3.2 Méthodes d'estimation du flot optique . . . 42
3.2.1 Estimation du flot optique pour les images perspectives . . . 43
3.2.2 Estimation du flot optique sur les images omnidirectionnelles . . . 47
3.3 Estimation des foyers d'expansion et de contraction . . . 51
3.3.1 Estimation de foyer d'expansion sur les images perspectives . . . 51
3.3.2 Estimation du foyer d'expansion et du foyer de contraction sur les images catadioptriques. . . . 55
3.4 Méthodes d'estimation du TAC . . . 56
3.4.1 Méthode d'estimation du temps à collision dans le cas perspectif . . . 57
3.4.2 Méthode d'estimation du TAC dans le cas omnidirectionnel . . . 74
3.4.3 Discussion et synthèse . . . 77
3.5 Conclusion . . . 78

4 Estimation du TAC en vision catadioptrique . . . 81
4.1 Introduction . . . 82
4.2 Définition du TAC dans le cas catadioptrique . . . 83
4.3 Modèle global du TAC (MGG) . . . 83
4.3.1 Equation de la profondeur de la surface plane en fonction des coordonnées images . . . 85
4.3.2 Modèle du mouvement apparent en fonction du TAC . . . 86
4.3.3 Modèle adapté de TAC basé sur les gradients pour les surfaces planes 87
4.4 Modèle local du TAC (MLFO) . . . 90
4.5 Algorithme de stratégie de ré-orientation de l'image catadioptrique . . . 91
4.5.1 Algorithme d'extraction des foyers d'expansion et de contraction . . . 92
4.5.2 Stratégie de la ré-orientation . . . 95
4.6 Conclusion . . . 97

5 Validation expérimentale et Application . . . 99
5.1 Introduction . . . 99
5.2 Résultats expérimentaux . . . 100
5.2.1 Plate-forme expérimentale utilisée . . . 100

5.2.2 Mouvement axial du capteur catadioptrique vers l'obstacle 103
5.2.3 Mouvement général du capteur catadioptrique : stratégie de ré-orientation de l'image 121
5.2.4 Comparaison des deux méthodes MGG et MLFO 132
5.3 Conclusion 134

6 Conclusion générale et perspectives 135
6.1 Bilan des contributions. 136
6.2 Perspectives 137

Bibliographie 139

Liste des tableaux

3.1 Les différentes approches existantes pour l'estimation du TAC 79

4.1 La sortie du filtre A 95

4.2 Approches du TAC dans la vision catadioptrique 97

5.1 Caractéristiques des capteurs para-catadioptriques virtuel et réel. 102

5.2 Caractéristiques des capteurs hyper-catadioptriques virtuel et réel. 102

5.3 Séquences de synthèse utilisées pour la MGG (translation axiale) 104

5.4 Description des séquences réelles utilisées dans les expérimentations. 108

5.5 Séquences de synthèse dans le cas de la MLFO (translation axiale) 111

5.6 Séquences réelles utilisées dans les expérimentations de la méthode MLFO. 117

5.7 $(\widehat{ARE})$ du TAC avec et sans ré-orientation pour la MGG 127

5.8 Séquence réelle dans le cas de la MGG (mouvement général) 128

5.9 Séquences de synthèse dans le cas de la MLFO (mouvement général) 129

5.10 $\widehat{ARE}$ du TAC avec et sans ré-orientation pour la MLFO 131

5.11 Avantages et inconvénients pour les méthodes MGG et MLFO 133

Table des figures

1.1 (a) : vue perspective. (b) : vue omnidirectionnelle. 22

2.1 Paysage de la ville d'Edimbourg. 28
2.3 Système de caméras multiples 31
2.4 Exemple de lentille fish-eye à gauche, d'image Fish-eye à droite. 31
2.5 Capteur catadioptrique utilisant un miroir de type parabolique. 32
2.6 Modèle perspectif d'une caméra 34
2.7 Configuration géométrique dans le cas para-catadioptrique. 35
2.8 Configuration géométrique dans le cas hyper-catadioptrique 37
2.9 Modèle unifié 38

3.1 Mouvement de translation planaire du capteur para-catadioptrique. 48
3.2 Voisinage adapté au miroir parabolique projeté ensuite sur le plan image . 50
3.3 Foyer d'expansion (FOE). 52
3.4 Distances entre FOE et les points de correspondance 52
3.5 Les positions des FOE et FOC 56
3.6 Mouvement sphérique du système robot-caméra [Stratmann 2004]. 56
3.7 Le champ de mouvement sphérique 57
3.8 Représentation géométrique de la translation d'une surface rigide. 58
3.9 Configuration géométrique dans le cas perspectif. 63
3.10 Cartes des TACs obtenues en utilisant la méthode de Camus 65

3.11 Théoreme de Green appliqué pour calculer la divergence (div). 66
3.12 Arrangement topologique log-polaire 67
3.13 Transformation log-polaire. 68
3.14 Cas 1 70
3.15 Cas 2 71
3.16 Cas 3 72
3.17 Cas 4 73
3.18 Estimation du TAC dans les différents cas du mouvement 74
3.19 Projection perspective et sphérique 75

4.1 Robot Pioneer-AT muni d'une caméra catadioptrique 84
4.2 Projection para-catadioptrique d'une surface plane 85
4.3 Diagramme résumant l'algorithme de la stratégie de réorientation. 92
4.4 Flot optique et les vecteurs FOE(1) et FOC(2) 93
4.5 Binarisation de la matrice des orientations 94
4.6 Histogramme du flot optique 95
4.7 Axes du robot et de la caméra (X' et u) alignés 96
4.8 Axes du robot et du capteur (X' et u) non alignés 96

5.1 Exemple d'une séquence de synthèse paracatadioptrique 101
5.2 Image paracatadioptrique réelle 101
5.3 Deux situations de mouvement du capteur para-catadioptrique 103
5.4 Estimation du TAC pour la MGG 105
5.5 ARE du TAC avec la MGG en présence de bruit 106
5.6 Surfaces planes à droite et à gauche 107
5.7 Estimation du TAC pour les (S_r) et (S_l) 108
5.8 Comparaison entre τ_{S_r} et $(\nu - \tau_{S_l})$ 108
5.9 Deux images des séquences réelles traitées dans les expérimentations 109
5.10 Estimation du TAC dans le cas réel avec la MGG 109
5.11 Suivi des points 3D de la mire 110
5.12 Voisinages utilisés pour le calcul des flots optiques 113
5.13 Estimation du TAC sur les séquences de synthèse avec la MLFO 114
5.14 Les cartes d'erreurs relatives (ARE) des TLK et TALK 115

5.15 ($\overline{ARE}$) du TAC dans le cas du mouvement axial du capteur virtuel 116
5.16 ($\overline{ARE}$) pour la séquence hyper-catadioptrique bruitée 117
5.17 TLK pour les deux séquences réelles : hyper- et para- catadioptriques . . . 119
5.18 Détection des obstacles basée sur la binarisation des cartes de TAC 120
5.19 Mouvement général du capteur catadioptrique embarqué sur le robot . . . 121
5.20 Séquence réelle utilisée dans l'expérimentation de ré-orientation 122
5.21 Images panoramiques 123
5.22 Flot optique de Lucas-Kanade à partir des deux images panoramiques . . . 123
5.23 Matrice d'orientation du flot optique 123
5.24 Matrice d'orientation après avoir effectué biniarisation et filtrage. 124
5.25 Histogramme des orientations du flot optique 124
5.26 Point FOE estimé dans l'image panoramique 125
5.27 FOE et l'angle $\widehat{\rho}$ dans l'image omnidirectionnelle 125
5.28 Quatre images de synthèse paracatadioptrique virtuel en mouvement dans le plan **(X-Y)** 126
5.29 Les vecteurs de flot optique estimés et la position du FOE 127
5.30 Comparaison du TAC (4.28) sans et avec ré-orientation 127
5.31 Comparaison dans le cas réel entre l'estimateur du TAC sans et avec ré-orientation 128
5.32 TLK dans le cas réel sans et avec réorientation 130
5.33 ($\overline{ARE}$) du TLK avec et sans la technique de la ré-orientation 130
5.34 Estimation du TAC sur la séquence réelle avec $|\rho| = 46°$ 131
5.35 Comparaison entre les deux méthodes MGG et MLFO 132

Notations et abréviations

Notations

P	: Point 3D de l'environnement.
m_p	: Point projeté sur le miroir.
m	: Point dans le repère caméra.
p	: Point en pixel dans le repère image.
$P_r^C()$	: Modèle perspectif.
$p_r^\zeta()$	: Modèle parabolique.
$p_r^H()$	: Modèle hyperbolique.
$p_r^S()$	: Modèle unifié.

Abréviations

TAC	: Temps à collision
ECMA	: Equation de la contrainte du mouvement apparent.
FOE	: Foyer d'expansion.
FOC	: Foyer de contraction.
HOF	: Histogramme de flot optique.
PVU	: Point de vue unique.

Publications

Journaux

— F. Benamar, S. Elfkihi, C. Demonceaux, E. Mouaddib, D. Aboutajdine. Estimation des Cartes du Temps de Collision (TTC) basée sur flot optique en Vision Para-catadioptrique. Trait. Signal, vol. 31,pp. no. 197-219, 2014.

— F. Benamar, S. Elfkihi, C. Demonceaux, E. Mouaddib, D. Aboutajdine. Visual Contact with Catadioptric Cameras. Robotics and Autonomous Systems, RAS'15, vol. 64, pp. no.100-119, 2015.

Conférences

— F. Benamar, C. Demonceaux, S. Elfkihi, E. Mouaddib, D. Aboutajdine. Time to Contact Estimation on Paracatadioptric Cameras. In IEEE Int. Conf on Pattern Recognition, ICPR'12, Tsukuba, Japan, November 2012.

— F. Benamar, S. Elfkihi, C. Demonceaux, E. Mouaddib, D. Aboutajdine. Gradient-based Time to Contact on Para-catadioptric Camera. In IEEE, 20th international Conference on Image Processing (ICIP'13), Melbourne, Australia, September 2013.

— F. Benamar, C. Demonceaux, S. Elfkihi, E. Mouaddib, D. Aboutajdine. Estimation des Cartes du Temps de Collision (TTC) en Vision Para-catadioptrique, Congrès des jeunes chercheurs en vision par ordinateur (ORASIS'13), (AFRIF, Bourgogne, France), Juin 2013, l'Abbaye de Cluny.

Chapitre

1 Contexte général et problématique

Sommaire

1.1 Contexte général 21
1.2 Définition générale 23
1.3 Problématique 24
1.4 Contributions 24
1.5 Plan du mémoire 26

1.1 Contexte général

La navigation autonome d'un robot dans un environnement inconnu fait l'objet de nombreux travaux depuis plus de quarante ans. Cet intérêt est motivé par la volonté d'accroître l'efficacité de l'activité humaine et de remplacer l'humain par le robot dans des endroits dangereux. Les premières applications robotiques ont été développées dans l'industrie comme par exemple la production automobile, dans des environnements très contraints. Récemment, les systèmes robotiques sont utilisés dans des environements moins contrôlés, souvent difficiles d'accès par l'homme. Les exemples incluent l'exploration spatiale (mission "curiosity"), l'exploration sous-marine, l'exploration souterraine.

L'usage de la vision pour la navigation du robot offre plusieurs avantages. En général, les caméras sont moins chères que les autres capteurs comme le télémètre laser, sonar et

le radar. La vision offre une diversité d'utilisation qui n'est pas toujours présente dans d'autres dispositifs. Elle permet d'acquérir des informations sur la couleur, la texture et le mouvement. Ces éléments peuvent être combinés pour percevoir la structure de la scène, connaître la forme de la surface et la nature du mouvement.

Avant tout traitement des images acquises par les caméras, la compréhension du système de formation d'images est très important car il intervient crucialement dans les modélisations selon la nature de la caméra employée. De nombreux types de caméras existent, souvent classées en deux catégories « perspectives » et « omnidirectionnelles ».

— Les caméras perspectives (voir l'exemple de la Fig. 1.1a) fournissent un champ de vision restreint de l'environnement. Un point 3D est projeté sur un plan photosensible par une projection centrale.

— Les caméras omnidirectionnelles fournissent une perception à très large champ de vue qui peut atteindre 360° autour de l'axe de la caméra. L'image omnidirectionnelle peut être obtenue par l'ajout d'un miroir au dessus d'une caméra classique (voir l'exemple de la Fig. 1.1b).

(a) (b)

FIGURE 1.1 – (a) : vue perspective. (b) : vue omnidirectionnelle.

Une caractéristique fondamentale de tout système de navigation visuel (artificiel ou biologique) en mouvement dans un environnement dynamique est la capacité de percevoir des objets fixes ou mobiles. Ceci est particulièrement important lorsque l'objet constitue une menace imminente de collision avec l'observateur.

Le mouvement d'un observateur vers une surface, ou de manière analogue, une surface vers un observateur, induit une expansion apparente ou effet imminent (en anglais : looming effect) dans l'image. Le taux de cette expansion fournit un moyen direct de l'estimation de

collision avec la surface. Ceci est généralement appelé « Temps à Collision » ou « Temps à Contact ».

1.2 Définition générale

Le Temps à Collision (TAC) (souvent désigné comme τ) est défini comme le rapport entre le déplacement d qui sépare l'observateur de l'obstacle et la vitesse V de son déplacement [Lee 1976]. Ainsi, pour une caméra sténopée se déplaçant le long de son axe optique, le temps à collision est défini comme suit :

$$\tau = -\frac{d}{V} \tag{1.1}$$

Autrement dit, la vitesse relative du robot est inversement proportionnelle à la distance qui sépare l'obstacle du robot.

Il existe de fortes preuves biologiques qui suggèrent que le TAC est un indice couramment utilisé pour détecter des objets imminents et contrôler le mouvement par rapport aux surfaces. De nombreux travaux ont étudié le comportement animal pour expliquer sur quels principes certains animaux peuvent se baser sur leurs propres déplacements et sur la « parallaxe du mouvement » pour éviter les obtstacles. En particulier les travaux de Franceschini [Fabrizio 1994] ont permis de comprendre comment l'organisation de l'oeil de la mouche et les traitements neuronaux des informations visuelles acquises en vol permettent à cet insecte d'estimer la distance des obstacles perçus frontalement et latéralement et de les éviter. Ces connaissances se sont avérées suffisamment précises pour donner lieu à la réalisation d'un système opto-électronique inspiré du système visuel de la mouche et permettant à divers robots roulants ou volants d'éviter les obstacles. D'autres études qui ont porté sur les sauterelles et d'autres insectes volants ont fait ressortir des réponses neuronales visuelles qui permettent de générer des actions d'évitement d'obstacles [Meldrum 1993]. Srinivasan et al. [Srinivasan 2000] observent comment les abeilles utilisent le mouvement visuel pour ralentir et effectuer des atterrissages en douceur.
Dans tous les cas, la réponse neuronale des insectes a montré la correspondance directe avec l'effet d'expansion. Les chercheurs essayent d'analyser cette proprieté inspirée de la nature afin de formaliser des algorithmes d'évitement d'obstacles pouvant

être implémentés dans des robots. [Lee 1976] a montré qu'un conducteur humain peut contrôler visuellement le freinage du véhicule en utilisant l'estimation du TAC. L'un des principaux avantages à noter de l'équation du TAC comme nous le verrons dans le chapitre 3, est que cette équation peut être entièrement exprimée dans le repère du plan image, sans avoir à mesurer des quantités réelles telles que la distance et la vitesse. Ainsi, les approches d'estimation du TAC à l'aide de la vision par ordinateur, peuvent être utiles pour éviter les obstacles et pour détecter d'éventuelles collisions pour la navigation autonome en robotique mobile.

1.3 Problématique

De nombreuses approches ont été adoptées pour le calcul du TAC dans la vision perspective, peu en omnidirectionnelle. Dans cette thèse nous nous intéressons aux capteurs ommnidrectionnels, en particulier le capteur catadioptrique (miroir convexe couplé à une caméra classique), nous cherchons à estimer le TAC à partir des images catadioptriques. Le TAC est une quantité dérivée du modèle de projection des points de la scène 3D dans le plan image. Dans le cas perspectif, les rayons lumineux des points 3D se projettent suivant la loi perspective, alors que dans le cas catadioptrique les rayons lumineux se projettent dans un premier temps sur le miroir puis sur le plan image. Ceci implique que les expressions obtenues dans le cas classique ne sont pas appropriées dans le cas catadioptrique. Pour répondre à cette problématique, nous proposons dans ce travail de thèse d'introduire de nouveaux modèles d'estimation du TAC basées, directement et indirectement, sur le flot optique et adaptés aux images catadioptriques, en s'inspirant des approches qui ont déjà prouvé leur robustesse dans la vision classique et en tenant compte de la nature de projection de ces images ainsi que leur caractéristiques géométriques et topologiques. Dans ce manuscrit, nous avons choisi d'adapter deux méthodes qui répondent aux mieux aux applications robotiques.

1.4 Contributions

Ainsi, nous avons choisi d'adapter deux méthodes pour l'estimation du TAC développées pour les images perspectives.

— La première méthode a été développée par Horn et al. [Horn 2009]. C'est une méthode basée sur l'estimation implicite du flot optique, calculée à partir de l'équation du flot optique (ECMA) et qui ne nécessite aucun traitement de haut niveau (suivi, détection, estimation explicite du flot optique, etc). Nous verrons que la méthode adaptée de celle-ci sur les caméras omnidirectionnelles a l'avantage d'un calcul direct du TAC en utilisant seulement les dérivées spatiales et temporelles mais elle n'est valable que dans le cas d' images para-catadioptriques et de surfaces planes. Le choix de cette méthode est motivé par le fait qu'elle est robuste, rapide et qu'elle peut être implémentée en temps réel.

— La seconde méthode est la méthode adaptée de Camus [Camus 1995], basée sur l'estimation explicite du flot optique. Cette approche représente une approche plus générale qui est valable pour tout type de capteur catadioptrique à point de vue unique (PVU) et permet de fournir une valeur du TAC en chaque point de l'image indépendamment de la géométrie de l'obstacle. Nous verrons que dans le cas catadioptrique, les méthodes usuelles ne sont pas directement applicables à cause des distorsions introduites par l'ajout d'un miroir. Pour cela, nous introduirons une nouvelle définition du TAC catadioptrique, par suite nous formalisons une expression mathématique dérivée du modèle de projection générique et qui est adaptée aux images catadioptriques.

— Les deux contributions présentées ci-dessus supposent que le mouvement du robot est aligné avec le mouvement du capteur catadioptrique. Pour les généraliser à tout type de mouvement, nous introduisons la stratégie de « ré-orientation » qui est basée sur l'estimation du foyer d'expansion (FOE).

En résumé, nous présentons dans cette thèse deux méthodes d'estimation du TAC qui sont basées sur des contraintes différentes et s'appliquent dans des conditions différentes. La premiè re méthode est globale, destinée aux surfaces planes et qui estime le TAC et les paramètres d'orientation de la surface. Le second est un modèle local capable de fournir une carte de TAC (360°) de l'environnement et valable pour tout type de capteur catadioptrique à point de vue unique (notion précisée dans le chapitre 2 section 2.2.3).

1.5 Plan du mémoire

Ce rapport s'articule autour de six chapitres.

Le chapitre 1 présente une introduction générale, qui décrit les différentes parties traitées dans ce rapport.

Le chapitre 2 traite des capteurs spécifiques utilisés dans ce travail. Il s'agit de la vision omnidirectionnelle, nous décrivons les différentes solutions technologiques existantes pour produire des images omnidirectionnelles. Ensuite, nous verrons comment modéliser ces capteurs.

Le chapitre 3 est consacré à l'état de l'art des méthodes d'estimation du flot optique et d'estimation du temps à collision (TAC) sur les images perspectives et également les images omnidirectionnelles. Nous avons profité de cet état de l'art pour classer les méthodes existentes d'estimation du TAC selon les hypothèses adoptées pour chacune. Ceci nous a permis de relever des constations et des motivations afin de choisir deux méthodes d'estimation du TAC dans le cas perspectif à adapter dans le cas catadioptrique.

Le chapitre 4 fournit l'adaptation de la méthode globale d'estimation du TAC choisie dans le chapitre 3, elle s'agit de la méthode de [Horn 2009]. Cette adaptation commence tout d'abord par formaliser l'équation de la surface plane en coordonnées images, puis exprimer le flot optique qui dérive du modèle parabolique en fonction du TAC. Ces deux outils nous permettent d'exploiter l'équation du flot optique (ECMA) afin d'estimer le TAC et les paramètres d'orientation de la surface plane. Ensuite, nous présentons une méthode plus générale d'estimation du TAC pour tout capteur à point de vue unique, il s'agit de l'adaptation de la méthode locale de [Camus 1995]. Cette adaptation est basée sur l'utilisation d'un modèle géométrique de projection adapté aux images catadioptriques.

Le chapitre 5 présente une série d'expérimentations de nos deux méthodes presentées dans le chapitre 4 sur des séquences d'images de synthèse et sur des séquences réelles. Nous avons imaginé plusieurs *scénarii* afin de valider nos deux modèles. Ensuite ces deux méthodes sont comparées entre elles.

Enfin, le **chapitre 6** clôture ce mémoire de thèse, en donnant des conclusions et des perspectives, des évolutions possibles et des applications envisageables.

Chapitre

2 Caméras omnidirectionnelles : principes et modélisation

Les miroirs feraient bien de
réfléchir un peu plus avant
de renvoyer les images
Le Gone

Sommaire

2.1 Introduction **28**

2.2 Vision omnidirectionnelle : principe et conceptions **29**

2.2.1 Images multiples pour l'obtention d'une image omnidirectionnelle 29

2.2.2 Objectifs grand-angle (fish-eye) 30

2.2.3 Capteurs catadioptriques 32

2.3 Modélisation géométrique des capteurs à point de vue unique **33**

2.3.1 Modélisation des caméras perspectives 33

2.3.2 Modélisation des caméras para-catadioptriques 35

2.3.3 Modélisation des caméras catadioptriques avec miroir hyperboloïde 36

2.3.4 Modélisation des caméras catadioptriques : modèle générique . 38

2.4 Conclusion **39**

2.1 Introduction

Depuis le début de la photographie, les photographes essaient de capturer l'environnement de la manière la plus fidèle et la plus complète. Plusieurs techniques et dispositifs ont été développés pour concrétiser cet objectif. Depuis l'invention de la chambre noire (en latin *camera obscura*) par Ibn El Haytham [1], il suffit de percer un petit trou (sténope) dans une chambre noire, d'y insérer une lentille, pour voir apparaître une image inversée dans le fond blanc de la boîte. Plusieurs procédés ont été développés par la suite jusqu'à ce que D. Rees ait réalisé le premier capteur omnidirectionnel en 1970. La vision panoramique est une réduction de la vision omnidirectionnelle. Elles ont toutes les deux comme objectif de fournir un champ de vision de 360° par rapport à l'axe vertical [Mouaddib 2005a]. La première image panoramique a vu le jour grâce au peintre irlandais Robert Baker (1739 – 1806), le créateur du terme *panorama* [2]. Son oeuvre représente un grand tableau circulaire de la ville d'Edimbourg au Royaume-Uni (voir la Fig. 2.1). La photographie panoramique consiste à créer une image avec un champ de vision de 360°. C'est ce qu'on appelle image panoramique.

FIGURE 2.1 – Paysage de la ville d'Edimbourg.

Le champ de vision panoramique de l'être humain varie entre 170° et 190° alors qu'on trouve dans la nature des animaux qui possèdent un champ de vue presque omnidirectionnel (à titre d'exemple, le lapin a un champ de vue approchant les 360°, la mouche avec ses 3000 lentilles possède une vision de 360°). La vision omnidirectionnelle est le procédé de vision qui fournit une sphère de vue du monde observé à partir de son centre [Mouaddib 2005c]. Le bénéfice de l'omnidirectionnalité pour la navigation de robots mobiles et volants est évident puisqu'elle offre des conditions de navigation autonome plus

1. Scientifique arabe ((965 – 1039)) et père de l'optique moderne.
2. Panorama : du Grec *pan* signifiant toute et *horama* la vue.

avantageuses que celles de la vision perspective. Dans ce chapitre, tout d'abord, nous présentons les différentes conceptions de capteurs produisant des images omnidirectionnelles. Par la suite, nous présentons quelques notions qui seront utilisées dans le cadre de cette thèse sur la modélisation des capteurs perspectifs et des capteurs omnidirectionnels. Enfin, nous proposons un aperçu des trois techniques de formation des images omnidirectionnelles et détaillons le cas des caméras catadioptriques.

2.2 Vision omnidirectionnelle : principe et conceptions

Les approches existantes pour obtenir des images avec large champ de vision peuvent être classées selon les trois catégories suivantes :

— Utilisation des images multiples (système rotatif ou à plusieurs caméras) ;
— Utilisation d'objectifs spéciaux (caméras grand angle « fish eye ») ;
— Utilisation d'un miroir (caméra catadioptrique).

Ce qui caractérise ces différentes catégories est le compromis entre la résolution de l'image obtenue et la vitesse d'acquisition. L'emploi des images multiples a l'avantage de fournir une grande résolution, mais au prix de temps d'acquisition assez élevé. Avec l'utilisation de miroirs convexes ou d'objectifs spéciaux, l'acquisition des images à la fréquence d'acquisition de la caméra utilisée est possible, au détriment d'une résolution plus faible. Toutefois, pour la navigation autonome d'un robot, une acquisition rapide est indispensable pendant les déplacements et une bonne résolution des images n'est pas toujours nécessaire. C'est pourquoi dans la communauté robotique, les caméras panoramiques catadioptriques sont largement utilisées.

Dans ce qui suit, nous donnons une brève description de chacune des trois techniques de formation d'image omnidirectionnelle.

2.2.1 Images multiples pour l'obtention d'une image omnidirectionnelle

Une première approche pour obtenir une image omnidirectionnelle consiste à utiliser plusieurs images acquises par une ou plusieurs caméras.

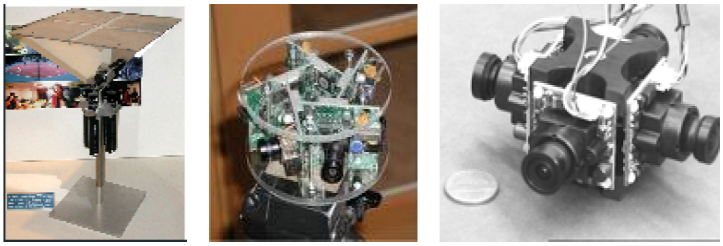
— Dans la littérature, on peut trouver des systèmes qui fournissent des images omnidirectionnelles en utilisant une seule caméra montée autour d'un axe vertical à vitesse angulaire constante (voir Fig. 2.2) [Sarachik 1989], [Jiang 1990], [Matsumoto 1996]. Par juxtaposition des images ainsi obtenues, on peut directement construire une vue panoramique de la scène [Ishiguro 1992] par un procédé de « mosaicing ». Cette technique possède plusieurs inconvénients : la rotation de la caméra provoque un flou dans les images acquises. Ensuite, le temps d'acquisition est assez long. Ces systèmes sont donc destinés à des scènes statiques, et non à des robots en mouvement (on trouve cependant dans [Barth 1996] une méthode qui fait l'acquisition rapide d'images panoramiques).

FIGURE 2.2 – Principe de système rotatif générant une image panoramique

— La seconde méthode consiste à utiliser plusieurs caméras (la Fig. 2.3 présente quelques exemples) : Le système de caméras multiples repose sur la fusion des images acquises avec plusieurs caméras dont chacune recouvre un champ de vue restreint. Pour ce type de capteurs, nous pouvons citer le système développé dans [Nalwa 1996] ou encore la RingCam de Microsoft^R [Cutler 2002] composée de 5 caméras. Pour reconstruire parfaitement l'image omnidirectionnelle, il faut s'assurer de l'alignement des centres optiques des caméras ce qui implique des difficultés dans la fabrication du système.

2.2.2 Objectifs grand-angle (fish-eye)

Une caméra à oeil de poisson (en anglais : « fish-eye ») utilise une lentille spéciale sur la caméra CCD afin de dévier les rayons lumineux. Les objectifs « fish-eye » sont des



−a− *−b−* *−c−*

FIGURE 2.3 – a- Le système de caméra « fullview ». b- Le système « RingCam ». c- La caméra « Poly-camera ».

objectifs à très petites focales, ce qui permet un cadrage large de la scène (voir Fig. 2.4). L'avantage de cette technique est qu'elle est facile d'utilisation et n'occulte pas une partie du champ visuel par la caméra (contrairement aux capteurs catadioptriques par exemple). Les images acquises par ce genre d'objectif sont déformées, et l'analyse de ces images est difficile, principalement, sur la périphérie, là où la résolution est faible et aussi par leur modèle de projection difficile à modéliser.

FIGURE 2.4 – Exemple de lentille fish-eye à gauche, d'image Fish-eye à droite.

La troisième technique de formation d'image omnidirectionnelle qui nous intéresse plus particulièrement, consiste à combiner des miroirs (paraboloïde, hyperboloïde, ellipsoïdal,...) à un système de capteur d'image classique [Hecht 1974]. L'objectif résultant est alors dit catadioptrique : dioptres pour la réfraction (lentilles) et catoptrique pour la réflexion (miroirs). Ce capteur se compose d'un miroir de révolution, dans lequel se reflète une scène qui est filmée par une caméra classique.

2.2.3 Capteurs catadioptriques

L'association d'un ou de plusieurs miroirs à une caméra conventionnelle dans une configuration bien définie forme un dispositif appelé catadioptrique. Les systèmes omnidirectionnels "catadioptriques" sont le résultat de la combinaison des lentilles et des miroirs. La dioptrique est la science des éléments réfractants (lentilles) alors que la catoptrique est la science des surfaces réfléchissantes (miroirs). Les images ainsi obtenues couvrent un champ de vue de 360° (voir la Fig. 2.5).

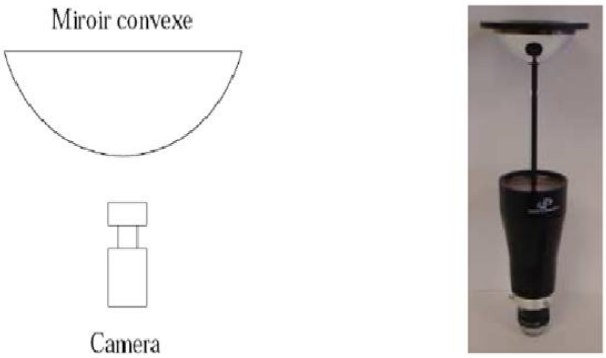


FIGURE 2.5 – Capteur catadioptrique utilisant un miroir de type parabolique.

Dans ce qui suit nous allons utiliser un capteur catadioptrique central, c'est-à-dire un miroir convexe couplé à caméra classique qui a l'avantage d'être un système à PVU (un centre de projection unique) dans certains cas.
Cette théorie du point de vue unique, publiée par Nayar et Baker dans [Baker 1998] signifie que chaque rayon lumineux de la scène passe dans une seule direction à travers ce point de vue, ou de façon équivalente, les rayons réfléchis convergent vers le même point. La contrainte de PVU est réalisée seulement pour des formes spécifiques de miroirs : le miroir elliptique, le miroir conique, le miroir parabolique et le miroir hyperbolique.
Les caméras catadioptriques à centre de projection unique combinent deux caractéristiques importantes : un centre de projection unique et un champ de vision important. Par contre, la vision omnidirectionnelle présente un inconvénient majeur matérialisé par les distorsions inhérentes liées au miroir. Celles-ci rendent inapplicables les outils classiques de traitement d'images.

2.3 Modélisation géométrique des capteurs à point de vue unique

Dans cette section, nous décrivons dans un premier temps la configuration géométrique des projections dans le cas des caméras classiques et dans un deuxième temps, nous présentons la projection dans le cas catadioptrique avec miroir paraboloïde, hyperboloïde et sphérique.

2.3.1 Modélisation des caméras perspectives

Un des modèles classiques de formation des images les plus connus dans la littérature est le modèle sténopé (également appelé modèle perspectif ou en anglais : "pinhole"). Le modèle géométrique sténopé est constitué d'un plan, appelé plan rétinien, dans lequel l'image se forme à l'aide d'une projection perspective. Il s'agit d'un modèle central où tous les rayons optiques s'intersectent en un même point appelé centre optique (ou centre de projection). Ainsi, tout point 3D se projette selon la droite le reliant avec le centre optique de la caméra. Le point image s'obtient à l'intersection de cette droite et du plan image (situé à la distance focale f, voir la Fig. 2.6). L'image se focalise à une distance $Z = -f$ derrière le centre optique sur le capteur (CCD ou CMOS). Notons que le centre optique peut être placé devant ou derrière le plan rétinien. Par simple symétrie centrale (de centre optique), on considère que le centre optique est situé à $Z = f$ comme illustré dans la figure 2.6. Toutefois, les deux configurations sont modélisées avec les mêmes paramètres à un signe près sur la distance focale.

Soit $\widetilde{P} = (X, Y, Z, 1)^\intercal$ les coordonnées homogènes du point P exprimées dans le repère caméra R_c, défini par le centre O_c et les axes $\mathbf{X}_c$, $\mathbf{Y}_c$ et $\mathbf{Z}_c$ (si le point 3D est exprimé dans un repère monde, il suffit d'appliquer un changement de base). Le point 3D $\widetilde{P} = (X, Y, Z, 1)$, est projeté dans le plan image normalisé en $\widetilde{m} = (x, y, 1)^\intercal$ par la fonction de projection perspective $p_r^C()$, tel que :

$$\widetilde{m} = p_r^C(\widetilde{P}) \quad \text{avec} \quad \begin{cases} x = f\frac{X}{Z} \\ y = f\frac{Y}{Z} \end{cases} \tag{2.1}$$

pour passer du plan normalisé (repère caméra) au repère image (repère pixellique), on utilise la projection orthographique. Il s'agit d'une transformation affine donnée par un en-

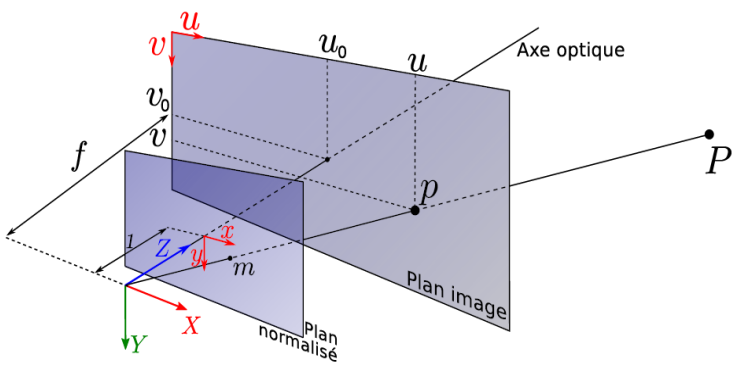


FIGURE 2.6 – Modèle perspectif d'une caméra. Un point 3D P est projeté dans le plan image de la caméra en un point p

semble de rayons parallèles orthogonaux au plan image. La projection orthographique a été introduite comme une approximation de la projection perspective par [Aloimonos 1990]. Le modèle de caméra affine a été formulé par [Mundy 1992] et souvent utilisé dans les problèmes d'estimation de mouvement. Cette projection est un cas limite de la projection perspective [Faugeras 1993]. Donc le passage au plan image se fait comme suit :

$$\widetilde{p} = K\widetilde{m} \tag{2.2}$$

avec K représentant la matrice des paramètres intrinsèques définie par l'équation suivante :

$$K = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.3}$$

Notons que, les paramètres intrinsèques sont les paramètres qui modélisent la géométrie interne et les caractéristiques optiques du capteur. L'ensemble de ces paramètres contient :

— La distance focale f : c'est la distance (en mm) qui sépare le plan rétinien (plan image) et le centre optique O_c.

— Les paramètres de conversion α_u et α_v : représentent l'ajustement horizontal et vertical qui permettent de passer d'un repère du plan rétinien (exprimé en mm) au repère image (exprimé en pixel).

— Le point principal (u_0, v_0) : point défini comme étant le projeté du centre optique O_c sur le plan image, exprimé en pixels.

Enfin, signalons que le modèle de projection perspectif est un modèle à point de vue

unique. En effet, toutes les lignes de vue s'intersectent en un seul point qui est le centre de projection.

2.3.2 Modélisation des caméras para-catadioptriques

La modélisation d'une caméra catadioptrique panoramique relève de la même démarche que celle d'une caméra perspective en ajoutant la réflexion des rayons lumineux sur le miroir de révolution. Cette réflexion introduit une transformation supplémentaire dont il faut tenir compte et qui dépend de la nature de la surface du miroir [Mouaddib 2005c], [Mouaddib 2005b].

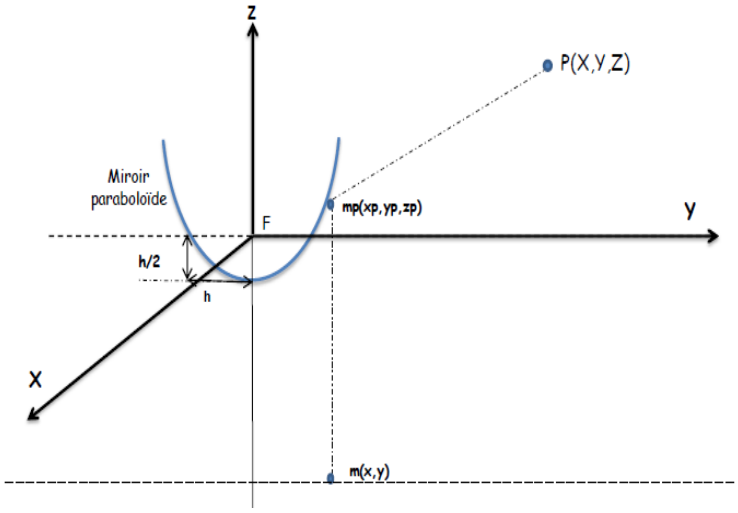


FIGURE 2.7 – Configuration géométrique dans le cas para-catadioptrique.

Le modèle parabolique est illustré par la figure 2.7. Par convention, le repère (**X, Y, Z**) est centré au foyer du paraboloïde en un point F. Les axes du repère sont choisis tels que **Z** soit confondu avec l'axe principal du miroir. Soit $P(X, Y, Z)$ un point $3D$ de l'environnement (voir Fig. 2.7). Le point P se projette sur le miroir paraboloïde en un point $m_p = (x_p, y_p, z_p)$ suivant le modèle suivant :

$$m_p = p_r^{\zeta}(P) \quad \text{avec} \quad m_p : \begin{cases} x_p = \frac{h\ X}{\rho - Z} \\ y_p = \frac{h\ Y}{\rho - Z} \\ z_p = \frac{h\ Z}{\rho - Z} \end{cases} \tag{2.4}$$

avec $\rho = \sqrt{X^2 + Y^2 + Z^2}$ et h est le paramètre du miroir d'équation :

$$z_p = \frac{x_p^2 + y_p^2 - h^2}{2h} \tag{2.5}$$

Ensuite, puisque nous avons une projection orthographique $p_r^O()$ (pour respecter le PVU), les coordonnées sur le plan image au point $m = (x, y)$ sont :

$$m = p_r^O(m_p) \quad \text{avec} \quad m : \begin{cases} x = x_p = \frac{h\ X}{\rho - Z} \\ y = y_p = \frac{h\ Y}{\rho - Z} \end{cases} \tag{2.6}$$

Pour passer aux coordonnées pixelliques $p = (u, v)$, il suffit de multiplier le vecteur m par la matrice K (eq. (2.3)) des paramètres intrinsèques : $p = Km$. Enfin, la fonction complète de projection d'un point 3D (exprimé dans le repère caméra R_c) dans le plan image pixellique est :

$$p = Kp_r^O(p_r^\zeta(P)) \tag{2.7}$$

2.3.3 Modélisation des caméras catadioptriques avec miroir hyperboloïde

Yamazawa et al. [Yamazawa 1993] ont proposé un capteur d'images omnidirectionnelles en utilisant un miroir hyperboloïde, appellé *HyperOmniVision*. Le miroir hyperbolique a deux points focaux, F et F' (voir Fig. 2.8). L'équation d'un hyperboloïde dans un repère centré en son foyer est :

$$\frac{(z_p + e)^2}{a^2} - \frac{x_p^2 + y_p^2}{b^2} = 1 \tag{2.8}$$

avec e représentant l'excentricité : $e = \sqrt{a^2 + b^2}$, où $a = 1/2(\sqrt{4e^2 + 4h^2} - 2h)$ et $b = \sqrt{h\sqrt{4e^2 + 4h^2} - 2h^2}$.

La droite passant par le foyer de l'hyperboloïde et par le point $P(X, Y, Z)$, donné dans le repère du miroir est :

$$x_p = \lambda X, y_p = \lambda Y, z_p = \lambda Z. \tag{2.9}$$

On remplace ces termes dans l'équation (2.8) et on déduit les solutions de λ qui correspondent aux deux intersections avec le miroir :

$$\lambda = \frac{b^2(-eZ \pm a\rho)}{b^2 Z^2 - a^2(X^2 + Y^2)} \tag{2.10}$$

avec $\rho = \sqrt{X^2 + Y^2 + Z^2}$.

La projection perspective p_r^C se fait de la manière suivante (en tenant compte du fait que la distance entre les points focaux est égale à $2e$) :

$$m = p_r^C(m_p) \quad \text{avec} \quad \begin{cases} x/f = \frac{x_p}{2e + z_p} \\ y/f = \frac{y_p}{2e + z_p} \end{cases} \tag{2.11}$$

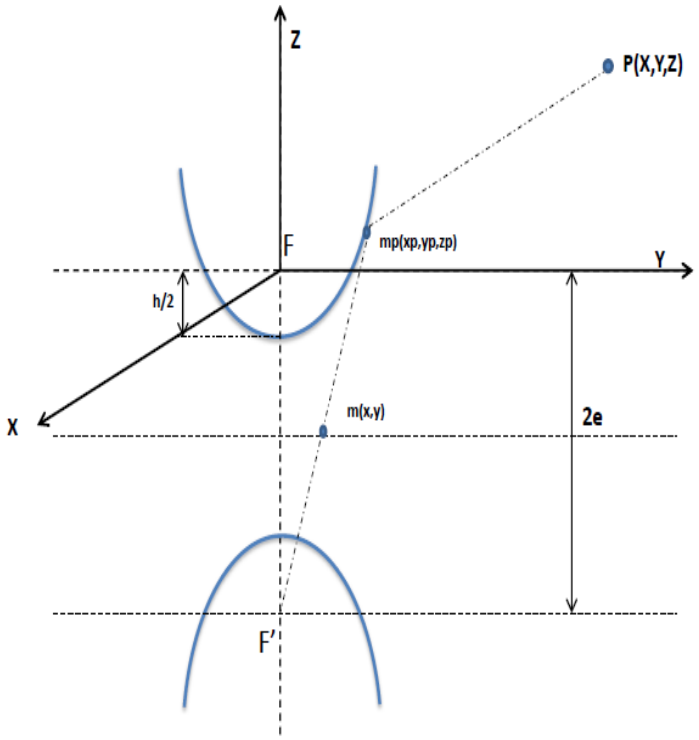


FIGURE 2.8 – Configuration géométrique dans le cas catadioptrique avec miroir hyperbolique.

Sans perte de généralité on pose $f = 1$ et on remplace les autres variables pour obtenir les coordonnées (x, y) en fonction des coordonnées du point $3D$ et les paramètres du miroir :

$$m = p_r^C(m_p) \quad \text{avec} \quad \begin{cases} x = \frac{x_p}{2e+z_p} \\ y = \frac{y_p}{2e+z_p} \end{cases} \tag{2.12}$$

Puis, on remplace λ,x_p, y_p et z_p pour obtenir les coordonnées du point image sur la caméra normalisée :

$$m = p_r^H(P) \quad \text{avec} \quad \begin{cases} x = \frac{4Xeh/\sqrt{4e^2+4h}}{2e/\sqrt{4e^2+4h^2}(\rho \pm Z)} \\ y = \frac{4Yeh/\sqrt{4e^2+4h}}{2e/\sqrt{4e^2+4h^2}(\rho \pm Z)} \end{cases} \tag{2.13}$$

De la même manière que l'équation (2.7), le passage aux coordonnées pixelliques se fait comme suit :

$$p = K(p_r^H(P)) \tag{2.14}$$

2.3.4 Modélisation des caméras catadioptriques : modèle générique

Geyer et Daniilidis [Geyer 2001] ont introduit une théorie d'unification pour les caméras catadioptriques ayant un point de vue unique. En substance, ils ont montré qu'une projection catadioptrique centrale est équivalente à un mapping en deux étapes via une sphère. Le théorème d'équivalence [Geyer 2001] entre toute projection centrale et une projection sphérique suivie d'une projection sur le plan fournit un modèle unifié qui modélise toutes les caméras à point de vue unique.

Dans la figure 2.9, le projeté P_s du point 3D $P(X, Y, Z)$ sur la sphère est obtenu par une projection centrale via le centre de la sphère O_c. La projection stéréographique du point $P_s(x_s, y_s, z_s)$ via un point O_p qui appartient à l'axe **Z** de la sphère donne l'image $m(x, y)$ de $P(X, Y, Z)$ sur l'image catadioptrique. Le paramètre ξ définit la forme du miroir et φ représente la distance entre le centre O_c et le plan image.

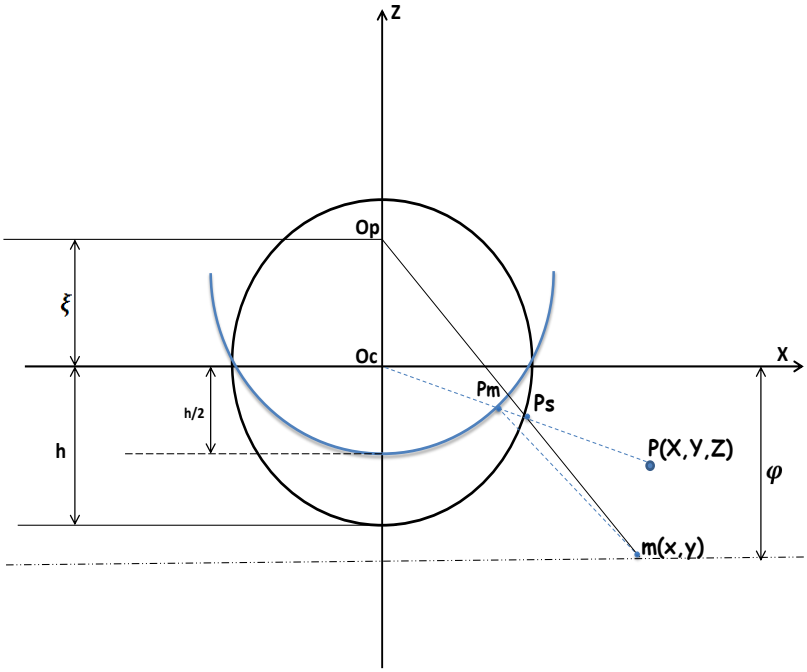


FIGURE 2.9 – Equivalence entre un modèle de projection catadioptrique et la projection sur la sphère suivie par une projection sur le plan image.

Ainsi, le point P_s appartenant à la droite passant par le centre de la sphère et le point

3D P vérifie :

$$x_s = \lambda X; y_s = \lambda Y; z_s = \lambda Z \tag{2.15}$$

L'équation de la sphère est :

$$x_s^2 + y_s^2 + z_s^2 = R^2 \tag{2.16}$$

Sans perte de généralité on pose $R = 1$. D'où :

$$\lambda = \frac{\pm 1}{\sqrt{X^2 + Y^2 + Z^2}} \tag{2.17}$$

La projection sur le plan de l'image est obtenue tout d'abord par :

$$m = p_r(m_s) \quad \text{avec} \quad \begin{cases} \frac{x_s}{\xi - z_s} = \frac{x}{\xi + \varphi} \\ \frac{y_s}{\xi - z_s} = \frac{y}{\xi + \varphi} \end{cases} \tag{2.18}$$

Enfin, le modèle générique est le suivant :

$$m = p_r^S(P) \quad \text{avec} \quad \begin{cases} x = \frac{(\xi+\phi)X}{\xi\sqrt{X^2+Y^2+Z^2} - Z} \\ y = \frac{(\xi+\phi)Y}{\xi\sqrt{X^2+Y^2+Z^2} - Z} \end{cases} \tag{2.19}$$

Par passage aux coordonnées pixelliques, nous utilisons la même transformation que l'équation (2.14) :

$$p = K(p_r^S(P)) \tag{2.20}$$

2.4 Conclusion

Dans l'espoir d'élargir le champ visuel, plusieurs conceptions de capteurs ont été mises en place. Nous avons donné un aperçu des différentes catégories connues dans la vision omnidirectionnelle pour construire une image à grand champ de vue. Ceci a commencé par une approche naïve comme le système rotatif, ensuite elle s'est étendue aux techniques de mosaïque et l'utilisation des lentilles spéciales (lentille « fish-eye ») jusqu' à la conception d'un capteur catadioptrique qui permet d'acquérir un champ de vision de 360°. Ce chapitre a donné aussi un aperçu de la modélisation des capteurs à PVU, que nous utiliserons par la suite dans le chapitre 4.

Chapitre

3 Estimation du temps à collision : Etat de l'art

Il faut éclairer l'histoire par les lois
et les lois par l'histoire
Montesquieu

Sommaire

3.1 Introduction **42**

3.2 Méthodes d'estimation du flot optique **42**

3.2.1 Estimation du flot optique pour les images perspectives 43

3.2.2 Estimation du flot optique sur les images omnidirectionnelles . 47

3.3 Estimation des foyers d'expansion et de contraction **51**

3.3.1 Estimation de foyer d'expansion sur les images perspectives . . 51

3.3.2 Estimation du foyer d'expansion et du foyer de contraction sur les images catadioptriques. 55

3.4 Méthodes d'estimation du TAC **56**

3.4.1 Méthode d'estimation du temps à collision dans le cas perspectif 57

3.4.2 Méthode d'estimation du TAC dans le cas omnidirectionnel . . 74

3.4.3 Discussion et synthèse 77

3.5 Conclusion **78**

3.1 Introduction

Dans ce chapitre, nous présentons un état de l'art composé de trois parties. La première partie est consacrée à lister les différentes méthodes d'estimation du flot optique sur les images perspectives et les images omnidirectionnelles. Nous en profiterons pour justifier le choix de la méthode d'estimation du flot optique adoptée. L'un des premiers intérêts du flot optique est qu'il permet d'estimer le foyer d'expansion (FOE). Ainsi, grâce au FOE nous pouvons déduire la direction du mouvement. Nous consacrons d'ailleurs la deuxième partie de ce chapitre à l'étude du FOE. La dernière grande partie est dédiée à la présentation des différentes méthodes d'estimation du TAC. Pour cela nous exposons dans un premier temps, les méthodes les plus connues dans la littérature basées explicitement et implicitement sur le flot optique. Ensuite nous donnerons une présentation des méthodes proposées pour l'estimation du TAC en vision omnidirectionnelle. Cet état de l'art nous permettra de conclure sur le choix de méthodes perspectives à adapter dans le cas para-catadioptrique.

3.2 Méthodes d'estimation du flot optique

Lorsqu'un observateur (un animal, une caméra, etc.) est en mouvement relatif par rapport aux objets de son environnement, un champ visuel de déplacement peut être perçu. Ce champ est appelé flot optique (ou flux optique). Il permet d'expliquer des variations dans une image animée en terme de déplacements de points images. Toutes les approches qui ont été proposées pour résoudre ce problème reposent soit sur la recherche d'un champ le plus régulier possible, soit sur l'ajout d'une hypothèse supplémentaire sur la forme du champ de flot optique que l'on souhaite trouver. Le flot optique est utilisé dans plusieurs applications comme : la reconstruction 3D [Yang 2012], la segmentation[Ibarra 2013], [Shanshan 2013], le suivi, l'estimation de l'ego-mouvement [Gluckman 1998], [Vassallo 2002], [Radgui 2009], la surveillance et la navigation des robots mobiles terrestres et volants [Li 2006].
Dans ce qui suit, nous présentons tout d'abord les méthodes d'estimation qui ont été développées pour les images perspectives et ensuite les méthodes d'estimation pour les images omnidirectionnelles.

3.2.1 Estimation du flot optique pour les images perspectives

De nombreuses approches ont été proposées pour estimer le mouvement apparent sur les images perspectives. Dans cette section, nous exposons les méthodes les plus connues dans la littérature [Mitiche 1996]. Elles sont regroupées suivant quatre catégories : les méthodes différentielles, les méthodes basées sur la mise en correspondance, les méthodes fréquentielles et finalement les méthodes basées sur une décomposition des images.

3.2.1.1 Méthodes différentielles

Les méthodes différentielles calculent le flot optique de l'image à partir de l'équation de contrainte du mouvement apparent. Cette équation est basée sur la contrainte de conservation de l'intensité avec le mouvement, en effet :
Soit un point 3D réel $P(t) = (X(t), Y(t), Z(t))$ projeté sur le plan image au point $p(t) = (x(t), y(t))$ suivant la fonction suivante :

$$\mathbf{R}^3 \longmapsto \mathbf{R}^2 \quad P(t) \longmapsto p(t) \tag{3.1}$$

Considérons une séquence d'images $I(x(t), y(t), t)$, le mouvement apparent dans l'image, plus communément appelé flot optique, est la vitesse de déplacement $\overrightarrow{v}(x(t), y(t), t)$ des points physiques dans la séquence d'images, c'est-à-dire :

$$\overrightarrow{v} = (v_1, v_2) = (\frac{dx}{dt}(t), \frac{dy}{dt}(t)) \tag{3.2}$$

La contrainte de l'illumination constante signifie que l'intensité lumineuse du point $p(t)$ ne change pas au cours du temps. Cela se traduit mathématiquement par :

$$\frac{dI}{dt}(x(t), y(t), t) = 0 \quad \forall t. \tag{3.3}$$

En utilisant la formule de Taylor et en négligeant les termes de second ordre, le flot optique est alors contraint par l'équation suivante :

$$\overrightarrow{\nabla I}.\overrightarrow{v} + \frac{\partial I}{\partial t} = 0 \tag{3.4}$$

où $\overrightarrow{\nabla I}$ est le gradient spatial de l'image. Cette équation est appelée l'équation de contrainte du mouvement apparent (notée ECMA). Elle génère un problème connu dans la littérature sous le nom « problème de l'ouverture » car on ne dispose que d'une seule contrainte pour

trouver deux inconnues v_1 et v_2. Pour résoudre ce problème de l'ouverture, plusieurs approches ont été proposées.

Régularisation globale :

En 1981, Horn et Schunk [Horn 1981] sont les premiers à introduire la notion de la régularisation dans une méthode de calcul du flot optique. La solution proposée revient à minimiser une fonctionnelle à deux termes : le premier terme consiste à déterminer le flot qui minimise l'ECMA :

$$\Gamma_1[\overrightarrow{v}] = \iint_{\Omega} (\overrightarrow{v}.\overrightarrow{\nabla} I + \frac{\partial I}{\partial t})^2 dxdy \tag{3.5}$$

où Ω représente le plan image.
le second terme est la fonctionnelle de régularisation de Tikhonov qui cherche à contraindre le flot à être le plus lisse possible :

$$\Gamma_2[\overrightarrow{v}] = \iint_{\Omega} \ ||\overrightarrow{\nabla} v \ ||^2 dxdy \tag{3.6}$$

La solution d'Horn et Schunk consiste à chercher le vecteur $\overrightarrow{v}$ en minimisant la fonctionnelle totale suivante :

$$\Gamma[\overrightarrow{v}] = \Gamma_1[\overrightarrow{v}] + \lambda \Gamma_2[\overrightarrow{v}] \tag{3.7}$$

où λ est un paramètre qui permet d'ajuster l'effet du terme (3.6). Cette solution résout le problème de l'ouverture, néanmoins, le terme de Tikhonov induit une forte pénalisation des gradients. Cela conduit à un lissage excessif du flot optique. Cette méthode est la première méthode différentielle globale d'estimation du flot optique développée dans la littérature.

Régularisation locale :

Pour remédier aux inconvénients des méthodes globales qui calculent un flot optique dense sur l'ensemble de l'image, Lucas et Kanade [Lucas 1981] ont proposé une méthode basée uniquement sur le terme (3.5) sans avoir recours au second terme. Ils supposent que le flot optique $\overrightarrow{v}$ est constant dans un voisinage autour du point $p(t)$. Cette méthode offre

une meilleure combinaison entre deux caractéristiques critiques à savoir, la précision et le temps de calcul. Sa fonctionnelle locale Γ^l est donnée par :

$$\Gamma^l[\overrightarrow{v}] = \Gamma_1[\overrightarrow{v}(q)] \quad \text{si} \quad q\epsilon\ \vartheta_p \tag{3.8}$$

avec ϑ_p un voisinage classique du point p de l'image I. L'adaptation de cette méthode en vision catadioptrique est traitée dans la section 3.2.2.2.

3.2.1.2 Méthodes basées sur la mise en correspondance

Ces méthodes consistent à mettre en correspondance des caractéristiques de l'image pour calculer le champ de déplacement $\overrightarrow{d}$ entre deux instants consécutifs t et $t+1$ [Anandan 1989], [Giachetti 2000]. Il s'agit de minimiser la somme des carrées (SSD) suivante :

$$SSD(p, \overrightarrow{d}) = \sum_{(i,j)\in V_p} [I((i,j),t) - I((i,j) + \overrightarrow{d}, t+1)]^2 \tag{3.9}$$

Où V_p est un voisinage centré autour du point $p = (x, y)$. Ces méthodes sont robustes et simples à mettre en oeuvre mais elles sont souvent coûteuses en terme de temps de calcul.

3.2.1.3 Méthodes fréquentielles

Ces méthodes étudient l'ECMA de l'image I en l'analysant dans le domaine fréquentiel. D'après l'hypothèse de contrainte de flot optique (3.3), on a la relation suivante :

$$I(p,t) = I(p + vt, 0) \tag{3.10}$$

avec $p = (x, y)$ est un point de l'image et $\overrightarrow{v} = (v_1, v_2)$ est le vecteur vitesse recherché. La transformée de Fourier $\widehat{I}$ appliquée à cette fonction est :

$$\widehat{I}(f_x, f_y, f_t) = \widehat{I}(f_x, f_y)\delta(v_1 f_x + v_2 f_y + f_t) \tag{3.11}$$

où δ est la fonction de Dirac et f_x, f_y et f_t représentent les fréquences spatiales et temporele de l'image I. Donc l'estimation du flot optique est extraite de l'équation (3.11) en identifiant dans l'espace des fréquences un plan de vitesse d'équation :

$$v_1 f_x + v_2 f_y + f_t = 0 \tag{3.12}$$

Le problème de l'ouverture est contourné en utilisant la constance du flot sur les supports de ces filtres. Ces méthodes sont divisées en deux catégories : méthodes basées sur l'energie

du signal filtré [Adelson 1985] ou basées sur la phase [Heeger 1988]. Généralement, ces méthodes donnent de bons résultats mais elles restent lentes en terme de temps de calcul.

3.2.1.4 Méthodes basées sur une décomposition des images

Une autre approche proposée pour contourner le problème de l'ouverture consiste à projeter l'équation du flot optique sur une base de fonctions $(\psi^n)_{n=1,.....,N}$ de $\mathbf{L}^2(\mathbb{R}^2)$. On construit une famille de fonctions translatées (ψ_u^n) comme :

$$\psi_u^n(x,y) = \psi_n(x-u_1, y-u_2) \quad \forall n = 1, ..., N, \quad u = (u_1, u_2) \tag{3.13}$$

et on suppose que le flot optique au point u est constant sur le support de ces fonctions. En général, ces fonctions sont des filtres spatiaux ou bien une famille d'ondelettes. Cette projection permet d'obtenir un système de N équations :

$$\iint_\Omega (\overrightarrow{\nabla} I . \overrightarrow{v}((x,y),t) + \frac{\partial I}{\partial t})\psi_u^n(x,y)dxdy = 0 \quad \forall n = 1, ..., N \tag{3.14}$$

Weber et Malik [Weber 1995] proposent d'utiliser plusieurs filtres spatio-temporels. Bernard [Bernard 1998] a utilisé une base d'ondelettes $\psi^n(x,y)_{n=1,.....,N}$, alors que Bruno et al. [Bruno 2002] ont utilisé une famille de fitre de Gabor pour contraindre l'ECMA sur le support de ces filtres.

3.2.1.5 Comparaison des méthodes

Les méthodes décrites auparavant et d'autres ont été evaluées et comparées qualitativement par Barron et al. [Barron 1994]. Ces auteurs ont examiné ces algorithmes en utilisant plusieurs séquences d'images. Une de leurs conclusions est que la méthode différentielle locale de Lucas et Kanade [Lucas 1981] et la méthode de Fleet et Jepson [David 1990] sont les plus performantes en robustesse et en temps d'exécution. Nous profiterons de cette remarque pour calculer le TAC developpé dans le chapitre 4 en utilisant le flot optique de [Lucas 1981]. Dans ce qui suit nous présentons l'estimation du flot optique dans le cas omnidirectionnel.

3.2.2 Estimation du flot optique sur les images omnidirectionnelles

Le problème de l'estimation du flot optique a été largement étudié pour les images perspectives. Les images omnidirectionnelles nous permettent d'avoir une information globale sur le mouvement et sur son sens via la présence de deux foyers d'expansion et/ou de contraction (FOE et FOC). De plus, l'analyse du mouvement à partir des images omnidirectionnelles permet de résoudre le problème de l'ambiguité lié au flot optique, présente dans les images perspectives, entre une rotation et une translation.
Nous présentons deux catégories d'estimation du flot optique dans le cas omnidirectionnel qui existent dans la littérature. La première catégorie calcule le flot sur la sphère unité alors que la deuxième le calcule directement sur le plan image.

3.2.2.1 Estimation sur la sphère

Comme nous l'avons signalé dans la section 3.2.1.1, les méthodes différentielles représentent un axe intéressant pour de nombreux auteurs pour les adapter en vision omnidirectionnelle. Nous citons par exemple : Daniilidis et al. [Daniilidis 2002] qui ont adapté la méthode locale de [Lucas 1981] sur la sphère en définissant une fonction gaussienne et un opérateur sphérique pour résoudre l'ECMA définie en coordonnées sphériques. Yosihiko et al. [Yoshizaki 2008] ont adapté la méthode globale de Horn et Schunk [Horn 1981] en minimisant leur fonctionnelle définie en coordonnées sphériques. Tosic et al. [Tosic 2005] ont calculé la corrélation entre deux images sphériques de la scène en se basant sur une technique de multirésolution de L-niveaux. Radgui et al. [Radgui 2011] se sont basés sur la méthode de Bruno et al. [Bruno 2002] et Bernard et al. [Bernard 1998] pour adapter le flot optique aux images catadioptriques en faisant une décomposition de l'ECMA avec une famille d'ondelettes sphériques. Récemment Alibouch et al. [Alibouch 2012] ont adapté la méthode de Gautama et al. [Gautama 2002]. Cette dernière est basée sur l'utilisation des filtres complexes en 2D pour extraire la phase spatiale, ensuite le gradient de la phase temporel est estimé à chaque position de la séquence alors qu'une mesure de fiabilité est appliquée pour déterminer les composantes valides de vitesse. L'adaptation de cette méthode pour les images catadioptriques consiste à reformuler l'équation de la contrainte de gradient de phase sur la sphère unité.

3.2.2.2 Estimation du flot optique sur le plan image

Cette section a pour but de présenter une des méthodes d'estimation du flot optique sur le plan de l'image omnidirectionnelle. C'est la méthode du flot optique adapté aux images para-catadioptriques. Cette contribution a été publiée par Radgui et al. [Radgui 2009] ; elle consiste à adapter le flot optique de Lucas Kanade en formulant un nouveau modèle de mouvement et un voisinage adapté qui tiennent compte de la géométrie du capteur para-catadioptrique. Notons que cette technique sera utilisée dans notre estimation de TAC. Nous donnons un aperçu sur le calcul de cette méthode.

Modèle du mouvement adapté :

Le modèle du mouvement adapté est basé sur le modèle de projection parabolique (2.7). La translation planaire 3D dans le plan **X-Y** de la caméra est notée $d = (d_x, d_y)$ (on suppose qu'il n'y a pas de translation suivant l'axe **Z**, voir la Fig. 3.1).

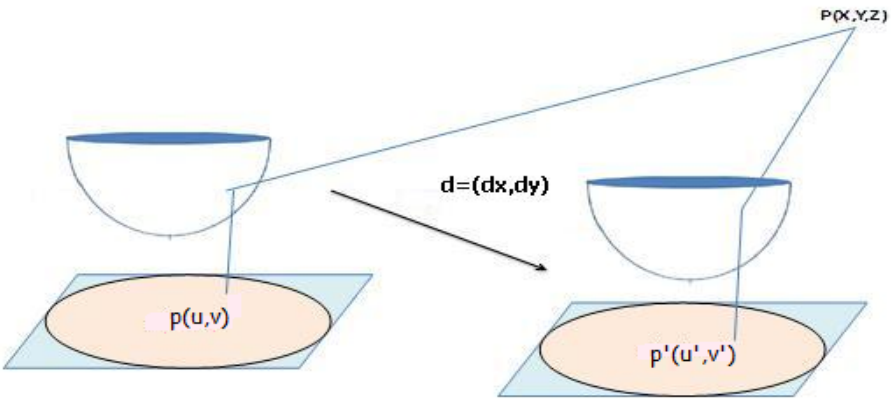


FIGURE 3.1 – Mouvement de translation planaire du capteur para-catadioptrique.

Notons que $p(u, v)$ et $p'(u', v')$ sont les projections (coordonnées pixelliques) du même point 3D P résultants du mouvement de la caméra dans le plan perpendiculaire à son axe optique.

La projection des points 3D déplacés dans l'image peut être exprimée comme suit :

$$p' = Kp_r^O(p_r^\zeta(P)) \quad \text{avec} \quad p' : \begin{cases} u' = \alpha_u \frac{h\ (X+d_x)}{\rho_d - Z} + u_0 \\ v' = \alpha_v \frac{h\ (Y+d_y)}{\rho_d - Z} + v_0 \end{cases} \tag{3.15}$$

avec $\rho_d = \sqrt{(X + d_x)^2 + (Y + d_y)^2 + Z^2}$ et (u_0, v_0) est le point principal. Pour simplifier cette modélisation, nous considérons que les composants $d = (d_x, d_y)$ peuvent être négligés

devant les coordonnées du point P. Nous considérons également que les facteurs d'échelle, suivant les deux directions, sont égaux $\alpha_u = \alpha_u = \alpha$. En utilisant ces hypothèses nous obtenons :

$$\begin{cases} u' \simeq u + \alpha \frac{\alpha h d_x}{\rho_d - Z} \\ v' \simeq v + \alpha \frac{\alpha h d_y}{\rho_d - Z} \end{cases} \tag{3.16}$$

Soient : $U = u - u'$ et $V = v - v'$. Les auteurs de [Radgui 2008] montrent alors que (U, V) peuvent s'écrire sous la forme :

$$\begin{cases} U = a(u - u_0)^2 + a(v - v_0)^2 + c \\ V = b(u - u_0)^2 + b(v - v_0)^2 + d \end{cases} \tag{3.17}$$

où a, b, c et d sont les paramètres du modèle du mouvement à estimer.
Suivant le modèle de mouvement obtenu (eq.(3.17)), la translation du point P dans l'espace décrit un cercle sur le plan image. Les paramètres du mouvement dépendent de Z, h, α et du déplacement (d_x, d_y), en effet :

$$\begin{cases} a = \frac{d_x}{2Zh\alpha} \\ b = \frac{d_y}{2Zh\alpha} \\ c = \frac{\alpha h d_x}{2Z} \\ d = \frac{\alpha h d_y}{2Z} \end{cases} \tag{3.18}$$

Ce modèle dérivé d'une projection parabolique, représente un nouveau modèle de mouvement pour les images para-catadioptriques. Ce modèle remplacera le modèle constant utilisé par Lucas-Kanade pour contraindre l'équation du flot optique. Pour pouvoir estimer les paramètres du mouvement quadratique (3.17) on a besoin d'un voisinage. Le voisinage classique utilisé par Lucas-Kanade est inapproprié pour les images catadioptriques. Dans la suite, nous présentons le voisinage adapté aux images para-catadioptriques utilisé par Radgui et al. [Radgui 2009].

Voisinage adapté :

Dans le cas perspectif, la projection des points sur le plan image suit une loi affine, un voisinage sous forme d'une fenêtre rectangulaire est donc approprié. Cependant, dans le cas de l'image catadioptrique la fonction de projection liant les points 3D et leurs projections dans l'image est non-linéaire. Dans ce cas un voisinage classique est inadéquat. Certains auteurs ont proposé des méthodes d'adaptation du voisinage tenant compte

de la résolution non-uniforme, par exemple ; [Svoboda 2001] et [Demonceaux 2006]. Le voisinage adapté du point $p(u,v)$ noté $\vartheta_{u,v}$ est défini comme suit :

$$(u_i, v_i) \epsilon\ \vartheta_{u,v} \quad \text{si} \quad |\theta - \theta_i| < d\theta \text{ et } |\varphi - \varphi_i| < d\varphi \tag{3.19}$$

où $(1, \theta, \varphi)$ sont les coordonnées sphériques du point p (voir la figure 3.2).

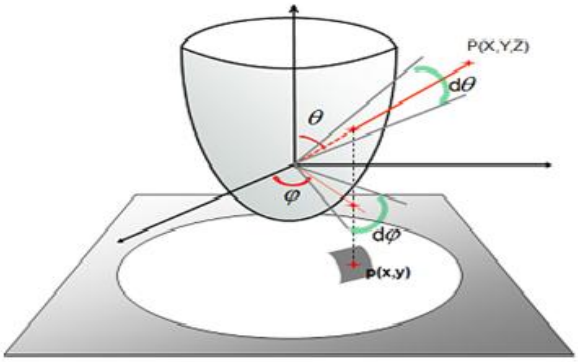


FIGURE 3.2 – Voisinage adapté au miroir parabolique projeté ensuite sur le plan image

Bien que la géométrie du voisinage adapté soit plus adaptée à la géométrie de l'image para-catadioptrique (voisinage plus grand dans les périphéries, plus petit au centre), le temps de calcul reste très élevé par rapport au « calcul » d'un voisinage classique.

Estimation du flot adapté sur les images para-catadioptriques :

Dans la suite de ce travail, le modèle de mouvement apparent (3.17) et le voisinage adapté (3.19) aux images para-catadioptriques seront utilisés pour calculer le mouvement apparent adapté d'un point $p(u,v)$. Calculer le flot adapté revient à estimer les quatre paramètres a, b, c et d du modèle. Ces paramètres sont obtenus par la méthode des moindres carrées :

$$\begin{aligned}
&(\widehat{\dot{u}}, \widehat{\dot{v}}) = \\
&\arg\min_{a,b,c,d} \textstyle\sum_{(u_i,v_i)\in\vartheta_{u,v}} \quad [\frac{\partial I}{\partial x}(a(u_i-u_0)^2 + a(v_i-v_0)^2 + c) \\
&\qquad\qquad\qquad\qquad\qquad\quad + \frac{\partial I}{\partial y}(b(u_i-u_0)^2 + b(v_i-v_0)^2 + d) \\
&\qquad\qquad\qquad\qquad\qquad\quad + \frac{\partial I}{\partial t}]^2
\end{aligned} \tag{3.20}$$

Remarque : comme les variables a, b, c et d sont corrélées deux à deux (eq. 3.18), on peut se contenter d'estimer seulement les deux paramètres a et b et déduire par la suite c et d par la formule suivante :

$$c = (\alpha h)^2 a \qquad d = (\alpha h)^2 b \tag{3.21}$$

Dans la section suivante nous donnons un bref état de l'art sur l'estimation du foyer d'expansion (FOE) qui résulte de la divergence du flot optique dans les images perspectives et aussi dans les images omnidirectionnelles. Une des caractéristiques du FOE est la possibilité de déduire la direction du mouvement du capteur dans l'image, ce qui s'avère intéressant à exploiter dans notre étude.

3.3 Estimation des foyers d'expansion et de contraction

3.3.1 Estimation de foyer d'expansion sur les images perspectives

Lors d'une translation de la caméra dans une scène statique, les directions de vitesse des points projetés sur le plan image converge vers un point du plan projectif appelé FOE (voir Fig. 3.3). Remarquons que, le flot optique a une grande amplitude dans la périphérie et une plus petite proche du FOE. Notons que, la propriété de divergence radiale nous donne suffisamment d'informations pour détecter la position du FOE dans le plan image, ce qui implique que, l'amplitude des vecteurs de flot peut être ignorée. En outre, il n'est pas nécessaire d'avoir le vecteur de flot optique exact en chaque point de l'image pour déterminer le FOE, une distribution judicieuse des vecteurs est suffisante [Sazbon 2004]. L'extraction du FOE se fait entre deux images en t et $t+1$ et joue un rôle très important dans différentes applications de la navigation visuelle. Le FOE est une information cruciale utilisée dans l'estimation du TAC [McCarthy 2008], la reconstruction 3D [Bak 2011] et par conséquent, dans le contrôle du mouvement et le système d'alerte de collision et l'évitement d'obstacles.

Les méthodes actuellement utilisées pour estimer le FOE entre deux images consécutives sont divisées en trois grandes classes. La première classe regroupe les méthodes dites discrètes, qui utilisent des caractéristiques correspondantes (points, lignes ou des courbes) [Jain 1983]. La deuxième classe contient les méthodes dites continues basées sur l'estimation explicite du flot optique [Sharma 1996], [Menéndez 1999], [Sazbon 2004] et la troisième classe regroupe les méthodes basées sur la contrainte de changement de luminosité [Negahdaripour 1989], [Joarder 1994].

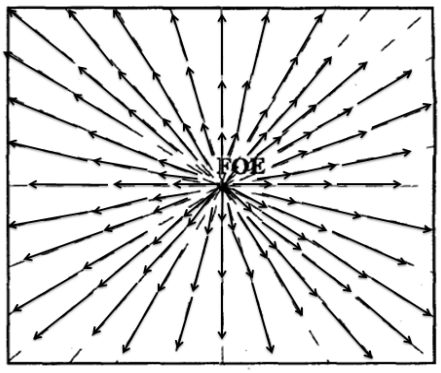


FIGURE 3.3 – Foyer d'expansion (FOE).

3.3.1.1 Méthodes discrètes

Soit une séquence d'image $I = [I_1, I_2, I_3....., I_N]$ de taille N acquise par la caméra en déplacement. Considérons deux images I_1 et I_2 avec n paires de points de correspondance (P_1, P_1').....(P_n, P_n'). La droite (P_1P_1') passe nécessairement de FOE. Pour déterminer le FOE on a besoin au minimum de deux paires de points correspondant, comme montré dans la figure 3.4[Jain 1983].

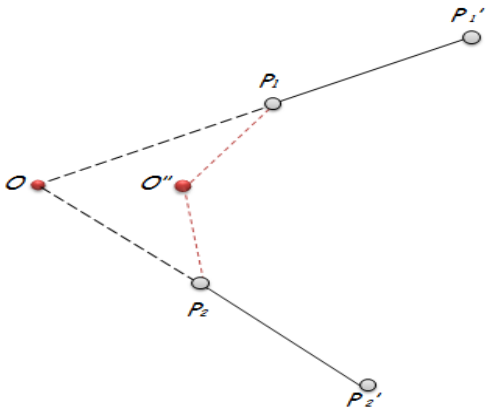


FIGURE 3.4 – Distances entre FOE et les points de correspondance et entre d'autres dans l'espace de recherche

Soit O' un point qui appartient à la droite (P_1P_1') et non au segment $[P_1P_1']$. Notons la distance euclidienne entre le point A et B par $L(A, B)$. On a :

$$L(P_1P_1') = |L(O'P_1') - L(O'P_1)|. \tag{3.22}$$

Pour un point O'' n'appartenant pas à la droite (P_1P_1'), l'inégalité triangulaire nous donne :

$$L(O''P_1') < L(O''P_1) + L(P_1P_1'). \tag{3.23}$$

donc

$$L(O''P_1') - L(O''P_1) < L(P_1P_1'). \tag{3.24}$$

de manière similaire avec un point O''' qui n'appartient pas à la ligne (P_2P_2'), on obtient :

$$L(O'''P_2') - L(O'''P_2) < L(P_2P_2'). \tag{3.25}$$

Remplaçant O' par O'' et O''' et sommant (3.24) et (3.25), on obtient :

$$L(O'P_1') - L(O'P_1) + L(O'P_2') - L(O'P_2) \leq L(P_1P_1') + L(P2P_2') \tag{3.26}$$

L'inégalité (3.26) atteint son maximum au point $O' = O$, c'est le point FOE. Ainsi, pour déterminer le FOE il suffit de maximiser la fonction suivante :

$$M(O') = [\sum_{k=1,n} L(O'P_k') - \sum_{k=1,n} L(O'P_k)] \tag{3.27}$$

Le principal inconvénient des méthodes discrètes réside dans le fait que la recherche des points de correspondance est une tâche difficile. Par ailleurs, ces méthodes ne sont pas robustes, car elles utilisent des informations locales de petites régions dans les images.

3.3.1.2 Méthodes continues

La deuxième classe des méthodes consiste à estimer le FOE en utilisant le flot optique. Ces techniques utilisent une information globale, elles sont donc plus robustes. L'une des méthodes de cette classe est basée sur l'usage de flot optique et d'un filtre adapté F. Ce filtre est de taille $(2w+1) \times (2w+1)$ et représente une grille cartésienne. Chaque point de la grille représente l'angle entre le point de la grille correspondant et l'origine (i.e. la direction radiale). Cela s'écrit :

$$F(m,n) = \arctan \frac{n}{m} \quad -w \leq m \leq w, -w \leq n \leq w \tag{3.28}$$

Le FOE est obtenu par la minimisation de la fonctionnelle S suivante :

$$(\widehat{x}_{FOE}, \widehat{y}_{FOE}) = \arg\min_{(x,y)} S(x,y) \tag{3.29}$$

Dans le cas où on peut utiliser l'équation (3.28) de manière idéale, $S(x, y)$ est la somme des carrés des différences entre F et les directions correspondantes de flot optique induites par un voisinage de taille $(2w + 1)^2$ pixels centré en (x, y). Puisque l'estimation de flot optique est imprécise, il est nécessaire d'introduire une méthode robuste pondérant le flot en chaque pixel en fonction de la précision de l'estimation. Par exemple, nous pouvons donner aux pixels périphériques les plus grands poids car ils ont une plus grande amplitude et sont donc moins sensibles au bruit. Par conséquent nous définissons $S(x, y)$ par :

$$S(x,y) = \Psi(u(x,y), v(x,y)) \sum_{m=-w}^{w} \sum_{n=-w}^{w} [(F(m,n) - \alpha(u(x+m, y+n), v(x+m, y+n)))^2$$

$$\Phi(u(x+m, y+n), v(x+m, y+n))] \tag{3.30}$$

les fonctions α, Ψ, Φ sont definies successivement par :

$$\alpha(u(x+m, y+n), v(x+m, y+n)) = \arctan \frac{v(x+m, y+n)}{u(x+m, y+n)}, \tag{3.31}$$

$$\Phi(u(x+m, y+n), v(x+m, y+n)) = \begin{cases} 1 & \text{si } u(x,y)^2 + v(x,y)^2 \geq t \\ 0 & \text{sinon.} \end{cases} \tag{3.32}$$

et

$$\Psi(u(x,y), v(x,y)) = [\sum_{m=-w}^{w} \sum_{n=-w}^{w} \Phi(u(x+m, y+n), v(x+m, y+n))]^{-1} \tag{3.33}$$

Le principal inconvénient des méthodes continues réside dans le temps d'exécution élevé, car elles essaient de résoudre un problème de minimisation des moindres carrées pour calculer le FOE sur un champ de vecteurs dense.

3.3.1.3 Méthode basée sur la contrainte du changement de l'intensité lumineuse

Une approche différente est utilisée pour l'estimation du FOE. Elle n'utilise ni le champ de flot optique ni les caractéristiques correspondantes, mais elle impose une contrainte sur la caméra qui doit se trouver en face de la scène observée (i.e, la profondeur est positive) [Negahdaripour 1989]. L'inconvénient de cette méthode est qu'elle fonctionne efficacement dans le cas d'une translation pure, mais pas lorsque le mouvement est combiné avec une rotation.

3.3.2 Estimation du foyer d'expansion et du foyer de contraction sur les images catadioptriques.

La structure topologique dense du flot optique omnidirectionnel permet d'acquérir des informations intéressantes. Le mouvement de translation pure produit un champ caractérisé par deux points singuliers, foyer d'expansion FOE et le foyer de contraction FOC, situés sur les pôles opposés de l'image sphérique. Le vecteur passant par ces deux points singuliers indique la direction de l'ego-mouvement du système caméra-robot. La rotation autour d'un axe génère un champ de mouvement circulaire autour de cet axe et elle change les deux positions de FOE et FOC dans la carte des vecteurs de flot optique de l'image omnidirectionnelle. Dans la figure 3.5, on visualise les vecteurs de flot optique d'un capteur omnidirectionnel qui subit un mouvement rigide dans un environement statique. Dans le cas d'un mouvement de translation pure (voir la Fig. 3.5.a), les deux positions de FOE et FOC sont opposées s'étendant sur un angle de 180°. Alors qu'un mouvement purement rotationnel (voir Fig. 3.5.b) provoque la disparition de ces deux points [Stratmann 2004].
Le mouvement du capteur peut être décrit en terme de trois composantes de translation $T = (T_x, T_y, T_z)$ et trois composantes de rotation $\omega = (\omega_x, \omega_y, \omega_z)$.
La vitesse angulaire $F(\Theta)$ du capteur en translation avec le vecteur $\overrightarrow{T}$ de direction Λ et en rotation avec le vecteur $\overrightarrow{\omega}$ a une répartition sinusoïdale de type :

$$F(\Theta) = A.\sin(\Theta - \Lambda) + \omega \tag{3.34}$$

avec $A = -\frac{||\overrightarrow{T}||}{r}$, les angles Θ et Λ sont illustrés dans la figure 3.6.
L'amplitude A de cette distribution sinusoïdale est proportionnelle à l'amplitude de la vitesse de translation et la distance r entre le point 3D de la scène et le centre de projection. La déviation de la phase du sinus Λ, indique le sens de la translation et les points singuliers de la distribution donnent le FOE et FOC (voir la Fig. 3.7).
Nous verrons dans le chapitre 4 une méthode basée sur l'analyse de la distribution de l'orientation du flot optique pour extraire les foyers FOE et FOC en utilisant le capteur catadioptrique.
L'estimation du flot optique et le foyer d'expansion sont deux outils importants pour l'estimation du TAC. Nous introduisons dans la section suivante un état de l'art des travaux les plus utilisées pour l'estimation du TAC qui existent dans la littérature.

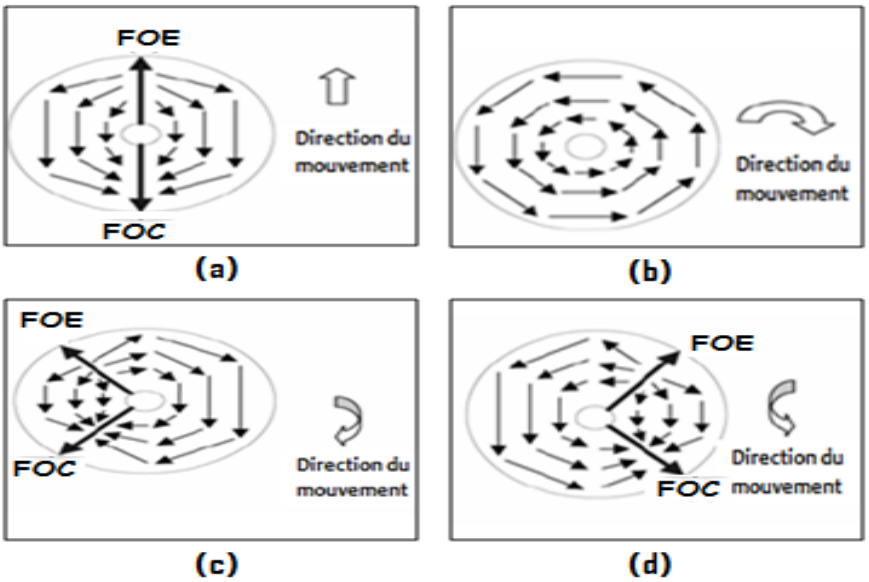


FIGURE 3.5 – Les positions des FOE et FOC. (a) : Positions de FOE et FOC dans le cas d'une translation pure. (b) : Disparitions de FOE et FOC dans le cas d'une rotation pure. (c) : Positions de FOE et FOC dans le cas d'un mouvement combiné d'une translation et une rotation, la translation dans ce cas est de droite à gauche. (d) : Positions de FOE et FOC dans le cas d'un mouvement combiné d'une translation et une rotation, la translation dans ce cas est de gauche à droite.

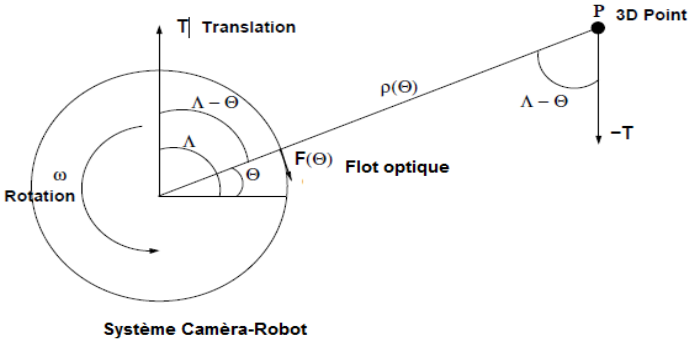


FIGURE 3.6 – Mouvement sphérique du système robot-caméra [Stratmann 2004].

3.4 Méthodes d'estimation du TAC

Dans cette section, nous exposons, dans un premier temps, les méthodes les plus connues dans la littérature pour l'estimation du TAC sur les images perspectives. Puis, nous présentons les méthodes proposées dans le cas omnidirectionnel.

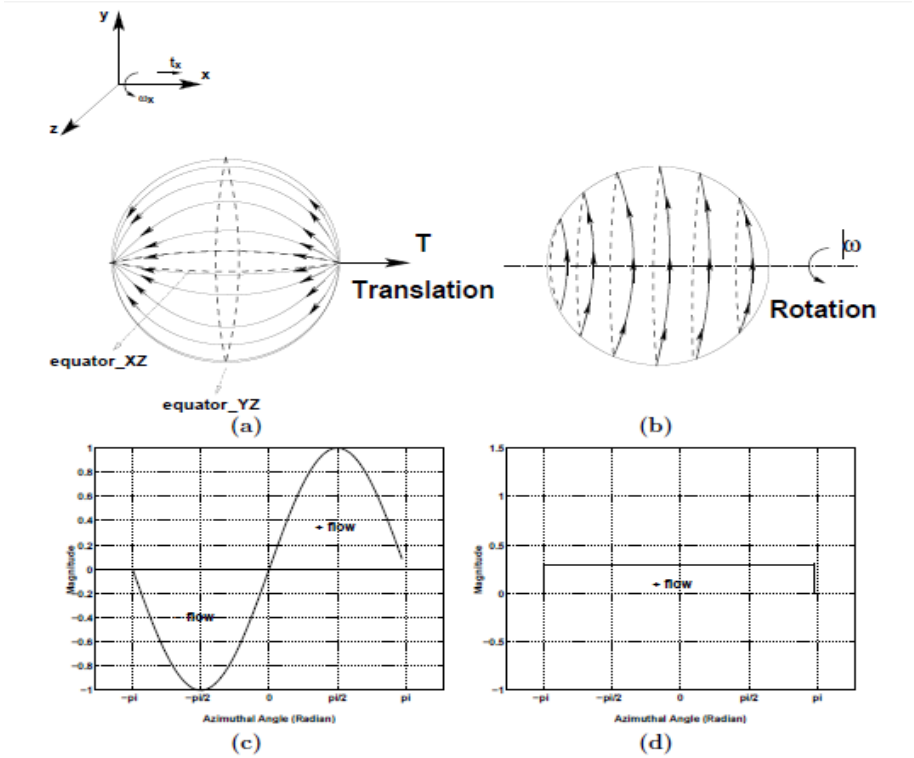


FIGURE 3.7 – Le champ de mouvement sphérique. (a) La sphère est en translation selon l'axe **X** dans la scène 3D. (b) La rotation de la sphère autour de l'un de ses axes produit une distribution uniforme du flot optique. (c) Distribution sinusoïdale du flot sur l'équateur XZ. (d) La distribution du flot sur l'équateur XZ est constante et proportionnelle à la vitesse de la rotation ([Stratmann 2004]).

3.4.1 Méthode d'estimation du temps à collision dans le cas perspectif

Pour les caméras perspectives, un certain nombre d'approches pour l'estimation du TAC ont été proposées. Nous classons ces méthodes en quatre catégories : les méthodes basées sur le flot optique, les méthodes basées sur les gradients d'image, les méthodes basées sur les contours fermés et les méthodes basées sur la représentation log-polaire.

3.4.1.1 Méthodes basées sur l'estimation explicite du flot optique

Nous exposons tout d'abord les méthodes basées sur les invariants différentiels du flot optique puis les méthodes basées sur l'estimation directe du flot optique.

Méthodes basées sur les invariants différentiels du flot optique rigide

En 1990, Subbarao [Subbarao 1990] a été le premier à déterminer une expression théorique explicite des bornes supérieure et inférieure du TAC en utilisant les invariants différentiels du flot optique.

Les invariants différentiels *div* (divergence), *curl* (vorticité) et *def* (déformation) ont été introduits à partir de la décomposition du flot optique local relativement à une surface plane rigide et ils sont définis en fonction des dérivées partielles du flot optique. En effet, afin de profiter des propriétés du flot optique, il est utile de le représenter en fonction de ses dérivées partielles. Soit P de coordonnées (X, Y, Z) appartenant à une surface rigide (S) de vecteur normal $(-1, a, b)$ et qui se projette sur le plan image en $p = (x, y)$. Soit $T = (T_x, T_y, T_z)$ la translation instantanée du point P et $\Omega = (\omega_x, \omega_y, \omega_z)$ sa rotation instantanée autour du centre du répère monde $\mathbb{R}^3$ (voir la figure 3.8).

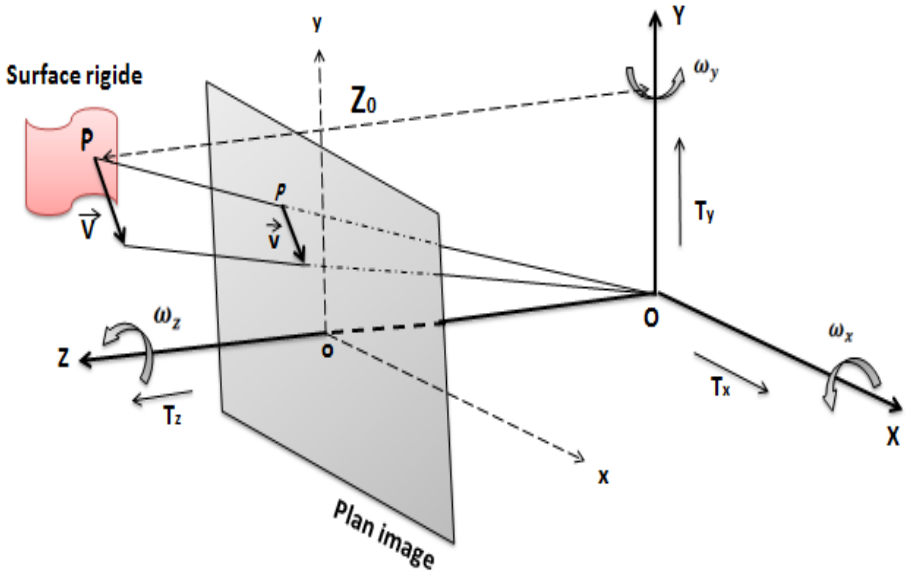


FIGURE 3.8 – Représentation géométrique de la translation d'une surface rigide.

[Longuet-Higgins 1980] a montré que dans ce cas, le flot optique pourrait s'approximer par :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} + \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{3.35}$$

avec

$$\begin{cases} u_0 = -\frac{T_x}{Z_0} - \omega_y, & v_0 = -\frac{T_y}{Z_0} + \omega_x, \\ u_x = \frac{T_z}{Z_0} + a\frac{T_x}{Z_0}, & u_y = \omega_z + b\frac{T_x}{Z_0}, \\ v_x = -\omega_z + a\frac{T_y}{Z_0}, & v_y = \frac{T_z}{Z_0} + b\frac{T_y}{Z_0}. \end{cases} \tag{3.36}$$

La matrice 2×2 dans l'équation (3.35) définit le tenseur du gradient de la vitesse. C'est dans la décomposition de cette matrice que les invariants différentiels locaux peuvent provenir. Koenderink et al. [Koenderink 1975] ont appliqué le théorème de la décomposition de Cauchy-Stokes [Aris 1965] qui stipule que toute matrice 2×2 peut être décomposée en une somme de deux matrices, une symétrique et l'autre antisymétrique, telle que :

$$\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = 1/2 \begin{pmatrix} 0 & u_y - v_x \\ -u_y + v_x & 0 \end{pmatrix} + 1/2 \begin{pmatrix} 2u_x & u_y + v_x \\ u_y + v_x & 2v_y \end{pmatrix} \tag{3.37}$$

La matrice symétrique peut être décomposée en la somme de multiple d'une matrice identité, I et d'une matrice symétrique de trace nulle, de telle sorte que :

$$\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = 1/2 \begin{pmatrix} 0 & u_y - v_x \\ -u_y + v_x & 0 \end{pmatrix} + 1/2 \begin{pmatrix} u_x + v_y & 0 \\ 0 & u_x + v_y \end{pmatrix} + 1/2 \begin{pmatrix} u_x - v_y & u_y + v_x \\ u_y + v_x & -u_x + v_y \end{pmatrix} \tag{3.38}$$

Cette décomposition nous permet de factoriser chaque matrice de (3.38) comme suit :

$$\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \frac{curl}{2} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} + \frac{div}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{def}{2} S \tag{3.39}$$

avec

$$div = u_x + v_y \tag{3.40}$$

$$curl = -u_y + v_x \tag{3.41}$$

$$def = \sqrt{(u_y + v_x)^2 + (u_x - v_y)^2} \tag{3.42}$$

S représente une matrice symétrique de trace nulle et de déterminant -1, tel que :

$$S = Q^{-1} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} Q \tag{3.43}$$

avec Q est la matrice de rotation définie comme suit :

$$\begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix} \tag{3.44}$$

et θ est l'angle de rotation (voir article [Cipolla 1997] page 17).
L'intérêt de cette décomposition est que, localement, le flot optique peut être caractérisé d'une manière indépendante de coordonnées images, en utilisant seulement les invariants différentiels.
En redéfinissant les composantes parallèles de translation comme suit :

$$\frac{T_x}{Z_0} = m\cos(\psi) \qquad \frac{T_y}{Z_0} = m\sin(\psi) \tag{3.45}$$

où m est l'amplitude du vecteur de la translation paralèlle $1/Z_0(T_x, T_y)$ et ψ est la direction de la translation parallèle au plan image.
Le vecteur normal de la surface plane peut être reécrit comme suit :

$$a = k\cos(\phi) \qquad b = k\sin(\phi) \tag{3.46}$$

avec ϕ, l'angle de sa direction et k son amplitude. Par suite, en utilisant les équations (3.36), (3.40), (3.42), (3.45) et (3.46), [Subbarao 1990] a défini les bornes maximum et minimum du TAC (τ) par la formule suivante :

$$1/2(div - def) \leq \tau^{-1} \leq 1/2(div + def). \tag{3.47}$$

Sans connaître la direction ψ du vecteur normal (a, b) et la direction ϕ de la translation parallèle, il n'est pas possible de calculer le TAC précisément. Les invariants différentiels sont exprimés en fonction des dérivées partielles du flot optique (3.40) et (3.42). Les bornes du TAC sont définies directement à partir du calcul du flot optique dans le cas d'un mouvement rigide le long de l'axe optique **Z** (direction de vue). La généralisation de l'expression (3.47) pour toute direction de vue est présentée dans la sous-section 3.4.2.

[Meyer 1994] et [Meyer 1992] propose une méthode de calcul des cartes du TAC basée sur le mouvement rigide et inspirée de la méthode de [Subbarao 1990]. Cette méthode est basée sur le modèle de mouvement quadratique de [Longuet-Higgins 1980]. Ainsi, à partir de l'équation de la surface plane, une expression est déduite en fonction des paramètres

d'orientation de la surface et les termes du premier ordre du champ de mouvement (3.36). Meyer fait remarquer que les paramètres d'orientations de la surface plane ne peuvent pas être récupérés uniquement par la connaissance du mouvement de translation de la caméra. Ainsi, pour obtenir les coefficients de premier ordre du champ de mouvement, un schéma de multi-résolution est utilisé pour affiner l'estimation des six paramètres, en supposant que le mouvement est linéaire dans une région locale, ce qui fournit un système d'équations linéaires sur-determiné qui peut être résolu en utilisant les moindres carrés. Cette méthode suppose que la composante verticale du mouvement 3D est nulle. Néanmoins, cette hypothèse est valable seulement quand l'axe optique de la caméra est parallèle au plan du sol.

En 2008, [McCarthy 2008] a montré qu'on peut calculer précisément le TAC en fonction de la divergence du flot optique div (3.40) pour les surfaces fronto-parallèles (ou à orientation connue). D'autre part, et contrairement à [Subbarao 1990] qui ne tient pas compte de l'effet des rotations instantanées pendant l'ego-motion et suppose que les points d'intérêt sont situés le long de l'axe optique de la caméra, [McCarthy 2008] a prouvé qu'on peut contourner ces restrictions en utilisant le suivi du FOE au cours du mouvement. En effet, compte tenu d'une caméra embarquée sur un robot, on peut supposer que la vitesse de translation T_c de la caméra, est proportionnelle à la vitesse de translation T_r du robot, c'est-à-dire que :

$$T_c = [\alpha T_r \quad \beta T_r \quad \gamma T_r] \tag{3.48}$$

et le vecteur de rotation de la caméra est donné par :

$$\omega_c = [\omega_x \quad \omega_y \quad \omega_z] \tag{3.49}$$

Le TAC est défini en fonction de la div par l'expression suivante [Coombs 1998] :

$$\tau = \frac{2}{div} = \frac{2}{u_x + v_y} \tag{3.50}$$

L'idée de [McCarthy 2008] est semblable à celle de [Subbarao 1990]. Elle consiste à exprimer le flot optique en utilisant la fonction de profondeur de la surface plane. Ce flot optique est donné par [Santos-Victor 1994] :

$$u(x,y) = f_x[\frac{\gamma T_r(\frac{x}{f_x} - \alpha)}{Z(x,y)} + \omega_x \frac{xy}{f_x f_y} - \omega_y(1 + \frac{x^2}{f_x^2}) + w_z \frac{y}{f_y}] \tag{3.51}$$

$$v(x,y) = f_y[\frac{\gamma T_r(\frac{y}{f_y} - \beta)}{Z(x,y)} + \omega_x(1 + \frac{y^2}{f_y^2}) - \omega_y \frac{xy}{f_x f_y} - w_z \frac{x}{f_x}] \tag{3.52}$$

avec f_x et f_y sont les facteurs de la focale en coordonnées pixelliques.
Alors, la fonction de la surface plane $Z(x, y)$ en fonction des coordonnées images et en fonction du point FOE (x', y') est représentée par :

$$Z(x, y) = \frac{Z(x', y')}{1 - a\frac{(x-x')}{f_x} - b\frac{(y-y')}{f_y}} \tag{3.53}$$

En remplaçant (3.53) dans (3.51) et (3.52), on obtient :

$$u(x, y) = \frac{\gamma T_r(x - f_x\alpha)}{Z(x', y')}[1 - \frac{a(x - x')}{f_x} - \frac{b(y - y')}{f_y}] + \omega_x\frac{xy}{f_y} - \omega_y(f_x + \frac{x^2}{f_x}) + w_z\frac{y}{f_x}] \tag{3.54}$$

$$v(x, y) = \frac{\gamma T_r(y - f_y\alpha)}{Z(x', y')}[1 - \frac{a(x - x')}{f_x} - \frac{b(y - y')}{f_y}] + \omega_x(f_y + \frac{y^2}{f_y}) - \omega_y\frac{xy}{f_x} - w_z\frac{x}{f_x}] \tag{3.55}$$

En utilisant la proprieté que le flot optique s'annule au point FOE, on peut exprimer les composantes de la rotation (ω_x, ω_y) en fonction des composantes de la translation T_c, de la composante de rotation ω_z, de la fonction $Z(x', y')$, des facteurs de la focale (f_x, f_y) et des coordonnées image. En remplaçant les composantes (ω_x, ω_y) dans les expressions des dérivées partielles de $u(x, y)$ et $v(x, y)$, il en résulte l'équation générale du TAC au point FOE suivante :

$$\tau = -\frac{\gamma}{div}[1 + a(\frac{x'}{f_x} + \alpha) + b(\frac{y'}{f_y} + \beta) - \frac{3}{\gamma x'}(-f_x\alpha + \frac{x'f_y + f_xf_y\alpha + x'y'\beta + \frac{y'^2 f_x\alpha}{f_y}}{f_y(1 + \frac{x'^2}{f_x^2} + \frac{y'^2}{f_y^2})} +$$

$$\frac{\omega_z y' T_r}{f_y Z(x', y')}(f_y + \frac{y'^2}{f_y} - \frac{x'^2}{f_x} - 1))] \tag{3.56}$$

L'équation (3.56) calcule le TAC même dans le cas où l'axe optique et l'axe du mouvement du robot ne sont pas alignés. Par contre, elle exige la connaissance des paramètres de l'orientation de la surface plane a et b et les paramètres de translation $\alpha = \beta = 0$, $\gamma = 1$ et $\omega_z = 0$.
Dans la méthode que nous allons présenter par la suite, l'estimation du TAC ne nécessite pas la connaissance des paramètres d'orientation de la surface et reste valable pour tout obstacle orienté de façon quelconque.

Méthodes basées sur l'utilisation directe du flot optique

La méthode basée sur le flot général [Camus 1995] a l'avantage de fournir une valeur du TAC pour chaque pixel, c'est-à-dire un TAC local et qui n'impose pas l'hypothèse que le mouvement doit être rigide comme les méthodes basées sur les invariants différentiels. Le calcul du TAC se fait sur deux images successives de la séquence. Afin de rappeler l'expression du TAC avec cette méthode, nous considérons le schéma de la figure 3.9. Soit P un point $3D$ fixe de coordonnées (X, Y, Z) et $p(x, y, f)$ sa projection sur le plan image. Le centre de projection O de coordonnées $(0, 0, 0)$ se déplace avec une vitesse $\frac{dZ}{dt}$ vers le point P selon l'axe **Z**.

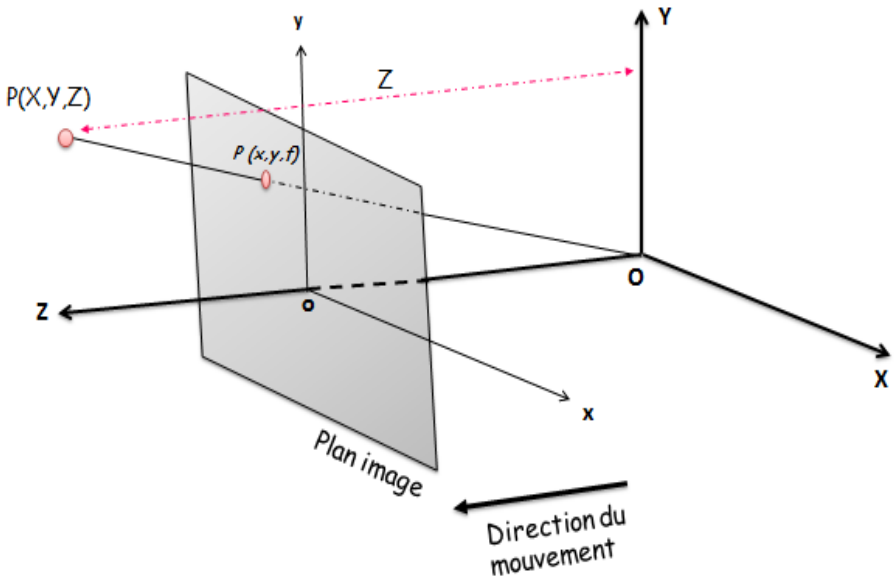


FIGURE 3.9 – Configuration géométrique dans le cas perspectif.

Le plan image est fixé à une distance f de l'origine C, par convention on prend $f = 1$. Rappelons l'équation de projection perspective (3.57) :

$$p = p_r(P) \quad \text{avec} \quad \begin{cases} x = f\frac{X}{Z} \\ y = f\frac{Y}{Z} \end{cases} \tag{3.57}$$

où (X, Y, Z) sont les coordonnées du point réel $3D$ et (x, y) représentent les coordonnées images dans le plan image.

En dérivant l'équation (3.57) par rapport au temps, on trouve l'expression du flot optique suivant l'axe y, donnée par :

$$\dot{y} = \frac{\dot{Y}}{Z} - Y(\frac{\dot{Z}}{Z^2}) \tag{3.58}$$

Etant donné que le mouvement se fait uniquement suivant l'axe Z, on a $\dot{Y} = 0$ et $\dot{X} = 0$. Ainsi, à partir des équations (3.57) et (3.58), on obtient :

$$\frac{y}{\dot{y}} = -\frac{Z}{\dot{Z}} = \tau \tag{3.59}$$

De manière similaire on peut retrouver l'expression du TAC suivant l'axe x. Notons que la quantité τ permet de calculer le TAC, dans le cas où l'axe optique est parallèle au sens du mouvement de la caméra perspective. Pour calculer le TAC local à chaque pixel de l'image, il faut estimer le mouvement apparent en ce point, et pour cela, [Camus 1995] a choisi d'utiliser le flot optique de [Lucas 1981]. La figure 3.10 montre un exemple des cartes de TAC obtenues en appliquant l'équation (3.59) où le flot optique a été calculé par [Lucas 1981]. La séquence perspective utilisée représente le déplacement de la caméra en mouvement de zoom vers un obstacle (la cathédrale).

3.4.1.2 Méthodes basées sur les contours fermés

Le TAC peut être estimé en utilisant des techniques basées sur le théorème de Green [Kaplan 1991], qui fournit un moyen d'estimer la divergence d'une région sans le calcul explicite du flot optique.

Considérons un patch d'une surface, $\mathbb{S} \in \mathbb{R}^3$, dont l'image est délimitée par un contour fermé, $C \in \mathbb{R}^2$ (voir Fig. 3.11), et avec un flot optique, $\overrightarrow{U} \in \mathbb{R}^2$, défini continument sur $\mathbb{S}$, le théorème de Green nous dit que l'intégrale des vecteurs de flot le long de C dans la direction de la normale à C, est égale à l'intégrale de la divergence des vecteurs de flot définis sur $\mathbb{S}$. Autrement dit, la divergence moyenne d'une surface $\mathbb{S}$ peut être obtenu en additionnant les vecteurs normaux le long de C. Cela peut être exprimé comme :

$$\int\int_{\mathbb{S}} div\overrightarrow{U}\, ds = \oint_C \overrightarrow{U}.\overrightarrow{n}\, dl \tag{3.60}$$

Où $\overrightarrow{n}$ est le vecteur normal de C, ds est un élément de la surface $\mathbb{S}$ et dl est un élément de C [Duric 1999]. Le TAC moyen de la surface $\mathbb{S}$ est l'inverse de la sommation (3.60).

Di Marco et al. [Di Marco 2003] proposent une approche basée sur la théorie des ensembles pour estimer le TAC en se basant sur les contours fermés. Le changement temporel des contours de la région est suivi par l'application d'un filtre récursif. A chaque intervalle de temps, le filtre calcule un ensemble de vecteurs d'état (c'est-à-dire les paramètres de la transformation affine et leurs dérivées partielles) conformes avec des mesures courantes

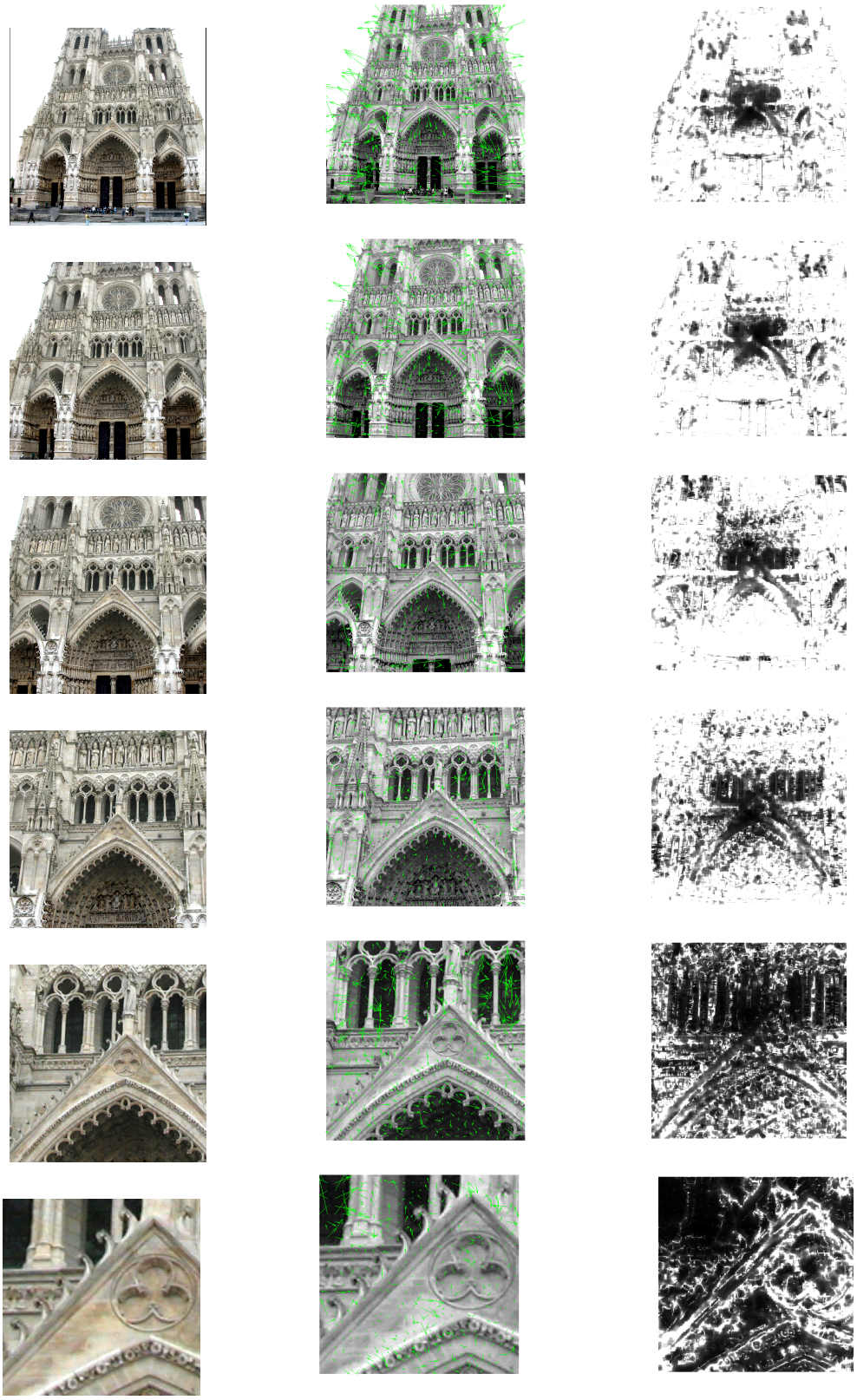


FIGURE 3.10 – Cartes des TACs de la séquence "Cathédrale" obtenues en utilisant la méthode de [Camus 1995]. 1ere colonne : Evolution de la séquence dans le temps. 2ème colonne : Flot optique de [Lucas 1981]. 3ème colonne : Les cartes du TAC. gris foncé : TAC faible (proche de l'obstacle), gris clair :TAC élevé (loin de l'obstacle) et le blanc : valeur aberrante

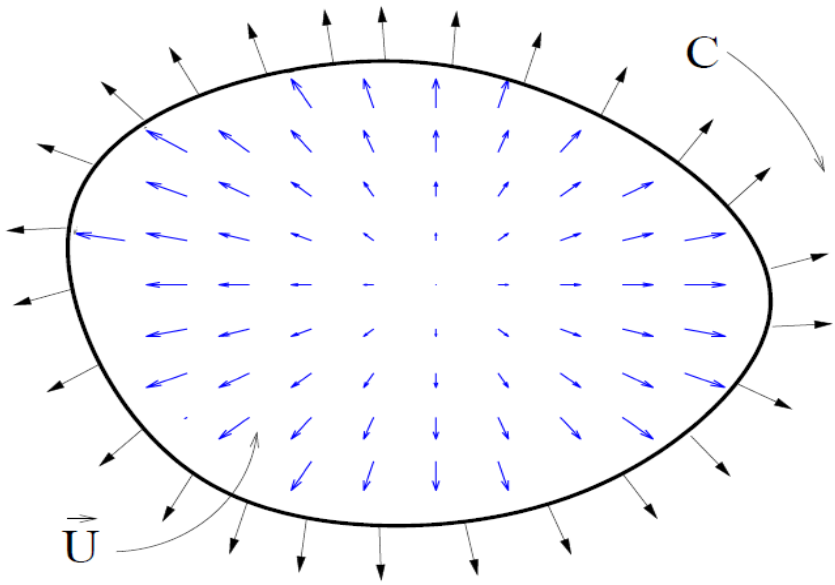


FIGURE 3.11 – Théoreme de Green appliqué pour calculer la divergence (*div*).

et des approximations calculées dans l'état précédent. La propagation des vecteurs d'état donne une indication directe de l'incertitude dans l'estimation. Les résultats montrent une certaine sensibilité aux valeurs initiales. Cependant l'algorithme est capable de détecter si le TAC estimé dépasse les bornes des limites de l'erreur de l'estimation. Les méthodes d'estimation du TAC basées sur les contours fermés sont généralement motivées par le désir d'éviter le calcul explicite du flot optique. En outre, il est important de signaler que le théorème de Green fournit seulement une approximation affine d'une surface projetée dans une petite région. Les modèles affines présentent une approximation inadéquate des surfaces non-planes. Ces modèles sont susceptibles de ne pas fonctionner en particulier si n'est pas fronto-paralèlle avec le plan image. Un autre problème aussi à noter, est la difficulté de trouver un contour à suivre si la surface est très proche et si le champ de vue (FOV) est restreint, les régions de contours fermés sont susceptibles de croître en dehors du plan image.

3.4.1.3 Méthodes basées sur la représentation log-polaire

Dans le système visuel humain les récepteurs de la rétine sont répartis d'une façon non uniforme. Ce sont les cônes, responsables de la vision. Ces cônes sont placés avec une densité élevée proche de la fovea et avec une densité basse dans la périphérie comme

representé dans la Fig. 3.12.

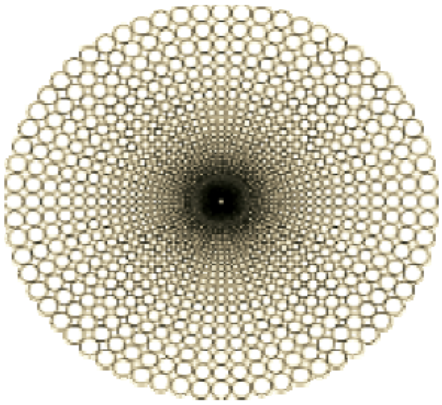


FIGURE 3.12 – Arrangement topologique log-polaire

Cette topologie présente une distribution radiale qui peut être simulée par une représentation polaire. Le système Log-polaire qui imite le fonctionnement de l'oeil, permet de réduire les données de traitement sans perdre leur signification [Sandini 1980]. Mathématiquement, le changement de coordonnées conformément à la transformation log-polaire peut être effectué en deux étapes. La première étape consiste à transformer le plan cartésien (x, y) (plan image) en plan polaire (ρ, θ) selon les équations suivantes (voir la Fig. 3.13) :

$$\rho = \sqrt{(x - x_o)^2 + (y - y_o)^2} \tag{3.61}$$

et

$$\theta = \tan^{-1}(\frac{y - y_o}{x - x_o}) \tag{3.62}$$

Dans la seconde étape, la transformation log-polaire est calculée à partir de la transformation polaire, selon les équations (3.63) et (3.64).

$$\eta = q\theta \tag{3.63}$$

et

$$\xi = log_a(\rho) - p \tag{3.64}$$

avec a, q et p sont des constantes déterminées par le dispositif matériel du capteur CCD. Tistarelli et Sandini [Tistarelli 1992] affirment que le flot optique estimé avec la représentation log-polaire est plus intéressant pour l'estimation du TTC. Plus précisément, ils montrent

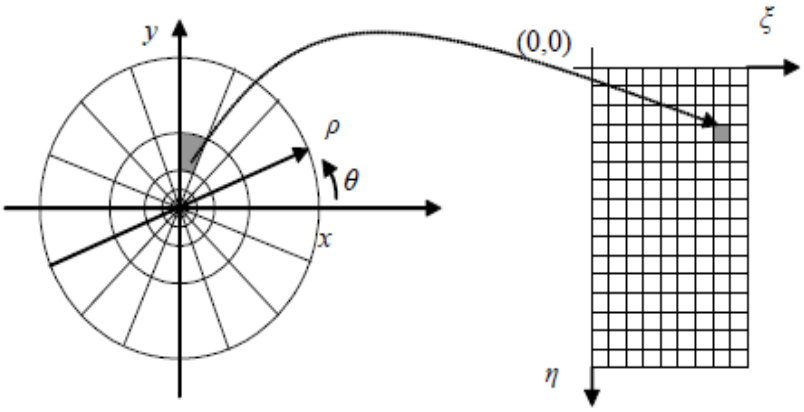


FIGURE 3.13 – Transformation log-polaire.

que l'estimation du TAC peut être obtenue explicitement et non pas seulement ses bornes comme dans le cas de [Subbarao 1990]. Ceci en exprimant les dérivées du flot optique dans la direction radiale. Il est démontré que seule la composante radiale du flot est liée au TAC. Notamment, cette méthode suppose que le gradient de la surface est locallement nul, ce qui n'est plus valable pour les surfaces inclinées [Colombo 2000].

3.4.1.4 Méthode basée sur les gradients de l'image

La méthode basée sur les gradients de l'image est introduite par Horn et al. [Horn 2007, Horn 2009]. Son objectif est d'estimer le TAC en évitant l'estimation explicite du flot optique, c'est une méthode simple qui ne nécessite aucun traitement de haut niveau comme le suivi ou la détection et est exprimée seulement en fonction des gradients de l'image. Supposons que la caméra se déplace vers une surface plane, quatre types de translations sont envisagées par la méthode de Horn et al. [Horn 2009] :

— Cas 1 : Mouvement de translation de la caméra le long de l'axe optique relativement à une surface plane perpendiculaire.

— Cas 2 : Mouvement de translation arbitraire de la caméra dont l'axe optique est perpendiculaire à la surface plane.

— Cas 3 : Mouvement de translation le long de l'axe optique avec une orientation arbitraire de la surface plane.

— Cas 4 : Mouvement de translation de la caméra et orientation de la surface plane

arbitraires.

Il est clair que le cas (1) est un cas particulier des deux cas (2) et (3) et que le cas (4) représente le cas général qui résume les trois autres. Par convention, prenons le repère caméra d'origine C et d'axes $\mathbf{X}, \mathbf{Y}$ et $\mathbf{Z}$ comme axe optique de la caméra. L'équation de la projection perspective utilisée est la même que l'équation (3.57). Posons $\nu = (\tilde{u}, \tilde{v})$ le flot optique et $(U, V, W) = (\dot{X}, \dot{Y}, \dot{Z})$ la vitesse relative du point réel P de la surface plane. Différentions l'équation (3.57) par rapport au temps :

$$\dot{m} = \dot{p_r}(P) \quad \text{avec} \quad \begin{cases} \tilde{u}=f(\frac{U}{Z} - \frac{x}{f}\frac{W}{Z}) \\ \tilde{v}=f(\frac{V}{Z} - \frac{y}{f}\frac{W}{Z}) \end{cases} \tag{3.65}$$

Ce qui donne :

$$\dot{m} = \dot{p_r}(P) \quad \text{avec} \quad \begin{cases} \tilde{u}=-\frac{W}{Z}(x - x_0) \\ \tilde{v}=-\frac{W}{Z}(y - y_0) \end{cases} \tag{3.66}$$

où (x_0, y_0) sont les coordonnées du FOE exprimées par :

$$\begin{cases} x_0=f\frac{U}{W} \\ y_0=f\frac{V}{W} \end{cases} \tag{3.67}$$

Dans ce qui suit, nous présentons les modèles d'estimation du TAC proposés par Horn et al. [Horn 2007] pour chaque cas.

Cas 1 : Mouvement de translation de la caméra le long de l'axe optique relativement à une surface plane perpendiculaire

Dans ce cas, l'axe optique est aligné avec le sens de translation de la caméra comme indiqué dans la figure 3.14 et on a $U = V = 0$.

D'aprés (3.65), le flot optique s'écrit alors :

$$\tilde{u} = -\frac{1}{Z}(xW) \qquad et \qquad \tilde{v} = -\frac{1}{Z}(yW) \tag{3.68}$$

remplaçons les équations (3.68) dans l'équation du flot optique (3.4). Nous trouvons :

$$-\frac{W}{Z}(xI_x + yI_y) + I_t = 0 \tag{3.69}$$

ce qui donne aussi :

$$CG + I_t = 0 \tag{3.70}$$

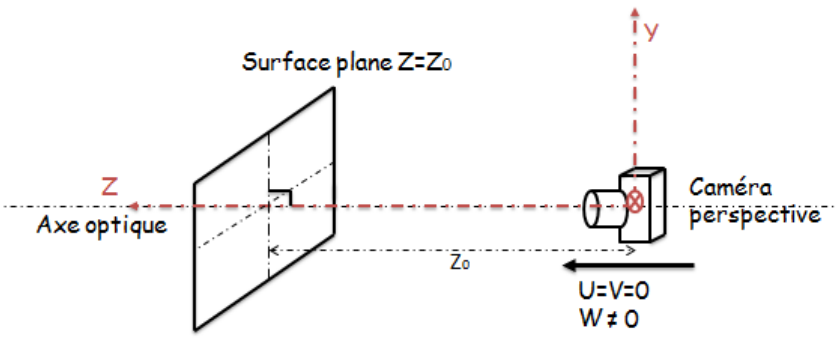


FIGURE 3.14 – Cas 1 : Translation de la caméra suivant l'axe optique relativement à une surface plane perpendiculaire.

avec $C = -\frac{W}{Z}$ l'inverse du TAC et $G = xI_x + yI_y$ le gradient radial.
Ainsi, résoudre le problème d'estimation du TAC revient à résoudre le problème de minimisation par les moindres carrés suivant :

$$min_C \sum_{(x,y)\in\Delta} (CG + I_t)^2 = 0 \tag{3.71}$$

où (Δ) est la région d'intérêt (elle peut être toute l'image) sur laquelle on souhaite calculer le TAC. Dérivons l'équation (3.71) par rapport à C, on trouve :

$$C = -\frac{\sum_{(x,y)\in\Delta} GI_t}{\sum_{(x,y)\in\Delta} G^2} \tag{3.72}$$

Par conséquent, calculer le TAC revient à calculer seulement le gradient radial et le gradient temporel.

Cas 2 : Mouvement de translation arbitraire de la caméra dont l'axe optique est perpendiculaire à la surface plane

Dans ce cas le mouvement de translation de la caméra n'est pas dans la direction de l'axe optique (voir la Fig. 3.15), c'est-à-dire : $U \neq 0, V \neq 0$ et $W \neq 0$.
Remplaçons les valeurs de $\tilde{u}$ et $\tilde{v}$ par leur expressions (3.66) dans l'équation ECMA (3.4), on trouve :

$$\frac{fU - xW}{Z} I_x + \frac{fV - yW}{Z} I_y + I_t = 0 \tag{3.73}$$

Posons $A = f\frac{U}{Z}$, $B = f\frac{W}{Z}$, $C = -\frac{W}{Z}$. Donc :

$$AI_x + BI_y + CG + I_t = 0 \tag{3.74}$$

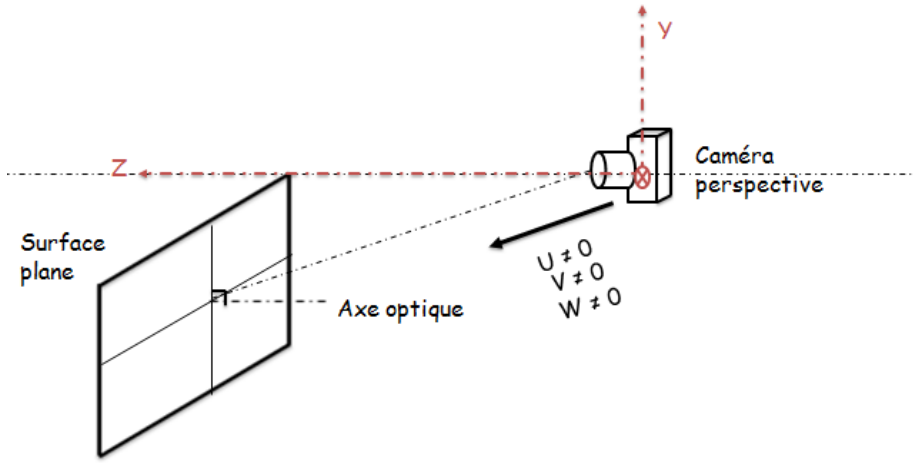


FIGURE 3.15 – Cas 2 : Translation arbitraire de la caméra relativement à une surface plane perpendiculaire.

La solution du système au sens des moindres carrés est :

$$min_{A,B,C} \sum_{(x,y)\in\Delta} (AI_x + BI_y + CG + I_t)^2 = 0 \tag{3.75}$$

Notons, si $U = V = 0$, c'est-à-dire, si le mouvement de translation de la caméra est aligné avec l'axe optique, le problème des moindres carrés se simplifie et on retrouve l'équation du premier cas (3.72) fonction seulement du terme C.

Cas 3 : Mouvement de translation le long de l'axe optique avec orientation arbitraire de la surface plane

Dans ce cas, le mouvement de la caméra est aligné avec l'axe optique, c'est-à-dire $U = 0, V = 0$ et $W \neq 0$, mais la surface plane n'est pas perpendiculaire à l'axe optique comme illustré dans la figure 3.16.

Soit l'équation de la surface plane :

$$Z = Z_0 + aX + bY \tag{3.76}$$

En utilisant le modèle de projection perspectif dans l'équation (3.76), nous trouvons :

$$Z(1 - a\frac{x}{f} - b\frac{y}{f}) = Z_0 \tag{3.77}$$

Remplaçons les valeurs de $\tilde{u}$ et $\tilde{v}$ dans l'équation ECMA (3.4),

$$-G\frac{W}{Z_0}(1 - a\frac{x}{f} - b\frac{y}{f}) + I_t = 0 \tag{3.78}$$

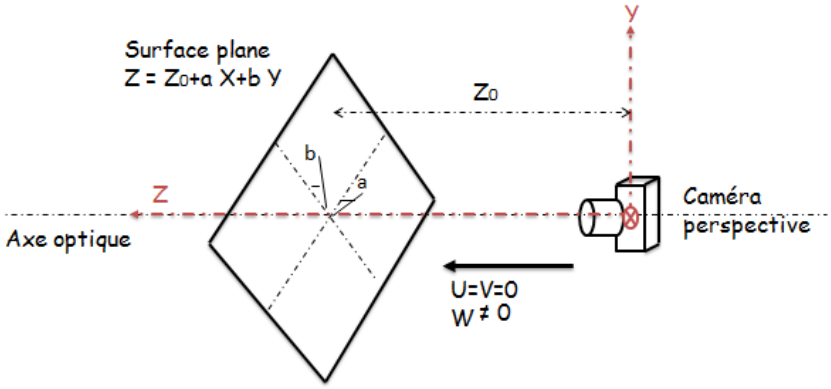


FIGURE 3.16 – Cas 3 : Translation de la caméra suivant l'axe optique relativement à une surface plane d'orientation quelconque.

par suite,

$$G(C + Px + Qy) + I_t = 0 \tag{3.79}$$

avec $P = (\frac{a}{f})\frac{W}{Z_0}$ et $Q = (\frac{b}{f})\frac{W}{Z_0}$. Formulons le problème de minimisation par la méthode des moindres carrés, ce qui donne l'équation suivante :

$$min_{C,P,Q} \sum_{(x,y)\in\Delta} (G(C + Px + Qy) + I_t)^2 \tag{3.80}$$

Nous remarquons que l'estimation du TAC ne nécessite pas la connaissance de la distance focale f. Si cette distance f est connue on peut déterminer l'équation de la surface plane en utilisant les deux paramètres $a = -f\frac{P}{C}$ et $b = -f\frac{Q}{C}$. Si $a = b = 0$ (surface plane est perpendiculaire à l'axe optique) le calcul du TAC est déjà discuté dans le cas 1.

Cas 4 : Mouvement de translation de la caméra et orientation de la surface plane arbitraires

Le cas qui généralise les trois cas $1, 2$ et 3 est le cas 4 dont le mouvement de la translation de la caméra n'est pas nécessairement aligné avec l'axe optique et l'orientation de la surface plane est arbitraire (voir Fig. 3.17).

On a bien : $U \neq 0, V \neq 0$ et $W \neq 0$. Reprenons un calcul identique au calcul fait pour le cas 3 en intégrant les équations (3.66) et (3.77) dans l'équation (3.4), ce qui donne :

$$\frac{-W}{Z_0}(1 - a\frac{x}{f} - b\frac{y}{f})[(x - x_0)I_x + (y - y_0)I_y] + I_t = 0 \tag{3.81}$$

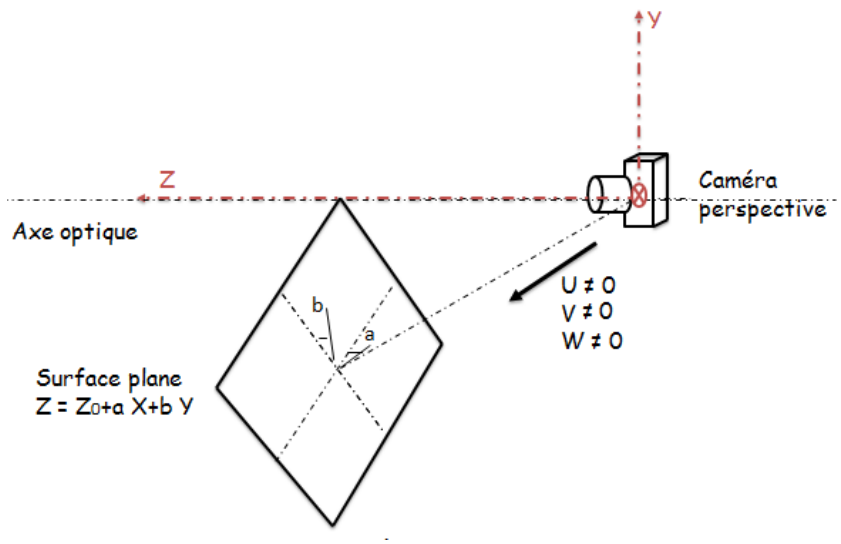


FIGURE 3.17 – Cas 4 : Translation arbitraire de la caméra relativement à une surface d'orientation quelconque

avec : $x_0 = f(\frac{U}{W})$ et $y_0 = f(\frac{V}{W})$
Reprenons les mêmes notations A, B, C, P et Q nous trouvons :

$$C(1 + x\frac{P}{C} + y\frac{Q}{C})[G + I_x\frac{A}{C} + I_y\frac{B}{C}] + I_t = 0 \tag{3.82}$$

Estimer les paramètres A, B, C, P et Q revient à résoudre l'équation suivante :

$$min_{C,P,A,B,Q} \sum_{(x,y)\in\Delta} (C(1 + x\frac{P}{C} + y\frac{Q}{C})[G + I_x\frac{A}{C} + I_y\frac{B}{C}] + I_t)^2 \tag{3.83}$$

ou bien, si on pose $F = 1 + x\frac{P}{C} + y\frac{Q}{C}$ et $D = G + I_x\frac{A}{C} + I_y\frac{B}{C}$, on trouve l'équation suivante :

$$\sum_{(x,y)\in\Delta} [CFD + I_t]^2 \tag{3.84}$$

Estimer les cinq paramètres revient à différencier la somme par rapport aux cinq paramètres ce qui génère un système d'équations non linéaires. Dans ce cas si $\frac{P}{C}$ et $\frac{Q}{C}$ sont connus, $F = 1 + x\frac{P}{C} + y\frac{Q}{C}$ est donc connu, l'équation (3.83) est linéaire, d'une manière similaire si $\frac{A}{C}$ et $\frac{B}{C}$, sont donnés alors $D = G + I_x\frac{A}{C} + I_y\frac{B}{C}$ est connu, par conséquent l'équation (3.83) est linéaire, cela revient à écrire :

$$min_{C,P,Q,D} \sum_{(x,y)\in\Delta} [(C + xP + yQ) * D + I_t]^2 \tag{3.85}$$

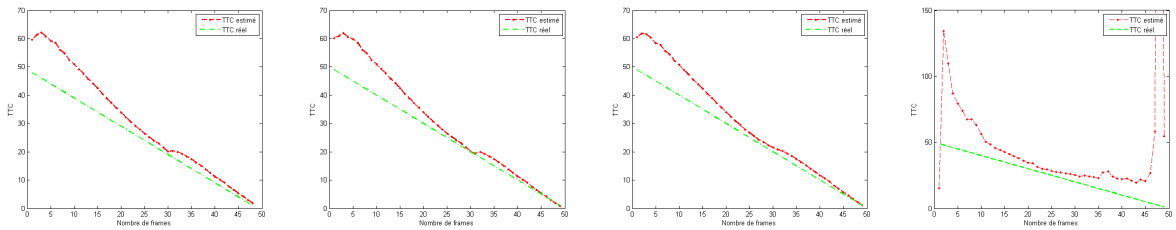


FIGURE 3.18 – Estimation du TAC dans les différents cas du mouvement de la caméra et de l'orientation de la surface : (a) cas 1, (b) cas 2, (c) cas 3 et (d) cas 4.

Les paramètres P, Q et C sont calculés à partir du système d'équation (3.85).

Récapitulatif

Nous avons présenté les différentes méthodes du calcul du TAC par la méthode des gradients d'image dans le cas perspectif, basées sur différentes contraintes et différents paramètres : $(U, V, W) = (\dot{X}, \dot{Y}, \dot{Z})$ est la vitesse relative du point P réel de l'objet et les deux inconnus a et b sont les paramètres de l'équation de la surface plane.

Pour le cas 1 : $U = V = 0$ et $a = b = 0$ implique un paramètre inconnu W.

Pour le cas 2 : $a = b = 0$ implique trois paramètres inconnus U, V, W.

Pour le cas 3 : $U = V = 0$, implique trois paramètres inconnus W, a, b.

Pour le cas 4 : pas d'hypothèses, cinq paramètres à estimer.

Nous remarquons que la méthode basée sur les gradients d'image estime le TAC globalement sur une zone d'intérêt dans l'image ou sur l'image toute entière, contrairement à la méthode basée flot optique qui donne une information locale du TAC pour chaque point de l'image. Dans la figure 3.18, nous présentons un exemple des simulations sur la séquence de la figure 3.10. Les différentes courbes du TAC correspondent aux différents *scénarii* présentés plus haut.

3.4.2 Méthode d'estimation du TAC dans le cas omnidirectionnel

Les méthodes décrites ci-dessus considèrent les images acquises avec une caméra perspective. Dans ce qui suit, nous donnons un aperçu de quelques méthodes de la littérature pour estimer le TAC avec un champ de vue plus large.

3.4.2.1 Généralisation de la méthode de Subbarao pour une direction de vue quelconque

La formulation (3.47) proposée par [Subbarao 1990] fournit une inégalité qui calcule les bornes du TAC, ceci n'est possible que si on suppose que le mouvement est rigide le long de l'axe optique **Z**. Une solution à ce problème a été proposée par [Colombo 2000] qui a démontré que les bornes du TAC peuvent être calculées dans le cas général de n'importe quelle direction de vue avec un champ de vue (FOV) plus grand.

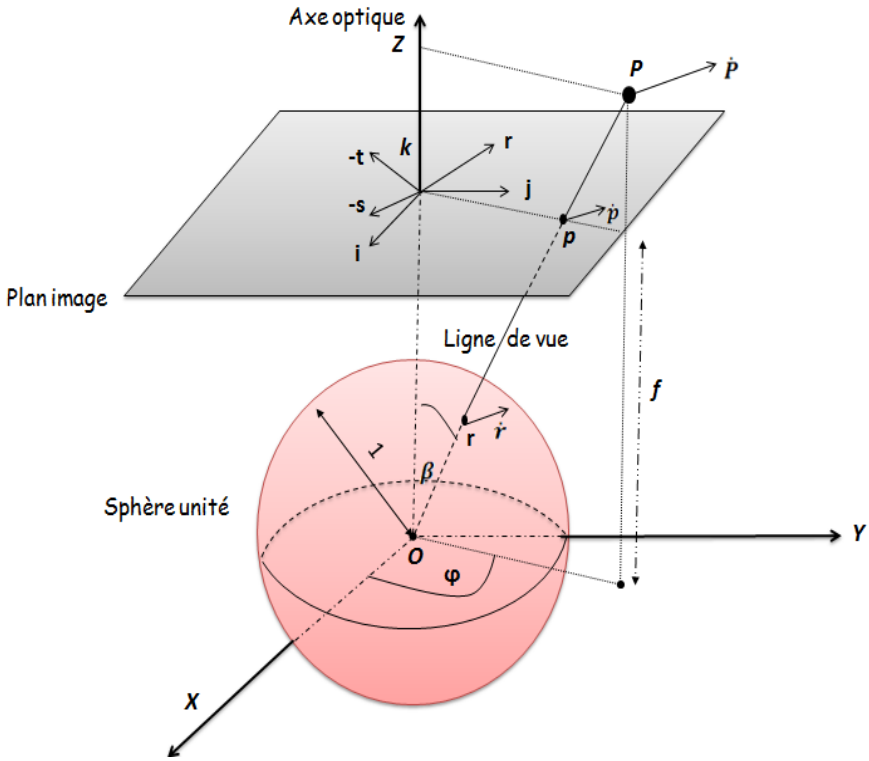


FIGURE 3.19 – Projection perspective et sphérique

Colombo [Colombo 2000] a introduit deux définitions du TAC. La première définition est valable dans le cas d'une projection perspective, dans le cas où la direction de vue est différente de l'axe optique. La seconde définition considère le cas d'une projection sphérique qui permet d'étendre le calcul du TAC sur n'importe quel champ de vue (FOV). En effet, soit $\dot{P}$ la vitesse totale du point P (voir Fig. 3.19), tel que :

$$\dot{P} = T_p + \Omega_p \wedge \overrightarrow{OP} \tag{3.86}$$

avec (T_p, Ω_p) est la vitesse et la rotation instantanées du point P. La définition du TAC

dans le cas d'une projection perspective est donnée par :

$$\tau_p = \frac{P_z}{|\dot{P}| \cos \beta} \tag{3.87}$$

où β est l'angle entre la ligne de vue passant par P et l'axe optique alors que P_z est la profondeur du point P dans la direction de l'axe optique Z.
Dans le cas sphérique, le TAC esr redéfinit comme suit :

$$\tau_s = \frac{|P|}{\dot{P}.r} \tag{3.88}$$

avec $|P|$ est la distance radiale entre le centre O de la sphère et le point P. τ_s donne une définition radiale du TTC.
Avec ces deux définitions, le calcul du TAC peut être fait dans toute direction de vue en fonction de la distance réelle $|P|$, la vitesse $\dot{P}$ et l'angle de direction de vue β. En introduisant la correspondance entre le mouvement planaire et sphérique (voir le lemme de correspondance [Colombo 2000]) et en utilisant (3.87) et (3.88), Colombo a généralisé la formule des bornes du TAC de [Subbarao 1990] comme suit :

$$\frac{div' - def'}{2} \leq \tau_s^{-1} \leq \frac{div' + def'}{2} \tag{3.89}$$

avec dev' et def' sont définis à partir des dérivées partielles du mouvement sphérique. La reformulation (3.89) représente la généralisation de l'équation (3.47).

Kazuaki et al. [Kondo 2005] ont employé la formule du TAC de [Tistarelli 1992] basée sur la représentation log-polaire et en utilisant un capteur omnidirectionnel non-isotropique couplé avec un miroir paraboloïde de type HBP (Horizontal fixed viewpoint Biconical Paraboloidal). La méthode de [Tistarelli 1992] consiste à utiliser le flot optique comme une fonction continue. Ceci n'est pas approprié dans le cas de la commande numérique d'un robot. C'est la raison pour laquelle [Kondo 2005] a proposé d'utiliser cette fonction en supposant que le mouvement relatif du robot est uniforme pendant une période de trois images. Cette méthode d'estimation du TAC est presentée dans le cadre de la vision omnidirectionnelle, mais basée sur l'hypothèse de [Tistarelli 1992] (surface fronto-parallèle), ce qui fait que cette méthode est limitée par une hypothèse forte qui n'est pas toujours verifiée dans le cadre de la robotique mobile.

3.4.3 Discussion et synthèse

Nous avons examiné les différentes techniques d'estimation du TAC qui existent dans la littérature, en citant les avantages et les limitations de chaque classe de méthodes. Le tableau 3.1 présente un résumé des approches abordées dans ce chapitre classées selon les hypothèses prises en compte. Nous avons aussi rajouté d'autres réfèrences de travaux que nous n'avons pas abordé dans ce chapitre mais qui apportent d'importantes contributions pour l'estimation du TAC.

Un inconvénient des approches d'estimations du TAC basées sur l'hypothèse du mouvement rigide est l'exigence d'une segmentation explicite afin d'estimer le mouvement de l'image. Une approche alternative nécessite de calculer le TAC à partir de l'estimation locale du flot optique général (par exemple [Camus 1995]). Ces méthodes ne nécessitent aucune connaissance *a priori* de la structure de la scène, ni d'une segmentation. En outre, les méthodes basées sur les contours fermés restent limitées par des hypothèses fortes qui ne sont pas toujours vraies dans les applications de la robotique mobile. Horn et al. [Horn 2009] proposent une méthode simple, basée sur les gradients de l'image et qui ne nécessite pas de traitements de haut niveau. Cette méthode est valable pour les surfaces planes d'orientation quelconque pour une translation arbitraire du robot. Il existe de nombreuses méthodes d'estimation du TAC dans le cas perspectif mais peu en omnidirectionnel. La méthode proposée par [Kondo 2005] d'estimation du TAC avec un capteur paracatadioptrique non isotropique est inspirée de la méthode de [Tistarelli 1992]. Néanmoins cette dernière suppose que la surface est fronto-parallèle pour chaque point de l'image, une hypothèse forte qui n'est pas valable dans les applications robotiques réelles.

Après avoir revisité les méthodes qui existent dans la littérature, nous proposons dans cette thèse deux méthodes d'estimation du TAC dans le cas catadioptrique. Une première méthode inspirée des travaux de [Horn 2007] et [Horn 2009] et la deuxième est inspirée du travail de [Camus 1995]. Ces deux méthodes fournissent deux modèles mathématiques adaptés aux images catadioptriques appliquées à des surfaces d'orientation quelconque, mais valables uniquement dans le cas d'une translation axiale (mouvement du robot aligné avec l'un des axes du capteur catadioptrique ($\mathbf{x}$ ou $\mathbf{y}$)) [Benamar 2012] et [Benamar 2013].

Afin de généraliser ces deux méthodes d'estimation du TAC, nous avons exploité les propriétés du FOE. En effet, dans des travaux antérieurs d'estimation du TAC basée sur les invariants différentiels [Ancona 1993], [Nelson 1989] et [Coombs 1998], la divergence du

flot optique (*div*) est mesurée souvent à la même position de l'image pour chaque frame. Ancona et al. [Ancona 1993] par exemple, utilisent des détecteurs de mouvement linéaires pour estimer le mouvement dans des régions disposées symétriquement autour du centre de l'image. Cependant, ceci ne tient pas compte de l'effet du changement du FOE sur la mesure de la divergence. De même Subbarao [Subbarao 1990], ne tient pas compte de l'effet du FOE et il suppose que les points d'intérêt sont placés sur l'axe optique. De telles hypothèses nécessitent un dispositif supplémentaire embarqué sur le robot pour aligner son mouvement avec l'axe optique. Dans de nombreux cas, un tel disposotif n'est pas disponible pour faciliter la fixation surtout quand le robot est à grande vitesse. Une approche alternative consiste à faire le suivi de la position du FOE. Van Leuwen et al. [Van Leeuwen 2000], [Van Leeuwen 2002] et [McCarthy 2008] montrent que l'utilisation du suivi de FOE pour corriger le désalignement de l'axe optique et l'axe de translation est important pour obtenir une estimation correcte du TAC. Nous avons profité de ces constatations pour rendre nos deux méthodes d'estimation du TAC valables dans le cas d'une translation quelconque et même dans le cas de la présence de rotations et ceci en se basant sur notre stratégie « ré-orientation » presentée dans le chapitre 4.

3.5 Conclusion

L'objectif dans cette thèse est d'estimer le temps à collision (TAC) sur des images catadioptriques. En premier lieu, nous avons presenté dans ce chapitre un bref état de l'art sur les techniques d'estimation du flot optique, car l'estimation du TAC est liée à l'estimation du flot optique d'une manière explicite ou implicite. Cet état de l'art nous a orienté sur les différentes contraintes et hypothèses supposées pour résoudre la problématique du TAC. Après avoir établi une liste de ces méthodes, nous avons choisi deux méthodes qui nécessitent moins d'hypothèses et qui sont de bonnes candidates pour des applications de robotique mobile. Nous avons aussi profité des propriétés du flot optique pour présenter la notion du foyer d'expansion (FOE). Ceci sera utile pour introduire notre stratégie de ré-orientation qui a pour but de généraliser nos deux méthodes de l'estimation du TAC adaptées aux images catadioptriques. Nous présentons ces deux méthodes dans le chapitre 4 de ce rapport.

TABLE 3.1 – Les différentes approches existantes pour l'estimation du TAC

Contraintes et restrictions	Réferences
Alignement de l'axe de mouvement avec l'axe optique	[Subbarao 1990],[Ancona 1993], [Coombs 1998], [Green 2004], [Zufferey 2006]
Surfaces fronto-parallèles	[Tistarelli 1992], [Santos-Victor 1997], [Colombo 2000], [Kondo 2005], [McCarthy 2008]
Connaissance de l'ego-motion et orientation de la surface	[Nelson 1989], [Meyer 1994], [Cipolla 1997], [Colombo 2000]
Translation arbitraire du robot avec une orientation de surface plane quelconque	[Horn 2007], [Horn 2009]
Translation arbitraire du robot avec surface quelconque	[Camus 1995]

Chapitre

4 Estimation du temps à collision dans les images catadioptriques

La perception d'un objet comme désirable ou indésirable ne réside pas dans l'objet lui-même, mais dans la façon dont on le perçoit.

Jean-François Ricard

Sommaire

4.1 Introduction **82**

4.2 Définition du TAC dans le cas catadioptrique **83**

4.3 Modèle global du TAC (MGG) **83**

4.3.1 Equation de la profondeur de la surface plane en fonction des coordonnées images 85

4.3.2 Modèle du mouvement apparent en fonction du TAC 86

4.3.3 Modèle adapté de TAC basé sur les gradients pour les surfaces planes 87

4.4 Modèle local du TAC (MLFO) **90**

4.5 Algorithme de stratégie de ré-orientation de l'image catadioptrique **91**

4.5.1 Algorithme d'extraction des foyers d'expansion et de contraction 92

4.5.2 Stratégie de la ré-orientation 95

4.6 Conclusion **97**

4.1 Introduction

Les techniques d'estimation du temps à collision (TAC) que nous avons présentées dans la section 3.4 considèrent des images acquises à partir de caméras perspectives, où la scène observée se projette sur une surface plane, représentée par des coordonnées cartésiennes et avec une résolution uniforme.

Dans ce chapitre nous présentons deux méthodes d'estimation du TAC adaptées aux images catadioptriques :

— La première méthode est appelée *méthode globale basée sur les gradients d'image*. Elle est valable pour un capteur para-catadioptrique et destinée pour les surfaces planes. Cette méthode est l'adaptation de la méthode de Horn et al. [Horn 2007, Horn 2009] déjà présentée dans le chapitre 3. Les équations d'estimation du TAC avec la méthode de Horn et al. ne sont pas applicables directement sur les images catadioptriques. Notre méthode introduit un modèle de TAC qui tient compte de la géométrie des images para-catadioptriques. Elle exploite la contrainte de la conservation lumineuse afin de calculer le TAC en utilisant seulement les dérivées spatiales et temporelles de l'intensité et sans aucun recours à un traitement de haut niveau comme la détection d'obstacle, le suivi ou l'estimation explicite du flot optique.

— La seconde méthode est appelée *méthode locale basée sur l'estimation explicite du flot optique*. Elle est plus générale, valable pour tout capteur catadioptrique PVU et valide pour n'importe quelle forme géométrique d'obstacle. Cette méthode est l'adaptation de l'approche de [Camus 1995] présentée dans le chapitre 3. Elle représente l'avantage d'être plus générale par rapport aux autres méthodes. Etant donné que l'expression du TAC dans le cas perspectif dérive du modèle géométrique de projection perspective, ce dernier n'est pas vérifié dans le cas des images catadioptriques à cause de l'utilisation du miroir. C'est pour cette raison que les méthodes usuelles de calcul du TAC ne sont pas applicables dans notre cas et que nous devons les reconsidérer en tenant compte de la géométrie particulière des images étudiées. Nous présentons une contribution plus générale d'estimation du TAC. Notre démarche consiste à reformuler un modèle adapté en utilisant un modèle géométrique de projection approprié aux images catadioptriques. Cette nouvelle approche permet de fournir une valeur du TAC en chaque point de l'image

(méthode locale) indépendamment de la forme géométrique de l'obstacle. Cette information locale du TAC, nous permet de construire des cartes de TAC de la scène 360° autour du capteur catadioptrique.

Tout d'abord, dans la deuxième section de ce chapitre, nous définissons le TAC dans le cas catadioptrique puis dans la troisième et la quatrième section nous présentons les deux méthodes adaptées du TAC dans le cas simple de translation axiale du capteur catadioptrique. Enfin, dans la cinquième section, nous introduisons la généralisation des deux modèles dans le cas de translation quelconque basée sur l'algorithme de ré-orientation de l'image.

4.2 Définition du TAC dans le cas catadioptrique

Considérons un robot mobile équipé d'une caméra catadioptrique, se déplaçant dans un plan perpendiculaire à son axe optique suivant la direction de l'axe $\mathbf{X}$ (voir la Fig. 4.1). Dans le cas perspectif, la méthode fondée sur le flot optique n'est valable que si l'axe optique de la caméra perspective correspond au sens du mouvement du robot. Le TAC est alors défini comme le temps nécessaire pour que l'obstacle atteigne le point O (voir la Fig. 3.9). Dans le cas de la caméra catadioptrique, on définit l'état de collision comme l'instant où l'obstacle percutant le foyer F du miroir. Le point de contact C sur l'obstacle est en $(X_0, 0, 0)^T$. La projection de C sur le plan image est le foyer d'expansion (FOE). Soit $P(X, Y, Z)$ un point appartenant à un obstacle fixe. Le TAC du point P est défini par l'équation :

$$\tau = -\frac{X}{\dot{X}} \tag{4.1}$$

4.3 Modèle global basé sur les gradients d'intensité dans le cas d'une translation axiale du capteur para-catadioptrique

Nous cherchons dans cette section, à décrire une méthode simple d'estimation du TAC valable pour les images para-catadioptriques, en utilisant seulement les gradients de l'image, sans avoir recours à des méthodes de haut niveau comme l'estimation explicite du flot

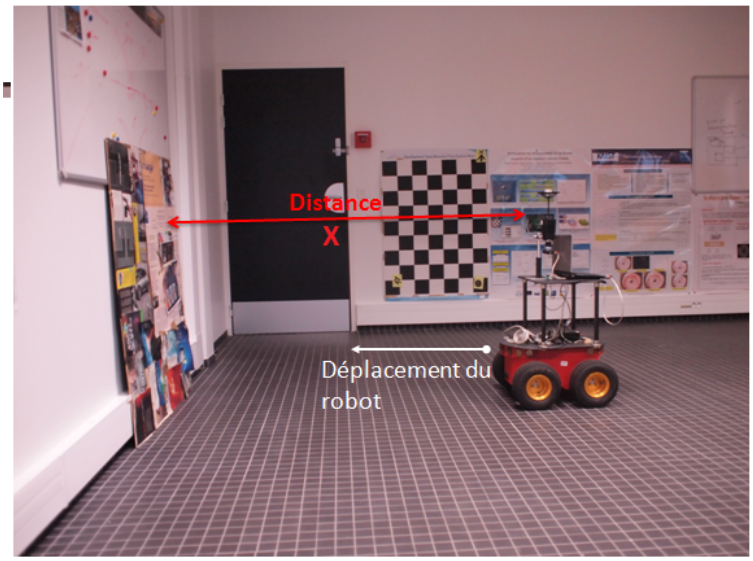


FIGURE 4.1 – Robot Pioneer-AT muni d'une caméra catadioptrique, s'approchant d'un obstacle

optique, le suivi ou les méthodes de détection d'obstacle. Cette approche est destinée aux surfaces planes dans le cas d'un mouvement de translation pure du capteur para-catadioptrique. Soit $P(X, Y, Z)$ un point de la surface plane dans l'espace $\mathbb{R}^3$. Le point de contact avec la surface plane est en $(X_0, 0, 0)^t$ (voir Fig. 4.2). L'équation du TAC en ce point de contact est définie par :

$$TAC = -\frac{X_0}{\dot{X}} \tag{4.2}$$

Les équations d'estimation du TAC dans le cas perspectif (3.72), (3.75), (3.80) et (3.85) sont introduites à partir de l'équation de la projection perspective. Cette loi géométrique de projection n'est pas applicable dans le cas para-catadioptrique à cause de l'introduction du miroir comme nous l'avons expliqué dans le chapitre 3. Dans la section suivante, nous allons revisiter cette méthode basée sur les gradients d'image pour l'adapter aux images para-catadioptriques. Pour cela nous utilisons des outils et des contraintes adaptés à la géométrie des images para-catadioptriques. Nous proposons dans ce chapitre d'établir un nouveau modèle d'estimation du TAC adapté à ces images en introduisant :

— L'équation de la profondeur de la surface plane en fonction des coordonnées images.
— Et le modèle du mouvement apparent en fonction du TAC.

4.3.1 Equation de la profondeur de la surface plane en fonction des coordonnées images

L'analyse présentée ci-dessous, traite le cas de la translation (le long de l'axe **X**) du capteur para-catadioptrique relativement à une surface plane inclinée, qui peut être parfaitement décrite à l'aide de deux paramètres d'inclinaison a et b (voir la Fig. 4.2).

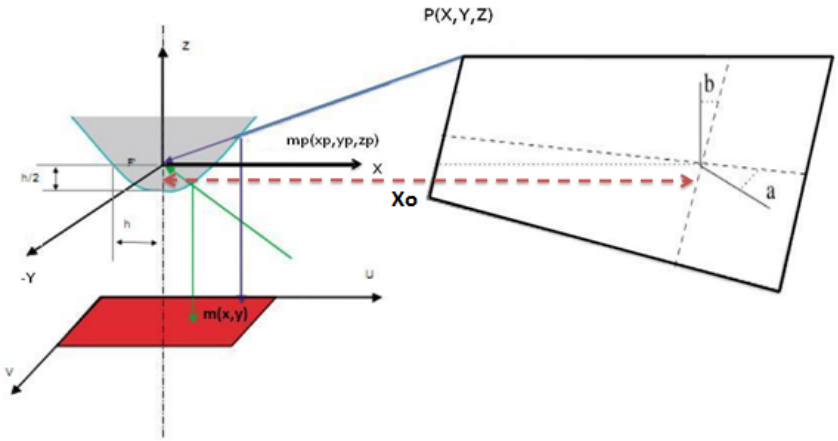


FIGURE 4.2 – Projection para-catadioptrique d'une surface plane

Soit $(S) \in \mathbb{R}^3$ une surface plane. Nous pouvons décrire la profondeur des points de la surface (S) par la fonction de profondeur suivante :

$$X(Y,Z) = X_0 + aY + bZ \tag{4.3}$$

Longuet-Higgins et Prazdny [Longuet-Higgins 1980] introduisent les équations de projection perspective dans l'équation de la surface plane, définissant ainsi la profondeur de la surface en fonction des coordonnées dans l'image. A cause de la présence du miroir de révolution, nous devons introduire les modèles de projection parabolique (2.4), (2.6) dans l'équation (4.3) de la surface plane (S), afin d'exprimer la fonction de profondeur en fonction des coordonnées image. Ainsi, nous obtenons l'équation suivante :

$$X = X_0 + ay\frac{\sqrt{X^2+Y^2+Z^2}-Z}{h} + bz_p\frac{\sqrt{X^2+Y^2+Z^2}-Z}{h} \tag{4.4}$$

Nous remplaçons l'équation du miroir paraboloïde z_p (2.5) dans l'équation (4.4). Nous obtenons donc :

$$X = X_0 + \frac{\sqrt{X^2+Y^2+Z^2}-Z}{h}(ay + b\frac{x^2+y^2-h^2}{2h}) \tag{4.5}$$

A partir de l'équation (2.6), nous constatons que :

$$\frac{\sqrt{X^2+Y^2+Z^2}-Z}{h}=\frac{X}{x} \tag{4.6}$$

Enfin, en remplaçant l'équation (4.6) dans l'équation (4.5), nous avons :

$$X(x,y)=\frac{X_0}{1-(a\frac{y}{x}+b\frac{(x^2+y^2-h^2)}{2hx}))} \tag{4.7}$$

L'équation (4.7) représente l'adaptation de l'équation (3.77), elle estime la profondeur de la surface (S) en fonction des coordonnées image dans le cas de la translation du capteur para-catadioptrique vers une surface plane d'orientation quelconque.

4.3.2 Modèle du mouvement apparent en fonction du TAC

Les équations qui lient les coordonnées images avec les coordonnées 3D sont définies par le modèle (2.6). Posant : $A=X^2+Y^2+Z^2$. Nous obtenons le mouvement apparent dans l'image $(\dot{x},\dot{y})$ en dérivant le système d'equation (2.6) par rapport au temps :

$$\dot{x}=\frac{h\dot{X}(\sqrt{A}-Z)-hX((X\dot{X}+Y\dot{Y}+Z\dot{Z})(A)^{-1/2}-\dot{Z})}{(\sqrt{A}-Z)^2} \tag{4.8}$$

De la même manière, nous obtenons la composante $\dot{y}$:

$$\dot{y}=\frac{h\dot{Y}(\sqrt{A}-Z)-hY((X\dot{X}+Y\dot{Y}+Z\dot{Z})(A)^{-1/2}-\dot{Z})}{(\sqrt{A}-Z)^2}. \tag{4.9}$$

En considérant que le mouvement général du capteur para-catadioptrique est une translation le long de l'axe $\mathbf{X}$, on a : $\dot{Z}=0$ et $\dot{Y}=0$, donc :

$$\dot{x}=\frac{h\dot{X}(\sqrt{A}-Z)-hX(X\dot{X})A^{-1/2}}{(\sqrt{A}-Z)^2} \tag{4.10}$$

et

$$\dot{y}=\frac{-hY(X\dot{X})A^{-1/2}}{(\sqrt{A}-Z)^2}. \tag{4.11}$$

Posons : $C_X=-\frac{\dot{X}}{X}$. Donc, (4.10) devient :

$$\dot{x}=xC_X(\frac{x}{h}(XA^{-1/2})-1) \tag{4.12}$$

et de la même manière, nous trouvons la composante $\dot{y}$ de l'équation (4.11) :

$$\dot{y}=C_X(\frac{xy}{h}(XA^{-1/2})) \tag{4.13}$$

Le terme $(XA^{-1/2})$ de (4.12) et (4.13) est exprimé en coordonnées 3D. Il peut être écrit comme suit :

$$XA^{-1/2} = \frac{X}{\sqrt{X^2+Y^2+Z^2}} \tag{4.14}$$

Donc, nous pouvons écrire le terme (4.14) comme suit :

$$XA^{-1/2} = \begin{cases} \frac{1}{\sqrt{\frac{X^2+Y^2+Z^2}{X^2}}} & \text{if} X > 0 \\ \frac{-1}{\sqrt{\frac{X^2+Y^2+Z^2}{X^2}}} & \text{if} X < 0 \\ 0 & \text{if } X = 0. \end{cases} \tag{4.15}$$

Nous avons :

$$\frac{1}{\sqrt{\frac{X^2+Y^2+Z^2}{X^2}}} = \frac{1}{\sqrt{\frac{(z_p+h)^2}{x^2}}} \tag{4.16}$$

Etant donné que $z_p + h$ est toujours positif et que x et X ont toujours le même signe. En utilisant (2.5), (2.4), (2.6) nous déduisons que :

$$XA^{-1/2} = \begin{cases} \frac{x}{(z_p+h)} & \text{if } X > 0 \\ \frac{x}{(z_p+h)} & \text{if } X < 0 \\ 0 & \text{if } X = 0. \end{cases} \tag{4.17}$$

Ainsi, $\forall X \in \mathbb{R}$, nous avons :

$$XA^{-1/2} = \frac{x}{(z_p+h)} = \frac{2hx}{x^2+y^2+h^2} \tag{4.18}$$

Par conséquent, $\dot{x}$ et $\dot{y}$ sont reformulées par :

$$\begin{cases} \dot{x} = xC_X(\frac{2x^2}{x^2+y^2+h^2} - 1) \\ \dot{y} = yC_X(\frac{2x^2}{x^2+y^2+h^2}) \end{cases} \tag{4.19}$$

A partir de l'équation de la profondeur en fonction des coordonnées images (4.7) et du modèle du mouvement apparent en fonction du TAC (4.19), nous pouvons introduire le modèle adapté d'estimation du TAC basé sur les gradients d'intensité pour les surfaces planes. Ce modèle sera présenté dans la section suivante.

4.3.3 Modèle adapté de TAC basé sur les gradients pour les surfaces planes

En se basant sur les travaux de [Horn 2009], l'idée est d'exploiter l'équation du flot optique (ECMA), le modèle de mouvement apparent en fonction du TAC (4.19) (dérivé des

expressions de $(\dot{x}, \dot{y})$) et l'équation de la surface en coordonnées images (4.7).
Nous supposons que la translation du robot est suivant l'axe $\mathbf{X}$ relativement à une surface d'orientation arbitraire. Nous rappelons l'équation (ECMA) déjà définie dans le chapitre 3 :

$$\mathbf{v}.\nabla I + \frac{\partial I}{\partial t} = 0 \tag{4.20}$$

avec $\mathbf{v} = (\dot{u}, \dot{v})$

$$\dot{u} = \alpha_u \dot{x}; \dot{v} = \alpha_v \dot{y} \tag{4.21}$$

Dans tout ce qui suit, nous considérons que les facteurs d'echelle sont égaux, c'est-à-dire $\alpha_u = \alpha_v = \alpha$. Les composantes $\dot{x}$ et $\dot{y}$ sont exprimées par l'equation (4.19). Donc, l'ECMA devient :

$$\alpha x C_X (\frac{2x^2}{x^2 + y^2 + h^2} - 1) I_u + \alpha y C_X (\frac{2x^2}{x^2 + y^2 + h^2}) I_v + I_t = 0 \tag{4.22}$$

Avec $C_X = -\frac{\dot{X}}{X}$. Si nous posons $xI_u + yI_v = G$ (gradient radial), nous obtenons :

$$\alpha C_X \frac{2x^2}{x^2 + y^2 + h^2} G - \alpha C_X x I_u + I_t = 0 \tag{4.23}$$

A partir de la formule (4.2), nous avons :

$$C = C_X \frac{X_0}{X} \tag{4.24}$$

avec C est l'inverse du TAC.
Par suite et en utilisant (4.7) et (4.24), on obtient l'équation suivante :

$$C_X = C.(1 - (a\frac{y}{x} + b\frac{x^2 + y^2 - h^2}{2hx})) \tag{4.25}$$

Ainsi en considérant les équations (4.23) et (4.25), nous avons :

$$S = C.[\frac{2\alpha x^2}{x^2 + y^2 + h^2} G - \alpha x I_u] - aC[\frac{2\alpha xy}{x^2 + y^2 + h^2} G - \alpha y I_u] -$$

$$bC[(\frac{\alpha x}{h})\frac{x^2 + y^2 - h^2}{x^2 + y^2 + h^2} G - \alpha \frac{x^2 + y^2 - h^2}{2h} I_u] + I_t = 0 \tag{4.26}$$

Posons : $P = -aC, Q = -bC$ et $G^* = (u - u_0)I_u + (v - v_0)I_v$. Finalement, l'équation (4.26) en coordonnées pixelliques s'écrit :

$$C[\frac{2(u - u_0)^2}{(u - u_0)^2 + (v - v_0)^2 + (\alpha h)^2} G^* - (u - u_0) I_u] +$$

$$P[\frac{2(u-u_0)(v-v_0)}{(u-u_0)^2+(v-v_0)^2+(\alpha h)^2}G^* - (v-v_0)I_u]+$$

$$+Q[\frac{(u-u_0)}{\alpha.h}\frac{(u-u_0)^2+(v-v_0)^2-(\alpha h)^2}{(u-u_0)^2+(v-v_0)^2+(\alpha h)^2}G^*-$$

$$\frac{(u-u_0)^2+(v-v_0)^2-(\alpha h)^2}{2.\alpha.h}I_u] + I_t = 0 \tag{4.27}$$

Dans cette dernière équation, nous avons trois variables à estimer. Pour estimer le C (inverse du TAC), la résolution de l'équation (4.27) se fait comme suit :

— Pour chaque point dans la région d'intérêt Δ, il faut écrire l'équation (4.27) correspondante.

— Construire un système d'équations correspondant composé d'autant d'équations que de points dans Δ ;

— Résoudre le système par minimisation des moindres carrées.

Il s'agit donc de résoudre au sens des moindres carrées, le système linéaire suivant :

$$H\Theta = B \tag{4.28}$$

avec

$$\Theta = \begin{pmatrix} C \\ P \\ Q \end{pmatrix} \tag{4.29}$$

et

$$B = [-I_t] \tag{4.30}$$

et

$$H = [H_1 \quad H_2 \quad H_3], \tag{4.31}$$

où

$H_1 = \frac{2(u-u_0)^2}{(u-u_0)^2+(v-v_0)^2+(\alpha h)^2}G^* - (u-u_0)I_u$

$H_2 = \frac{2(u-u_0)(v-v_0)}{(u-u0)^2+(v-v_0)^2+(\alpha h)^2}G^* - (v-v_0)I_u$

$H_3 = \frac{(u-u_0)}{\alpha.h}\frac{(u-u_0)^2+(v-v_0)^2-(\alpha h)^2}{(u-u_0)^2+(v-v_0)^2+(\alpha h)^2}G^* - \frac{(u-u_0)^2+(v-v_0)^2-(\alpha h)^2}{2.\alpha.h}I_u$

Finalement, si $a = b = 0$, c'est à dire la surface plane est perpendiculaire à l'axe **X**, l'équation (4.28) se simplifie en :

$$CH_1 = I_t \tag{4.32}$$

Normalement, lorsque nous n'avons aucune information préalable sur l'orientation de la surface plane, il convient d'utiliser (4.28). Cette méthode permet d'estimer trois variables seulement avec les gradients spatiaux, gradient temporel et les coordonnées images. Le principal avantage de cette méthode est que nous n'avons pas besoin d'informations supplémentaires pour calculer a et b. Dans la section suivante nous présentons une méthode d'estimation du TAC plus générale destinée à tous les capteurs catadioptriques PVU et valable pour toute forme géométrique d'obstacle.

4.4 Modèle du TAC local basé sur le flot optique dans le cas d'un mouvement axial du capteur catadioptrique

Dans cette section, nous proposons une méthode d'estimation du TAC général qui fonctionne pour toute caméra catadioptrique à PVU, pour ce fait, nous utilisons le modèle (2.19) qui modèlise toutes les caméras à point de vue unique. Le TAC local τ pour chaque point de la scène 3D est défini par l'équation suivante :

$$\tau = -\frac{X}{\dot{X}} \tag{4.33}$$

En utilisant le système d'équations (2.19), nous obtenons :

$$\frac{x}{y} = \frac{X}{Y} \tag{4.34}$$

Dérivons cette expression par rapport au temps (avec $\dot{x} = \frac{dx}{dt}$) :

$$\begin{aligned} \dot{x} &= \tfrac{Y(y\dot{X})-(yX)\dot{Y}}{Y^2} \\ &= \tfrac{Y(X\dot{y}+y\dot{X})}{Y^2} - \tfrac{(yX)\dot{Y}}{Y^2} \\ &= \tfrac{X}{Y}\dot{y} + \tfrac{y}{Y}\dot{X} - \tfrac{yX}{Y^2}\dot{Y} \\ &= \tfrac{X}{Y}(\dot{y} + \tfrac{y}{X}\dot{X} - \tfrac{y}{Y}\dot{Y}) \end{aligned} \tag{4.35}$$

Sachant que le robot se déplace le long de l'axe $\mathbf{X}$, nous avons $\dot{Y} = 0$. En combinant les équations (4.34) et (4.35), nous obtenons :

$$\dot{x} = \frac{x}{y}(\dot{y} - y\tau^{-1}) \tag{4.36}$$

Par conséquent l'expression du TAC est donnée par :

$$\tau = \frac{1}{\frac{\dot{y}}{y} - \frac{\dot{x}}{x}} \tag{4.37}$$

En coordonnées pixelliques, nous pouvons exprimer l'équation (4.37) par :

$$\tau = \frac{1}{\frac{\dot{v}}{v-v_0} - \frac{\dot{u}}{u-u_0}} \tag{4.38}$$

Notons que l'expression de τ est indépendante des paramètres h, α_u et α_v. Un étalonnage complet du capteur n'est donc pas nécessaire pour calculer le TAC, seules les coordonnées du point principal (u_0, v_0) doivent être connues.
Pour calculer le TAC au point (u, v), nous avons besoin d'estimer le mouvement apparent en ce point. Plusieurs méthodes d'estimation ont été développées dans la littérature. Nous avons choisi à titre d'exemple d'utiliser le modèle du flot adapté aux images para-catadioptriques proposé dans [Radgui 2008]. Cette approche permet d'améliorer l'estimation du flot optique sur des images para-catadioptriques en évitant le recours à la sphère d'équivalence [Geyer 2001] qui nécessite une phase d'interpolation coûteuse en terme de temps de calcul.
Dans la section suivante, nous décrivons un algorithme permettant de généraliser les deux modèles d'estimation du TAC global (4.38) et local (4.28) décrits ci-dessus.

4.5 Généralisation des modèles adaptés d'estimation du TAC dans le cas d'un mouvement de translation arbitraire du capteur catadioptrique

Pour généraliser les modèles adaptés d'estimation du TAC (4.28) et (4.38), nous utilisons une stratégie de « ré-orientation de l'image ». Elle nécessite d'estimer la rotation de l'image afin de pouvoir utiliser les modèles mathématiques (4.28) et (4.38). En effet, l'objectif est de détecter la direction du vecteur de foyer d'expansion (qui correspond à la direction du déplacement) puis de changer le repère de l'image (ré-orienter) afin de se retrouver dans le cas le plus simple d'une translation suivant l'axe **X**.
Dans le cas perspectif, les auteurs de [Van Leeuwen 2000], [Van Leeuwen 2002] et [McCarthy 2008] montrent que l'utilisation du suivi du FOE pour corriger le désalignement de l'axe optique

et de l'axe de translation est importante pour obtenir une estimation correcte du TAC. De la même manière, la position de FOE indique la direction de la translation sur le plan image omnidirectionnel. L'extraction du FOE a été calculée en utilisant l'algorithme de Kim [Kim 2007]. Dans ce qui suit, nous présentons la stratégie de ré-orientation proposée sous forme du diagramme donné à la figure 4.3. Les étapes principales de cet algorithme consistent à extraire le FOE de l'image catadioptrique en utilisant le flot optique de Lucas-Kanade et d'en déduire l'angle de rotation de l'image recherché.

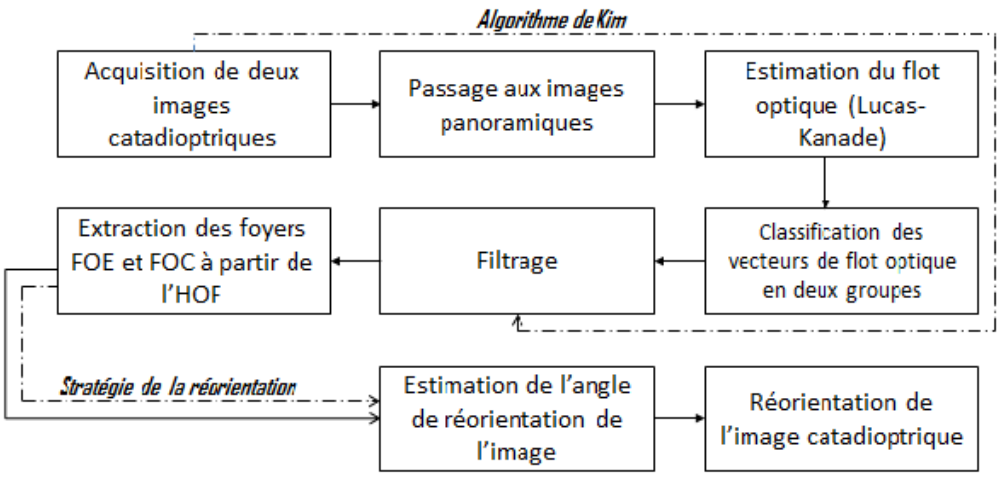


FIGURE 4.3 – Diagramme résumant l'algorithme de la stratégie de réorientation.

4.5.1 Algorithme d'extraction des foyers d'expansion et de contraction

Nous résumons l'algorithme de KIM [Kim 2007] dans les six étapes suivantes :

— Etape 1 : Acquisition de deux images omnidirectionnelles en niveaux de gris.

— Etape 2 : L'algorithme de Kim est basé sur le calcul de l'histogramme des orientations des vecteurs du flot optique sur deux images panoramiques. Nous avons besoin de transformer les images omnidirectionnelles en images panoramiques en appliquant une transformation polaire avec une interpolation cubique.

— Etape 3 : Calcul du flot optique à l'aide de l'algorithme de Lucas-Kanade sur les images panoramiques. La figure 4.4 montre les flots optiques et les vecteurs (FOE) et (FOC) qui sont indiqués par (1) et (2) respectivement, ils sont définis comme

des vecteurs verticaux de bas en haut dans l'image panoramique et du centre au point FOE et FOC dans l'image omnidirectionnelle.

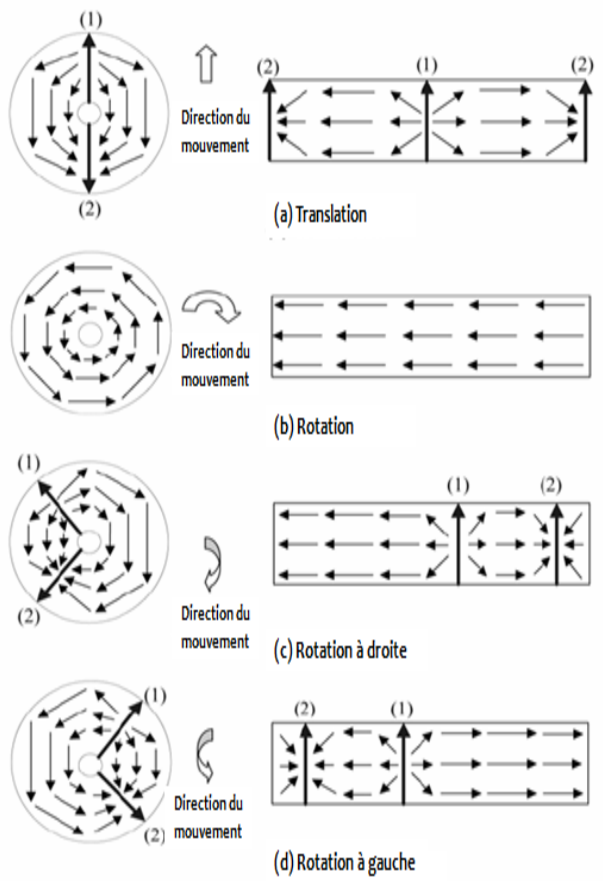


FIGURE 4.4 – Flot optique et les vecteurs FOE(1) et FOC(2) dans l'image omnidirectionnelle et l'image panoramique [Kim 2007].

— Etape 4 : L'histogramme du flot optique est basé sur le calcul de la matrice des orientations $\Theta = (\theta(u,v))_{u,v}$ du flot optique de l'image panoramique. L'orientation $\theta(u,v)$ du flot optique $\mathbf{v} = (\dot{u}, \dot{v})$ au point (u,v) de l'image panoramique est définie comme suit : $\theta(u,v) = \arctan(\dot{v}, \dot{u})$. Nous remarquons que les orientations du flot optique séparent l'image en deux régions, une première au milieu avec de faibles orientations et une deuxième avec orientations grandes (en valeur absolue) (voir chapitre 5). La binarisation de la matrice Θ, dans ce cas, consiste à affecter aux orientations élevées ($|\theta| \geq \pi/2$) la valeur 0 et aux orientations faibles ($|\theta| < \pi/2$) la valeur 1. La figure 4.5 montre un exemple de la matrice binaire obtenue notée A.

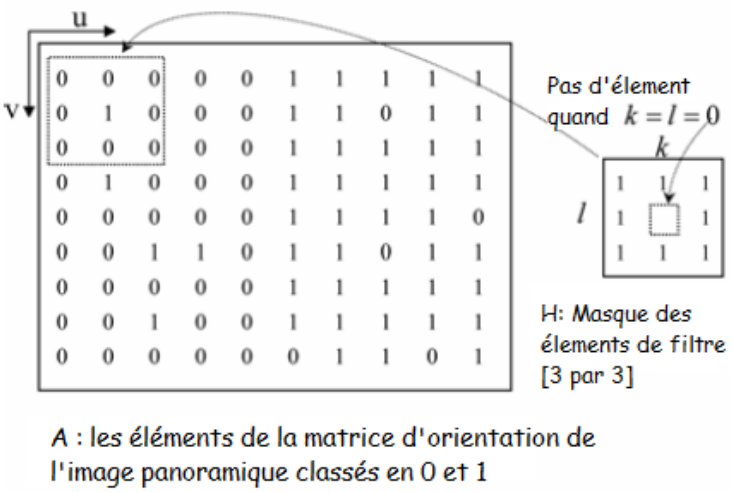


FIGURE 4.5 – Binarisation de la matrice des orientations Θ de l'image panoramique [Kim 2007].

— Etape 5 : Le but du filtrage est de corriger les valeurs mal estimées du flot optique dans l'image panoramique. Pour cela, nous optons pour une technique de filtrage avec le masque $H = [1 \quad 1 \quad 1; 1 \quad 0 \quad 1; 1 \quad 1 \quad 1]$. Il s'agit de convoluer l'image binaire A avec le masque H comme suit :

$$f(u,v) = \sum_{k=-1}^{l} \sum_{k=-1}^{l} A(u+k, v+l) H(k,l) \tag{4.39}$$

avec $k = l \neq 0$ et (u,v) est le point dans l'image panoramique. Le résultat de cette convolution est explicité dans le tableau 4.1.

Le filtrage est effectué comme suit :

$$\begin{cases} A(u,v) = & 1 \quad \text{Si} \quad f(u,v) \geq 6 \\ A(u,v) = & A(u,v) \quad \text{Si} \quad 2 < f(u,v) < 6 \\ A(u,v) = & 0 \quad \text{Si} \quad f(u,v) \leq 2 \end{cases} \tag{4.40}$$

Le tableau 4.1 décrit la sortie du filtre A. $\sharp 1$ et $\sharp 0$ sont les nombres d'élements 1 et 0 dans la matrice $[3 \times 3]$ du filtre $A(u,v)$ autour du pixel (u,v). Le maximum et le minimum de la fonction $f(u,v)$ sont 8 et 0. Les seuils choisis sont 6 et 2 pour affecter à $A(u,v)$ la valeur 1 ou 0, sinon l'élement $A(u,v)$ garde sa valeur courante.

— Etape 6 : Histogramme du flot optique (HOF) : Les vecteurs (FOE) et (FOC) sont déterminés en utilisant l'histogramme des orientations du flot optique [Chaudhry 2009].

$\sharp 1$	$\sharp 0$	$f(u,v)$	Sortie du filtre $A(u,v)$
8	0	8	1
7	1	7	1
6	2	6	1
5	3	5	$A(u,v)$
4	4	4	$A(u,v)$
3	5	3	$A(u,v)$
2	6	2	0
1	7	1	0
0	8	0	0

TABLE 4.1 – La sortie du filtre A

Après l'étape de binarisation et de filtrage, nous gardons seulement les vecteurs de flot optique de faibles orientations. Nous construisons l'HOF à partir de cette région. L'extraction des foyers (FOE) et (FOC) se fait en choisissant un seuil avec lequel nous pouvons déterminer le minimum et le maximum des bins de l'histogramme (voir Fig. 4.6).

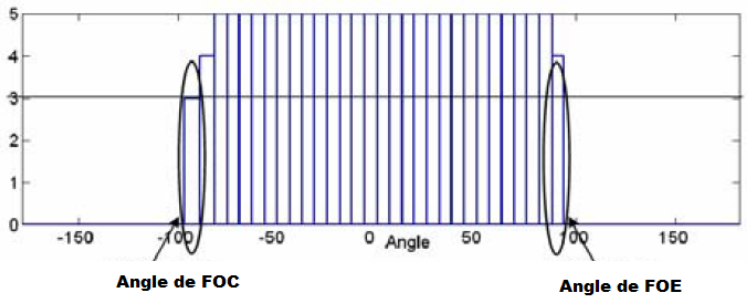


FIGURE 4.6 – Angle des vecteurs FOE et FOC extrait de l'histogramme du flot optique (HOF).

4.5.2 Stratégie de la ré-orientation

Le point FOE estimé dans l'image panoramique, nous permet de trouver son correspondant dans l'image omnidirectionnelle qui représente le point FOE recherché. L'angle de

rotation ρ cherché est l'angle entre l'axe **u** de l'image et l'axe passant par le FOE et le centre de l'image omnidirectionnelle.

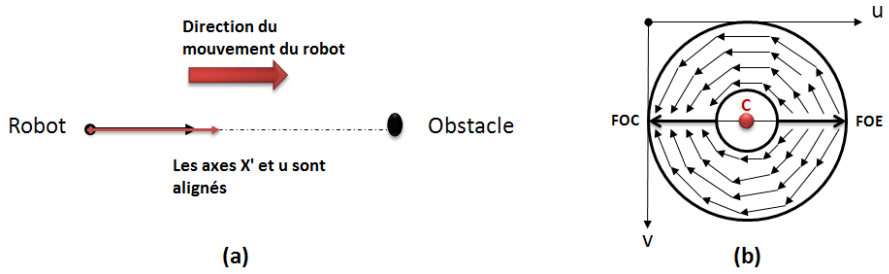


FIGURE 4.7 – Les positions des points FOE et FOC dans le cas où les axes du robot et de la caméra (X' et u) se sont alignés. (a) La translation du robot et de la caméra se sont dans la direction de l'obstacle. (b) Les positions de FOE et FOC dans l'image omnidirectionnelle.

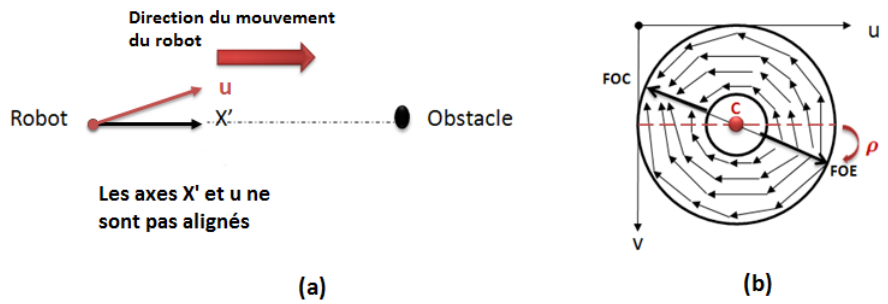


FIGURE 4.8 – Les positions des points FOE et FOC dans le cas où les axes du robot et du capteur (X' et u) ne sont pas alignés. (a) Les directions des axes de translation du robot et du capteur ne sont pas alignés. (b) Les positions de FOE et FOC dans l'image omnidirectionnelle.

L'angle ρ est utilisé pour ré-orienter l'image afin de transformer le mouvement en une translation suivant l'axe u (voir les Figs. 4.7 et 4.8). Ainsi, la contrainte $\dot{Y} = 0$ est satisfaite, et par conséquent nous pouvons appliquer les modèles (4.38) et (4.28). Dans le chapitre des résultats expérimentaux, nous illustrons les démarches décrites ci- dessus sur des images de synthèse et sur des images réelles.

4.6 Conclusion

Dans ce chapitre, nous avons presenté deux nouvelles contributions d'estimation de temps à collision (TAC) en vision catadioptrique. Ces approches sont la « méthode globale basée sur les gradients d'intensité » et la « méthode locale basée sur le flot optique » (voir la table 4.2).

TABLE 4.2 – Résumé des approches d'estimation du TAC en vision catadioptrique

Contraintes et restrictions	Réferences
Translation générale du robot par rapport à une surface plane d'orientation quelconque ($FOV = 360°$)	Chapitre 4
Translation générale du robot vers une surface quelconque ($FOV = 360°$)	Chapitre 4

La première méthode consiste à estimer le TAC dans le cas le plus simple d'une surface plane et un capteur para-catadioptrique. Cette approche fournit une information globale du TAC et elle l'estime seulement avec un calcul des gradients spatiaux et gradient temporel. Elle est basée principalement sur la formulation de l'équation de la profondeur en fonction des coordonnées images et la formulation d'une nouvelle expression du flot optique en fonction du TAC. Ces deux outils nous permettent de reformuler un nouveau modèle adapté pour l'estimation du TAC en exploitant la contrainte de la conservation lumineuse. Elle doit permettre non seulement une meilleure estimation en terme de temps de calcul mais aussi fournir une information supplémentaire sur le vecteur normal à surface plane, cette technique est simple, directe et elle peut être implementée en temps réel. Notre deuxième contribution nécessite d'introduire une nouvelle formulation plus générale du TAC dérivé d'un modèle géométrique de projection approprié aux images catadioptriques. Cette méthode reformule le TAC pour tous les capteurs à PVU, ainsi que pour toute forme géométrique d'obstacle. Cette méthode nous permet de former des cartes à (360°) de TAC de l'environnement et elle nécessite aucune segmentation à priori.
Ces deux méthodes developpées ont été introduites dans le cas d'une translation axiale. La généralisation à des translations quelconques dans le plan nécessite d'introduire la stratégie de ré-orientation de l'image catadioptrique en estimant tout d'abord le FOE et

par suite l'angle ρ de la rotation appliquée à l'image.

Les résultats expérimentaux obtenus sur des images de synthèse et réelles de ces approches sont presentés dans le chapitre 5.

Chapitre

5 Validation expérimentale et Application

Tout ce que l'esprit peut
concevoir et croire, il peut aussi
le réaliser.
Charles HAANEL

Sommaire

5.1 Introduction **99**

5.2 Résultats expérimentaux **100**

5.2.1 Plate-forme expérimentale utilisée 100

5.2.2 Mouvement axial du capteur catadioptrique vers l'obstacle . . 103

5.2.3 Mouvement général du capteur catadioptrique : stratégie de réorientation de l'image 121

5.2.4 Comparaison des deux méthodes MGG et MLFO 132

5.3 Conclusion **134**

5.1 Introduction

Les chapitres 2 et 3 ont permis de positionner notre travail et le chapitre 4, de proposer deux méthodologies d'estimation du TAC en vision catadioptrique. La première est

basée sur l'estimation implicite du flot optique appellée méthode "*globale basée sur les gradients d'image*" et la deuxième est basée sur l'estimation explicite du flot optique appellée méthode "*locale basée sur le flot optique*". Dans ce chapitre, nous présentons une série d'expérimentations réalisées pour valider et illustrer ces contributions ainsi qu'une comparaison entre les deux méthodes.

5.2 Résultats expérimentaux

Les résultats concernent la totalité des modèles développés. La première partie du présent chapitre est consacrée aux différentes simulations de la méthode locale et la méthode globale dans le cas d'un mouvement axial du capteur catadioptrique. La seconde partie sera consacrée à la présentation des résultats des deux méthodes proposées dans le cas d'un mouvement général du capteur.
Pour les deux méthodes, nous décrivons tout d'abord les outils et le matériel utilisés. Nos résultats théoriques sont validés sur des données générées avec un simulateur virtuel et avec des séquences vidéos réelles.

5.2.1 Plate-forme expérimentale utilisée

Nous générons des séquences d'images de synthèse en utilisant le logiciel POVRAY [1]. Des images de résolution $200 * 200$ sont obtenues en déplaçant une caméra catadioptrique virtuelle vers une surface plane (l'obstacle). Un exemple d'images générées avec ce logiciel est illustré à figure 5.1. Les paramètres intrinsèques de la caméra virtuelle sont fixés pour toutes les séquences aux valeurs presentées dans les tableaux 5.1 et 5.2. L'avantage principal des images de synthèse est que la vérité terrain peut être connue avec exactitude et les problèmes de changement de luminosité, de fluctuations de vitesse du robot et de taux d'acquisition peuvent être contrôlés.
Nous utilisons également des séquences d'images catadioptriques réelles dans nos expérimentations (voir la Fig. 5.2a). Ces images sont obtenues en utilisant une caméra Sony Uye dirigée vers un miroir convexe embarquée sur un robot mobile "Pioneer 3-AT". Le dispositif est illustré sur la figure 5.2b. La calibration de la caméra est effectuée avec la toolbox Hyscas [Caron 2011] (voir les tableaux 5.1 et 5.2).

1. Logiciel libre disponible sur http ://www.povray.org/

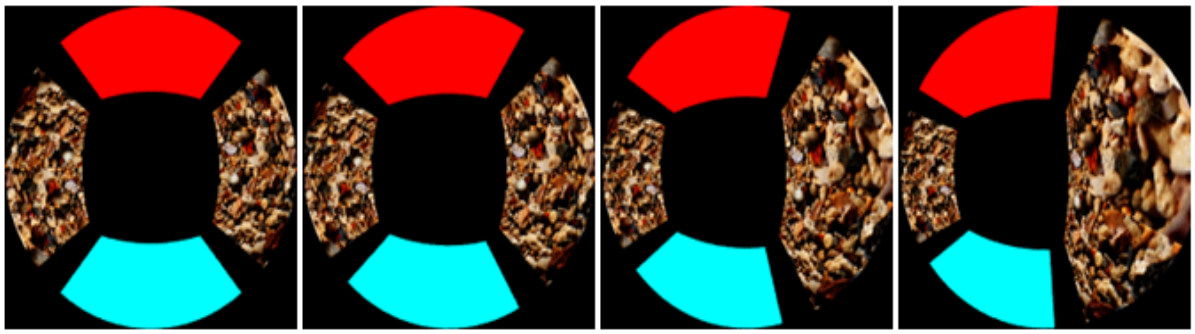
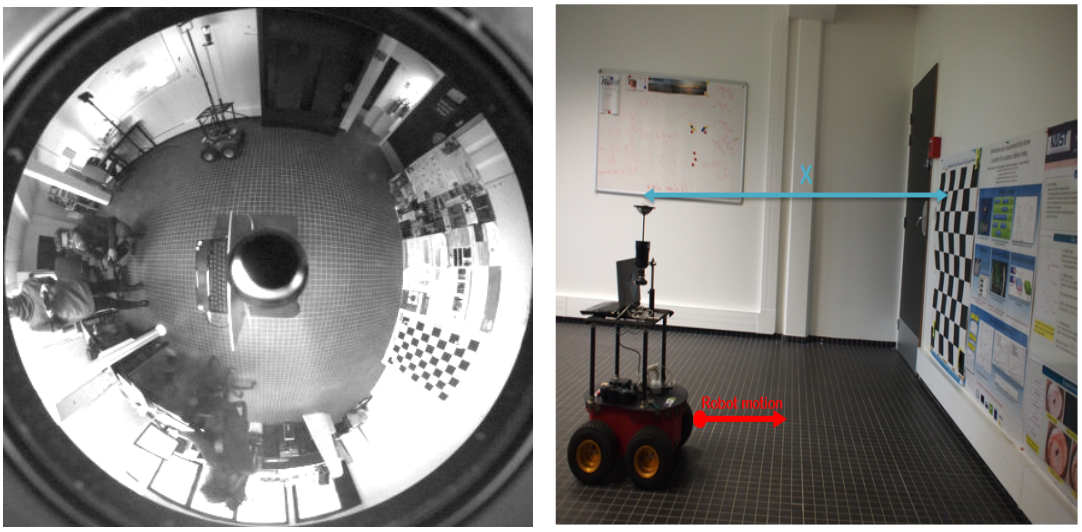


FIGURE 5.1 – Exemple d'une séquence de synthèse issue d'un capteur para-catadioptrique virtuel aux instants $t = 1, t = 11, t = 26$ et $t = 35$.



(a) (b)

FIGURE 5.2 – (a) Exemple d'image catadioptrique réelle acquise par une caméra catadioptrique. (b) Le robot pioneer 3-AT avec une caméra para-catadioptrique placée verticalement sur un mat.

Dans ce qui suit, nous validons expérimentalement les deux méthodes dans deux cas de mouvement du capteur catadioptrique. Tout d'abord nous traitons le cas du mouvement axial du capteur, ensuite nous traitons le cas général. Pour évaluer quantitativement les résultats des deux approches, nous avons utilisé l'estimateur suivant :

$$ARE = \left|\frac{\tau - \tau_r}{\tau_r}\right| \tag{5.1}$$

ARE représente l'erreur relative entre le TAC mesuré et le TAC_r réel. Cet estimateur

TABLE 5.1 – Caractéristiques des capteurs para-catadioptriques virtuel et réel.

Capteur	Para-catadioptrique virtuel	Para-catadioptrique réel
Résolution	200×200	440×440
Miroir paraboloïde	$h = 2.5$	$h = 1$
Paramètres de calibrage	avec le logiciel Povray	avec le logiciel Hyscas
Point principal (u_0, v_0)	(100,100)	(219.87,219.14)
Paramètre α	40	151.53

TABLE 5.2 – Caractéristiques des capteurs hyper-catadioptriques virtuel et réel.

Capteur	Hyper-catadioptrique virtuel	Hyper-catadioptrique réel
Résolution	200×200	640×512
Miroir hyperboloïde	$e = 3.65$	$e = 2.34$
Paramètres de calibrage	avec le logiciel Povray	avec le logiciel Hyscas
Point principal (u_0, v_0)	(100,100)	(340,266)
Paramètre α	363.63	175.63

d'erreur est calculé pour chaque pixel de l'image, ce qui permet d'obtenir une carte d'erreur du TAC pour laméthode locale et une courbe pour la méthode globale.
Afin d'évaluer l'erreur d'estimation sur le TAC pour chaque image de la séquence, nous calculons la moyenne de la carte ARE de la manière suivante :

$$\overline{ARE} = \frac{1}{N} \sum_{(i,j)} ARE(i,j) \tag{5.2}$$

N est le nombre de pixels (i, j) dans l'image.
Enfin, pour calculer l'erreur globale pour toute la séquence, nous utilisons l'estimateur $\widehat{ARE}$ qui est la moyenne de $\overline{ARE}$ défini par l'équation suivante :

$$\widehat{ARE} = \frac{1}{M} \sum_{k=1}^{M} \overline{ARE}(k) \tag{5.3}$$

où M est la taille de la séquence d'image k.

5.2.2 Mouvement axial du capteur catadioptrique vers l'obstacle

Dans cette section nous présentons les résultats dans le cas où le mouvement du capteur catadioptrique est aligné avec le mouvement du robot (mouvement axial) et ceci dans le cas des deux méthodes : locale et globale présentées dans le chapitre 4.

5.2.2.1 Méthode globale basée sur les gradients d'image : méthode MGG

Nous présentons dans cette section, les résultats des simulations obtenus en implémentant la méthode MGG. Nous rappelons que cette méthode adapte la méthode de Horn et al. [Horn 2009] pour les images para-catadioptriques.
Plusieurs situations ont été imaginées pour le mouvement du robot vers différentes orientations des surfaces planes. Dans cette section nous présentons deux situations dans le cas du mouvement axial du capteur, la première vers une surface plane perpendiculaire à l'axe de déplacement et la seconde vers une surface plane inclinée (voir la Fig. 5.3).

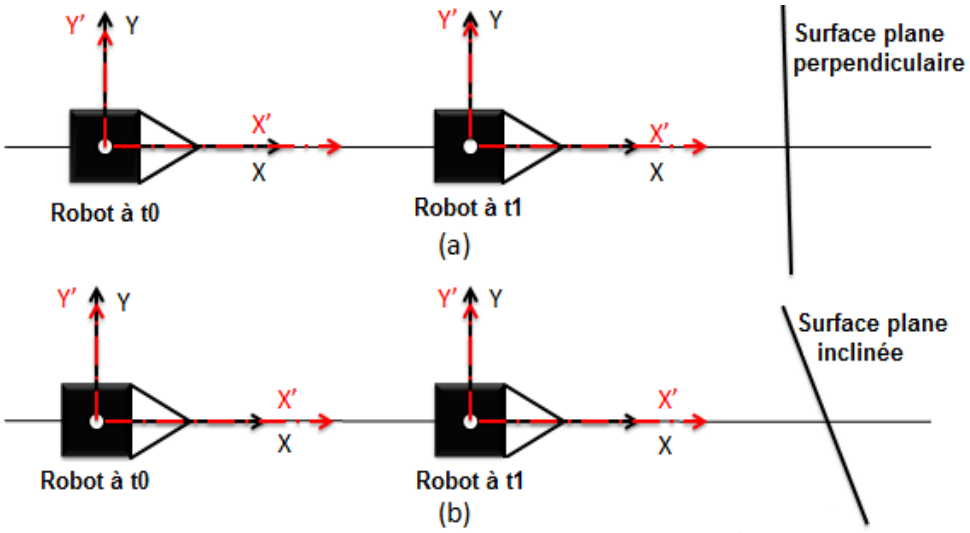


FIGURE 5.3 – Deux situations de mouvement du capteur para-catadioptrique (X, Y) embarqué sur un robot (X', Y'). (a) Translation le long de l'axe **X** relativement à une surface plane perpendiculaire à l'axe **X**. (b) Translation le long de l'axe **X** relativement à une surface plane d'orientation quelconque.

Nous validons expérimentalement cette approche tout d'abord avec des images de synthèse ensuite avec des images réelles.

A. Séquence de synthèse

Nous générons des séquences de synthèse avec le logiciel POVRAY en faisant bouger la caméra para-catadioptrique virtuelle vers des plans texturés. Les paramètres intrinsèques de la caméra para-catadioptrique virtuelle sont donnés dans la table 5.1.

La table 5.3 présente les caractéristiques des différentes séquences de synthèse utilisées

TABLE 5.3 – Caractéristiques des séquences de synthèse utilisées pour la MGG dans le cas de la translation axiale du capteur.

Séquence	miroir	X_0	$(\dot{X}, \dot{Y}, \dot{Z})$	$(-1, a, b)$	Angle $\lvert\widehat{\rho}\rvert$	σ_n
seq 1	paraboloïde	200	(1,0,0)	$(-1, 0, 0)$	0	0
seq 2	paraboloïde	200	(1,0,0)	(-1,-0.17,0.08)	0	0
seq 3	paraboloïde	200	(1,0,0)	(-1,-0.36,0.17)	0	0
seq 4	paraboloïde	200	(1,0,0)	(-1,-0.61,0.36)	0	0
seq 5	paraboloïde	50	(1,0,0)	$(-1, 0, 0)$	0	0
seq 6	paraboloïde	50	(1,0,0)	$(-1, 0, 0)$	0	6
seq 7	paraboloïde	50	(1,0,0)	$(-1, 0, 0)$	0	9
seq 8	paraboloïde	50	(1,0,0)	$(-1, 0, 0)$	0	11
seq 9	paraboloïde	50	(1,0,0)	$(-1, 0, 0)$	0	12
seq 10	paraboloïde	50	(1,0,0)	$(-1, 0, 0)$	0	14

pour valider la MGG. La distance X_0 représente la distance initiale entre le robot et la surface le long de l'axe $\mathbf{X}$, $(\dot{X}, \dot{Y}, \dot{Z})$ représente la vitesse du capteur, $(-1, a, b)$ est le vecteur normal de la surface plane, $\widehat{\rho}$ est l'angle de rotation estimé dans l'image et σ_n est l'écart type du bruit blanc ajouté à l'image. Notons que les distances sont exprimées en mètres, le temps est en unité d'images et l'angle est en degrés. Les séquences seq1, seq2, seq3 et seq4 sont composées d'une seule surface plane texturée par un damier. Dans ce cas la distance initiale X_0 entre le capteur et la surface plane est égale à $200m$ (voir la première ligne de la Fig. 5.4). Elles sont utilisées pour tester le modèle (4.28) avec différentes inclinaisons de la surface plane.

La séquence seq5 est composée de quatre plans séparés de 50m, deux plans parmi ces quatres sont texturés (voir la Fig. 5.1).
Les séquences seq6, seq7, seq8, seq9 et seq10 sont générées à partir de la séquence seq5, en ajoutant un bruit blanc d'écart type $\sigma_n = 6, 9, 11, 12$ et 14 respectivement.

Afin de tester la robustesse de la méthode basée sur les gradients d'image, nous avons effectué deux tests en variant les paramètres suivants :

— Les composantes du vecteur normal $(-1, a, b)$ de la surface plane en vue de tester l'influence de l'inclinaison de la surface sur l'estimation du TAC avec le modèle (4.28).

— L'écart type σ_n du bruit blanc ajouté aux images.

Le résultat de la simulation du premier test est montré à la figure 5.4. La vérité terrain est représentée par la courbe verte, le TAC calculé par la courbe rouge. Nous remarquons que pour les quatre séquences, l'algorithme donne de bonnes estimations du TAC même dans le cas d'une inclinaison forte de la surface plane.

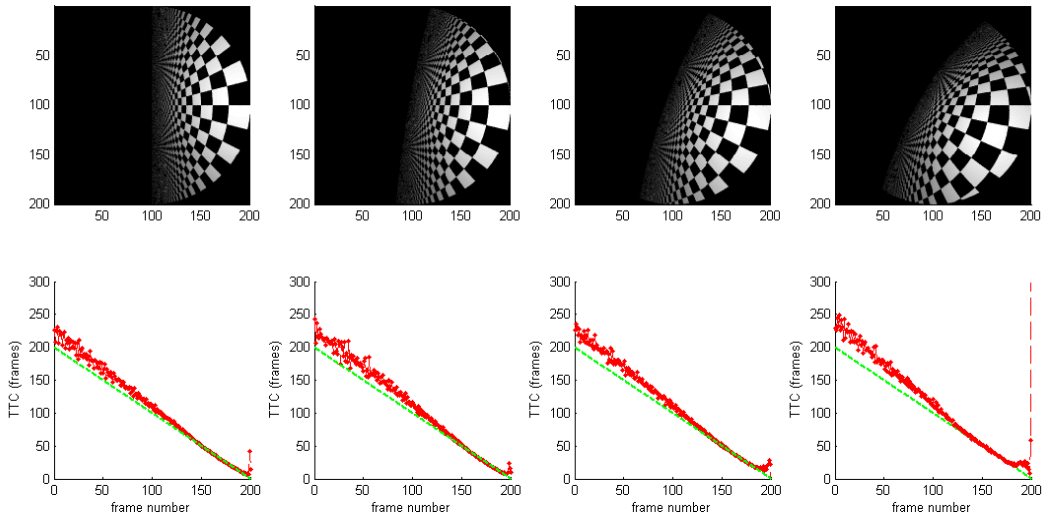


FIGURE 5.4 – Estimation du TAC. La première rangée montre respectivement l'image $n°1$ des séquences seq1, seq2, seq 3 et seq4 presentées dans le tableau 5.3. La deuxième ligne montre la courbe du TAC estimé (en rouge) et la vérité terrain (en vert) pour chaque séquence.

Le resultat du deuxième test est illustré dans la figure 5.5. Les erreurs ARE de la seq6 ($\sigma_n = 6$) varient entre 10 % et 35%, elles représentent des valeurs acceptables d'estimation du TAC. Pour la seq7 ($\sigma_n = 9$), ces valeurs oscillent entre 15% et 46%, alors que dans le cas des séquence seq8 ($\sigma_n = 11$) et seq9 ($\sigma_n = 12$), les valeurs d'ARE varient autour de 30% et elles sont élevées au début et à la fin de la séquence. Enfin pour la seq10 ($\sigma_n = 14$) les valeurs d'ARE sont élévées ($> 30\%$) le long de la séquence. Donc, les valeurs ARE de chaque séquence croissent avec le bruit, en particulier, au début des séquences à cause de la perturbation d'estimation des gradients et vers la fin de celles-ci à cause de l'aliasing temporel [2]. Au milieu des séquences, les valeurs des ARE en présence du bruit sont bonnes sauf pour le cas de la seq10. Enfin, nous constatons que plus les images sont fortement bruitées, plus le calcul du TAC est erroné.

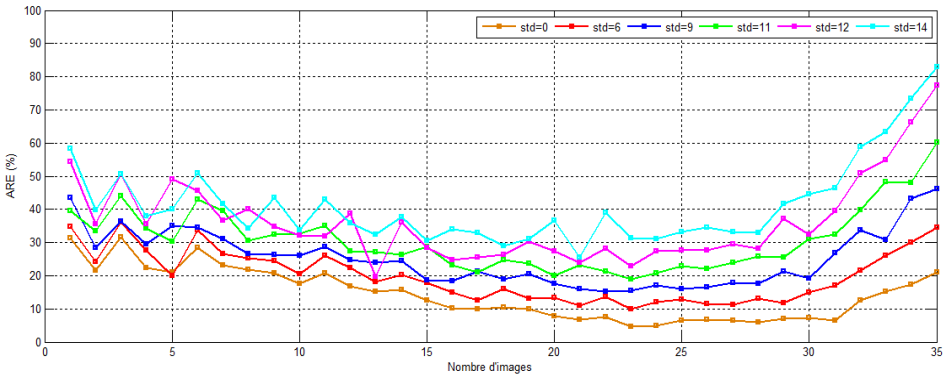


FIGURE 5.5 – L'erreur ARE sur le TAC par la méthode MGG en présence de bruit ajouté aux images para-catadioptriques avec $\sigma_n = \{0, 6, 9, 11, 12, 14\}$.

Dans tout ce qui précède, nous avons validé l'estimation du TAC pour la surface qui se trouve devant le capteur paracatadioptique (voir la Fig. 5.6). L'estimation du TAC en vision catadioptrique nous permet d'estimer le TAC même pour les obstacles qui se trouvent derrière le capteur, ce que les capteurs perspectifs ne fournissent pas.
Les résultats de cette double estimation sont montrés à Fig. 5.7. La résolution de l'équation (4.28) produit de bonne estimation du TAC entre l'image $n°1$ et l'image $n°35$. Le TAC calculé pour la surface de devant (S_r) nommé τ_{S_r} est décroissant le long de la séquence. Après l'image $n°35$ la courbe du TAC présente des aberrations significatives à cause du phénomène de l'aliasing temporel.

2. Erreur due à l'approximation de la dérivée temporelle de l'image par une différence finie.

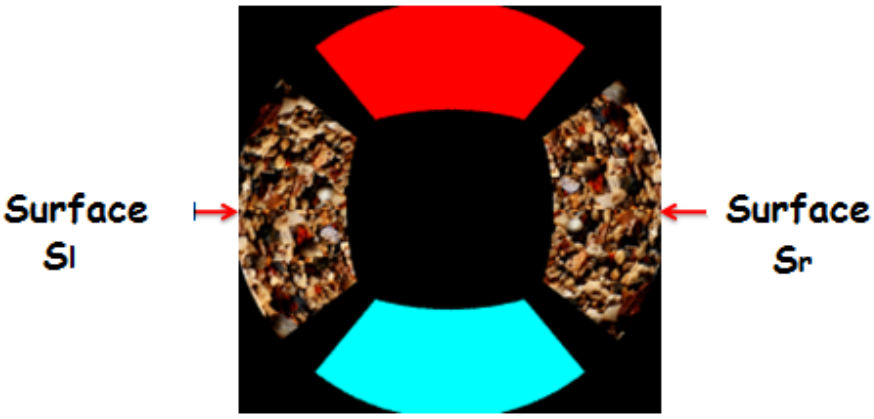


FIGURE 5.6 – Surfaces planes à droite et à gauche du capteur virtuel dans l'image para-catadioptrique (seq5)

A l'opposé, le TAC calculé pour la région (S_l) nommé τ_{S_l} est une courbe croissante, ceci est évident car le capteur s'éloigne de la surface (S_l). Dans ce cas, l'erreur ARE du τ_{S_l} est élevée en comparaison avec τ_{S_r} de l'image $n°1$ à l'image $n°35$. Cependant, après l'image $n°35$, l'ARE pour le τ_{S_l} est basse par rapport à celle du τ_{S_r}.
Nous avons déjà remarqué que l'estimation du TAC à la fin de la séquence devient difficile à cause des phénomènes signalés ci-haut. Pour remédier à ce problème, nous profitons de l'estimation du τ_{S_l} pour améliorer l'estimation du τ_{S_r}. En effet, nous avons la relation suivante :

$$\tau_{S_r} + \tau_{S_l} = \nu \tag{5.4}$$

avec ν est une constante. Donc nous pouvons déduire la nouvelle expression du TAC dans la direction du mouvement du robot par :

$$\tau = \frac{\widehat{\tau_{S_r}} + (\nu - \widehat{\tau_{S_l}})}{2} \tag{5.5}$$

La constante ν peut être calculé automatiquement le long de déplacement du capteur. Le résultat de cette nouvelle expression du TAC est montré dans la figure 5.7.

B. Séquence réelle

De la même manière que dans les exemples de synthèse, nous validons le critère (4.28) sur deux séquences réelles, la première avec une surface plane perpendiculaire (voir la Fig. 5.9a) et la seconde avec une surface plane inclinée autour de son axe **Z** avec un angle $\beta = 20°$ (voir la Fig. 5.9b). Les caractéristiques de ces deux séquences sont presentées dans la table. 5.4.

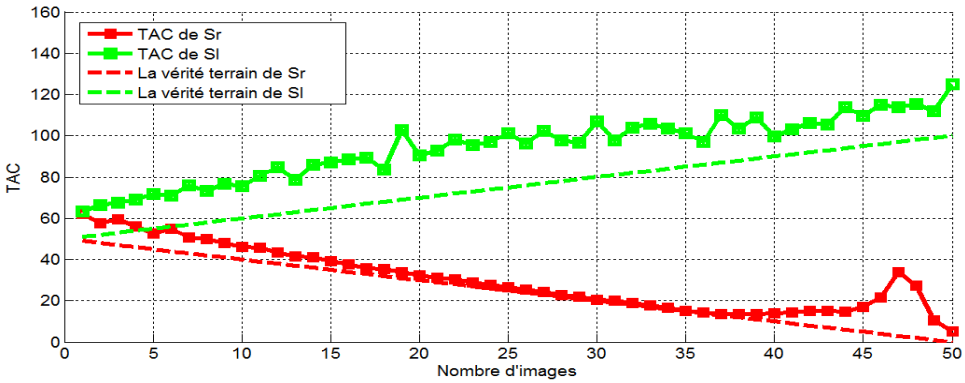


FIGURE 5.7 – Estimation du TAC avec la séquence para-catadioptrique (seq5) des obstacles (S_r) et (S_l). La courbe rouge illustre le TAC estimé de la surface plane (S_r). La courbe bleue montre le TAC estimé de la surface plane (S_l). La vérité terrain du τ_{S_l} est illustrée en bleu pointillé et la courbe rouge pointillé représente la vérité terrain du τ_{S_r}.

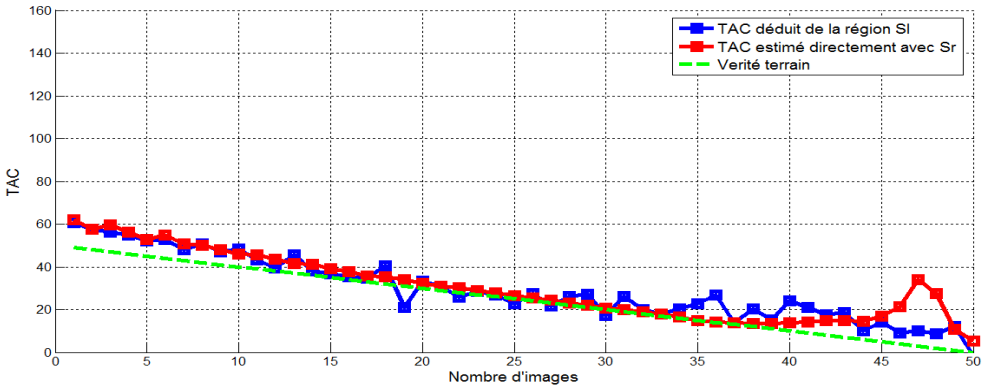
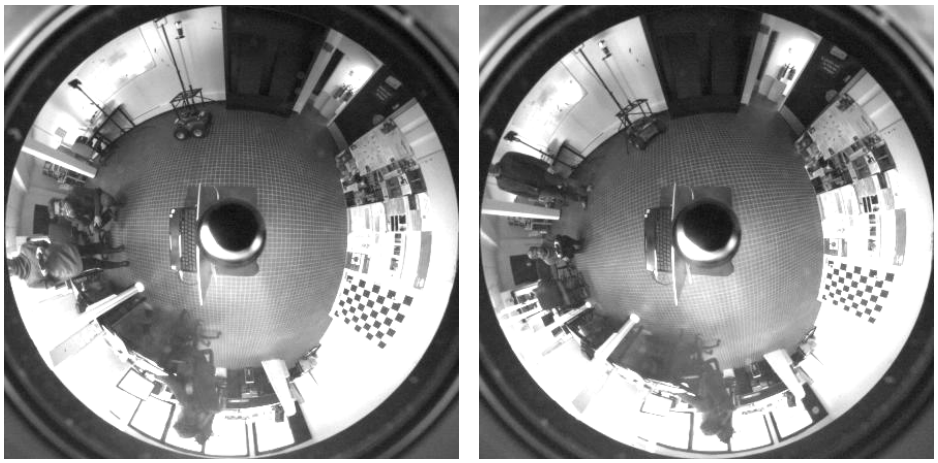


FIGURE 5.8 – Résultats de comparaison entre τ_{S_r} calculé pour la surface plane (S_r) avec le TAC déduit de τ_{S_l} $(\tau = \nu - \tau_{S_l})$

TABLE 5.4 – Description des séquences réelles utilisées dans les expérimentations.

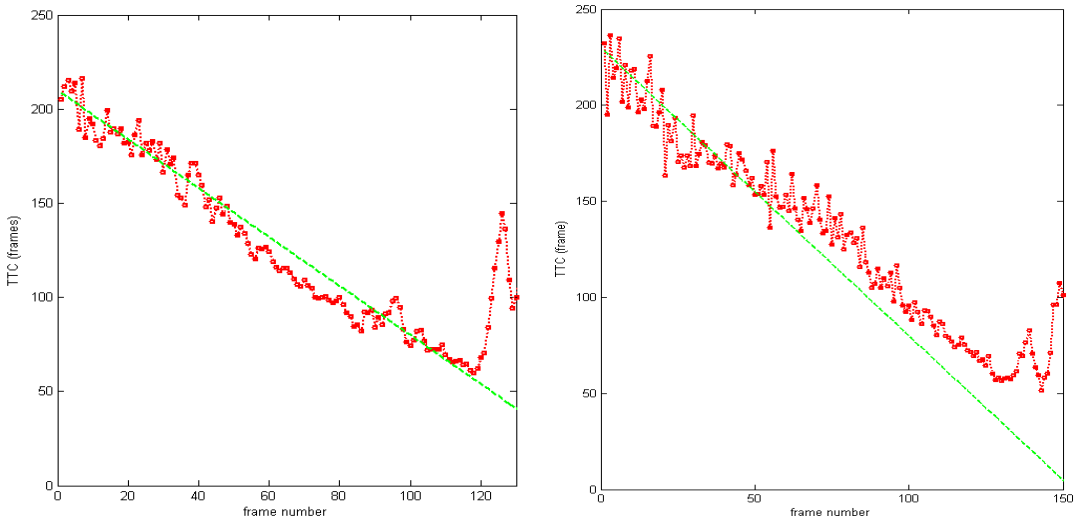
Séquence	Surface plane (obstacle)	(α, β)	Vecteur normal $(-1, a, b)$
seq 11 (voir Fig. 5.9a)	Perpendiculaire	$(0°, 0°)$	$(-1, 0, 0)$
seq 12 (voir Fig. 5.9b)	Inclinée	$(0°, 20°)$	$(-1, 0.36, 0)$

La figure. 5.10 contient les courbes de TAC estimés pour ces deux séquences réelles. Les résultats des simulations (courbes rouges) sont proches des vérités terrain (courbes vertes). Quelques fluctuations perturbent l'estimation du TAC au cours du déplacement du robot,



(a) (b)

FIGURE 5.9 – Deux images des séquences réelles traitées dans les expérimentations (voir la Table. 5.4) : (a) Image $n°1$ de la seq 11. (b) Image $n°1$ de la seq12.



(a) (b)

FIGURE 5.10 – Estimation du TAC dans le cas réel pour les séquences décrites dans la table 5.4, la courbe rouge fournit l'estimation du TAC et la courbe vertecorrespond à la vérité terrain : (a) TAC estimé pour la seq 11, (b) TAC estimé pour la seq 12.

ceci est dû au fait que la vitesse du robot n'est pas toujours constante.
La vérité terrain est obtenue à partir de l'estimation de la position de l'objet (la mire par exemple Fig. 5.11) par rapport au repère du monde (repère de la caméra) en utilisant les

mesures 2D dans l'image et leurs correspondances 3D.

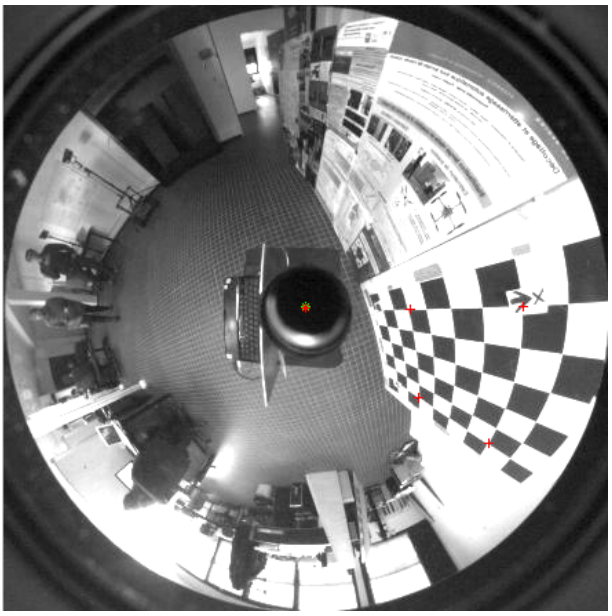


FIGURE 5.11 – Suivi des points 3D de la mire (marqués en rouge) au cour du déplacement du robot vers la surface plane (la surface plane qui inclut la mire), afin d'exprimer ses coordonnées dans le repère de la caméra para-catadioptrique.

Dans ce cas, le calcul repose sur la connaissance *a priori* d'une structure 3D, comme par exemple une mire dont les sommets sont connus. Pour calculer la matrice homogène du changement du repère de la mire vers le repère de la caméra (repère du monde), il faut estimer les paramètres extrinsèques du capteur para-catadioptrique calibré. Ce problème d'estimation a été abordé à l'aide d'une technique d'optimisation non linéaire basée sur l'algorithme de Levenberg-Marquardt (LM). L'estimation des coordonnées du point P dans le repère caméra nous permet de calculer la distance instantanée entre le robot et les points de la mire (surface plane). Sachant que la vitesse du robot est maintenue constante au cours du déplacement, le calcul du TAC réel sera déduit du calcul de la distance réelle à une échelle près (voir les courbes vertes dans les figures 5.10a, 5.10b, et 5.31).
Nous décrivons dans la suite les résultats expérimentaux de la méthode basée sur l'estimation du flot optique ou aussi appellée MLFO. Cette méthode est valide pour tout capteur catadioptrique à PVU et valable pour toute forme géométrique d'obstacle. Cela en fait donne une méthode générale d'estimation du TAC.

5.2.2.2 MLFO basée sur le flot optique : MLFO

Nous validons expérimentalement l'approche basée sur l'estimation du flot optique, tout d'abord avec des images de synthèse, puis avec des images réelles. Nous avons déjà montré que la MLFO est valable pour tous les capteurs à PVU. Dans les expérimentations que nous présentons dans ce chapitre, nous allons nous limiter aux capteurs para-catadioptriques et hyper-catadioptriques.

A. Séquence de synthèse

Pour tester expérimentalement la MLFO, nous avons utilisé un ensemble de séquences de synthèse hyper-catadioptriques et para-catadioptriques (voir la table 5.5). Les paramètres intrinsèques de la caméra virtuelle sont donnés dans les tables 5.1 et 5.2.

TABLE 5.5 – Caractéristiques des séquences de synthèse utilisées dans le cas de la MLFO avec translation axiale du capteur.

Séquence	miroir	X_0	$(\dot{X}, \dot{Y}, \dot{Z})$	Angle $\lvert\widehat{\rho}\rvert$	σ_n
seq 13	paraboloïde	50	(1,0,0)	0°	0
seq 14	hyperboloïde	50	(1,0,0)	0°	0
seq 15	hyperboloïde	50	(1,0,0)	0°	6
seq 16	hyperboloïde	50	(1,0,0)	0°	9
seq 17	hyperboloïde	50	(1,0,0)	0°	11
seq 18	hyperboloïde	50	(1,0,0)	0°	12
seq 19	hyperboloïde	50	(1,0,0)	0°	14

Nous présentons les résultats du TAC en utilisant l'équation (4.38). Afin d'évaluer la robustesse de la méthode, l'analyse a été effectuée en variant :

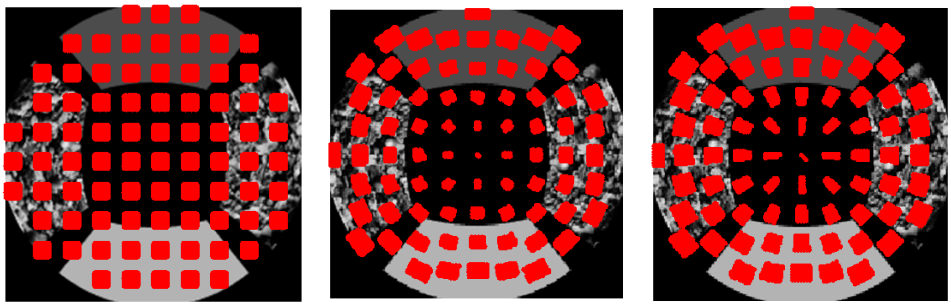
— La nature de l'image catadioptrique (para-catadioptrique ou hyper-catadioptrique).
— L'écart type (σ_n) du bruit blanc ajouté aux images.

La première expérimentation est réalisée à l'aide des séquences seq13 et seq14. L'objectif principal de ce premier test est de prouver que la relation (4.38) est valable pour tout

capteur catadioptrique à PVU [3] et aussi de produire une carte du TAC basée sur l'estimation du flot optique. Nous avons retenu la méthode de Lucas-Kanade qui est toujours l'une des solutions les plus utilisées dans les applications robotiques pour l'estimation du mouvement [Barron 1994]. Comme nous l'avons signalé au chapitre 3, section 3.2.2.2, Radgui et al. [Radgui 2008] proposent une adaptation de cette méthode pour les images para-catadioptriques. En effet, à cause de la distorsion introduite par le miroir, le champ de mouvement ne peut pas être considéré comme constant dans le voisinage. Pour cela, ces auteurs proposent d'adapter la méthode de Lucas-Kanade en remplaçant le modèle constant par le modèle quadratique et le voisinage classique par un voisinage adapté. De la même façon, nous avons proposé d'adapter la méthode de Lucas-Kanade sur les images hyper-catadioptriques en utilisant le modèle quadratique (3.17) et en remplaçant le voisinage classique par un voisinage adapté basé sur l'équation du miroir hyperboloïde.
Nous présentons des résultats quantitatifs et qualitatifs relatifs au modèle (4.38), calculés en utilisant le flot optique de Lucas-Kanade (TLK) et en utilisant le flot optique adapté (TALK). En termes de qualité des résultats obtenus, nous présentons des cartes de TAC avec TLK et TALK à figure 5.13. A noter que sur ces figures, la couleur bleu foncé indique une valeur faible du TAC (capteur virtuel proche du plan), un rouge clair désigne une valeur élevée du TAC (capteur virtuel loin du plan) et la couleur rouge foncé désigne les valeurs aberrantes alors que la couleur bleu extrêmement foncé indique les TAC nuls.

Notons que l'objectif principal de ce travail est l'adaptation du TAC en vision catadioptrique. Nous utilisons deux méthodes de calcul du flot optique pour vérifier son influence sur l'estimation du TAC. Il ne s'agit pas d'une comparaison de différentes techniques de flot optique.
Au niveau des résultats quantitatifs obtenus, nous avons utilisé les mesures d'erreur (5.1), (5.2) et (5.3) pour valider les méthodes de flot optique TLK et TALK.
La relation (4.38) fournit la valeur du TAC pour chaque pixel. La figure 5.12 montre le voisinage $\vartheta 1$ de taille $9 * 9$ utilisé pour la méthode TLK (voir Fig. 5.12a). Le voisinage adapté $\vartheta 2$ pour les images para-catadioptriques est fixé par $d\theta = \frac{\pi}{40}$ et $d\varphi = \frac{\pi}{60}$ (voir Fig. 5.12b) et le voisinage $\vartheta 3$ pour les images hyper-catadioptriques est fixé par $d\theta = \frac{\pi}{40}$ et $d\varphi = \frac{\pi}{180}$ (voir Fig. 5.12c).

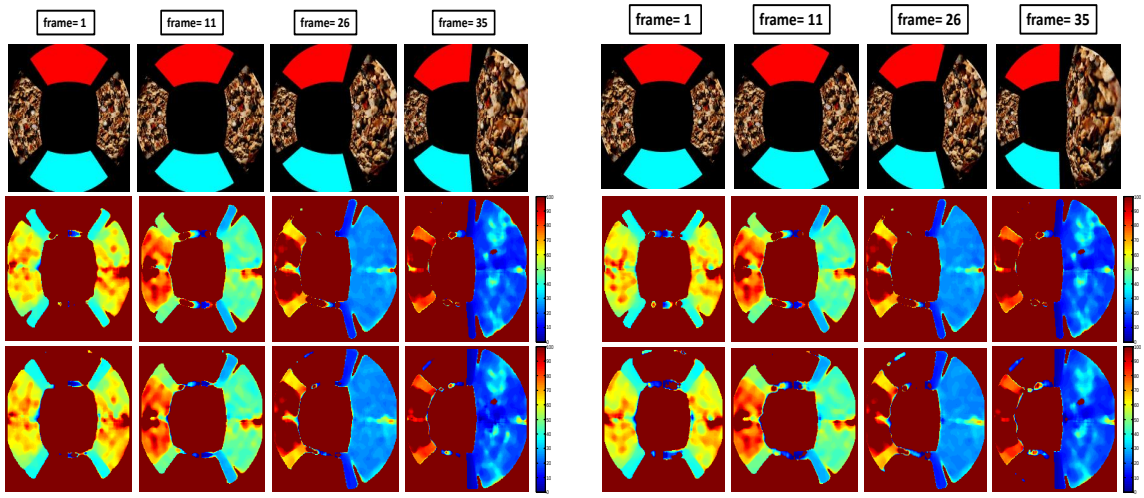
3. Dans ce manuscrit, nous nous limitons aux capteurs para-catadioptriques et hyper-catadioptriques car ces deux capteurs sont les plus utilisés dans la vision catadioptrique.



(a) (b) (c)

FIGURE 5.12 – Voisinages utilisés pour le calcul des flots optiques sur les images catadioptriques. (a) Voisinage classique $\vartheta 1$. (b) : Voisinage adapté $\vartheta 2$ dans le cas d'un miroir paraboloïde. (c) :Voisinage adapté $\vartheta 3$ dans le cas d'un miroir hyperboloïde.

En pratique, le calcul du TAC s'effectue à partir d'une estimation du flot optique qui n'est pas toujours précise (à cause du mauvais conditionnement du système numérique dans des régions peu texturées dans l'image ou à cause de l'aliasing temporel...). Pour diminuer l'effet des valeurs aberrantes, nous avons utilisé un estimateur robuste médian dans un voisinage de taille $9 * 9$ pour chaque pixel de l'image. La figure 5.13 présente les cartes d'estimation du TAC avec les méthodes de flot optique TLK et TALK pour les séquences de synthèse para-catadioptriques et hyper-catadioptriques.
Visuellement, nous remarquons que les deux méthodes conduisent à de bons résultats. Les couleurs des cartes varient de la couleur rouge clair (TAC élévé) à la couleur bleu foncé (TAC faible) quand le capteur virtuel s'approche de la surface à droite. De même, les valeurs du TAC des deux séquences augmentent quand le capteur s'éloigne de la surface de gauche (les couleurs des cartes passent de la couleur jaune à la couleur rouge). Ces constations montrent que l'algorithme donne de bons résultats d'estimation du TAC avec les capteurs à PVU.
Une quantification d'erreur locale des cartes du TAC est donnée dans la figure 5.14. La couleur de chaque point est calculée par la mesure d'erreur ARE (eq. (5.1)). La couleur la plus répandue sur toutes les cartes est la couleur bleu foncé. Cela signifie que les erreurs d'estimation du TAC sont faibles le long de la translation du capteur virtuel. Néanmoins, à la fin de la séquence, l'erreur sur le TAC de plusieurs régions dans l'image pour les deux séquences (seq13 et seq14) devient élevée. Ceci peut être justifié par le phénomène de l'aliasing temporel, ou à cause de la déformation de l'objet proche du capteur, qui engendre



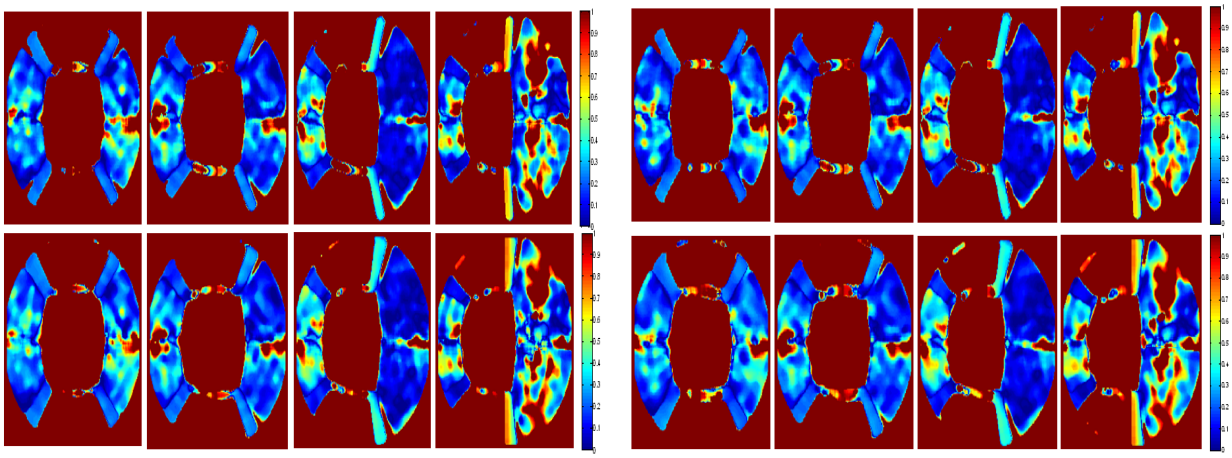
(a) (b)

FIGURE 5.13 – Estimation du TAC sur les séquences de synthèse seq13 et seq14. La première rangée illustre la progression au cours du temps de quatre images des deux séquences. La deuxième rangée donne les cartes de TAC en utilisant le flot optique classique TLK alors que la troisième rangée contient les cartes de TAC en utilisant le flot optique adapté TALK. (a) : Séquence para-catadioptrique. (b) : Séquence hyper-catadioptrique.

un manque de texture. Dans le cas d'une surface faiblement texturée, l'estimation du flot optique est en général entachée d'erreurs à cause de la mauvaise estimation des gradients spatiaux d'intensité.

Notons que dans toutes ces expérimentations, nous ne prenons pas en considération les points de l'image avec $ARE > 0.8$. Dans ces endroits, les deux méthodes génèrent des erreurs importantes. En dehors de ces endroits, notre algorithme du calcul du TAC prouve sa robustesse sur les deux types d'images catadioptriques. En effet, en tenant compte des résultats des figures 5.14 et 5.15, nous pouvons constater que, si nous nous limitons à la courbe avant l'aliasing temporel (ou au manque de texture à la fin de la séquence) qui débute vers la 35ème image, l'erreur varie entre 20% et 30%. Ces valeurs sont comparables à celles trouvées dans la littérature relative au cas des caméras perspectives [Galbraith 2005].

Il faut cependant souligner deux points. Le premier point est lié aux singularités dues aux

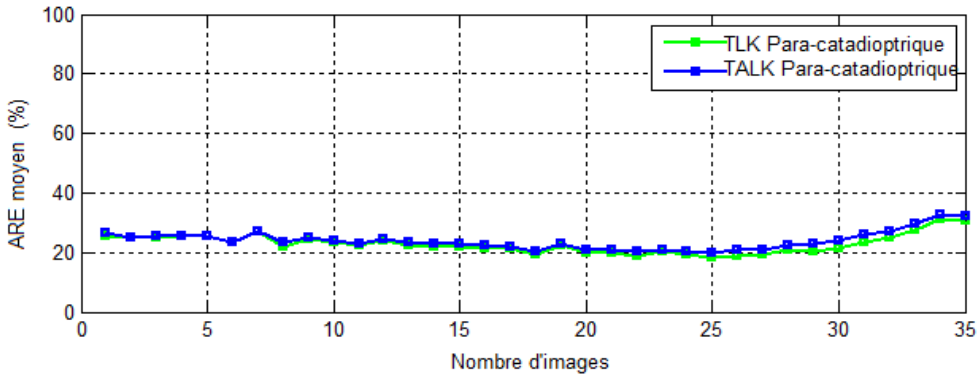


(a) (b)

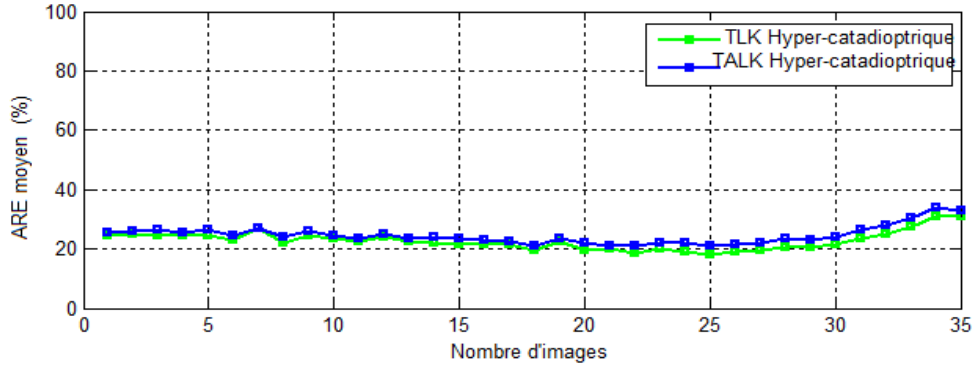
FIGURE 5.14 – Les cartes d'erreurs relatives (ARE) sur le TAC en utilisant la mesure (5.1) sur les séquences seq 13 et seq14. La première rangée contient l'erreur ARE pour la méthode TLK. Seconde rangée montre l'erreur ARE pour la méthode TALK. (a) : Séquence para-catadioptrique. (b) : Séquence hyper-catadioptrique.

égalités suivantes ($u - u_0 = 0$ et/ou $v - v_0 = 0$) qui rendent le modèle indéfini. Quant au second point à souligner, il concerne le foyer d'expansion (FOE). Nous constatons à partir des cartes de TAC présentées dans les figures 5.13 et 5.14 que loin du FOE, l'estimation avec TLK et TALK donne de bons résultats, alors que, proche du FOE l'estimation n'est pas correcte. Enfin, avec les deux séquences de synthèse para-catadioptrique et hyper-catadioptrique, les résultats sont bons aussi bien avec le calcul du TAC basé sur le flot optique classique TLK qu'avec celui basé sur le flot adapté TALK. Les résultats des deux estimateurs sont presque identiques. Donc, dans tout ce qui suit, nous choisissons d'utiliser le flot classique de Lucas-Kanade pour estimer le TAC avec la relation (4.38).

La deuxième expérimentation (voir Fig. 5.16) est effectuée pour évaluer notre algorithme de calcul du TAC en fonction du bruit. Nous ajoutons un bruit blanc gaussien à la séquence seq14 avec différents écarts type $\sigma_n = \{0, 6, 9, 11, 12, 14\}$. De même, nous générons les séquences : seq15 avec l'ajout d'un bruit d'écart type $\sigma_n = 6$, seq16 avec $\sigma_n = 9$, seq17 avec $\sigma_n = 11$, seq18 avec $\sigma_n = 12$ et seq19 avec $\sigma_n = 14$.
La figure 5.16 montre un bon comportement des courbes de TLK en présence de bruit pour $\sigma_n \in \{0, 6, 9, 11\}$, mais dès que $\sigma_n > 11$ nous remarquons une mauvaise estimation



(a)



(b)

FIGURE 5.15 – La moyenne de l'erreur ARE ($\overline{ARE}$) du TAC dans le cas du mouvement axial du capteur virtuel (séquences utilisées sont seq13 et seq14). La courbe verte indique le ($\overline{ARE}$) pour la méthode TLK et la courbe bleue indique le ($\overline{ARE}$) du TALK. (a) : séquence para-catadioptrique. (b) : séquence hyper-catadioptrique.

du TAC.

Nous avons montré avec des séquences de synthèse que la relation (4.38) permet d'estimer le TAC convenablement. Dans la section suivante, nous faisons de nouveaux tests dede notre méthode de calcul du TAC avec différentes séquences réelles.

B. Séquences réelles

De la même manière, nous faisons des expérimentations sur des séquences vidéos réelles dans le cas où le mouvement du capteur est axial le long de l'axe **X**. La plateforme utilisée est décrite dans la section 5.2.1. Nous avons réalisé des expériences avec deux types de capteurs (Fig. 5.2) :

— Capteur para-catadioptrique composé d'un miroir paraboloïde couplé à une lentille

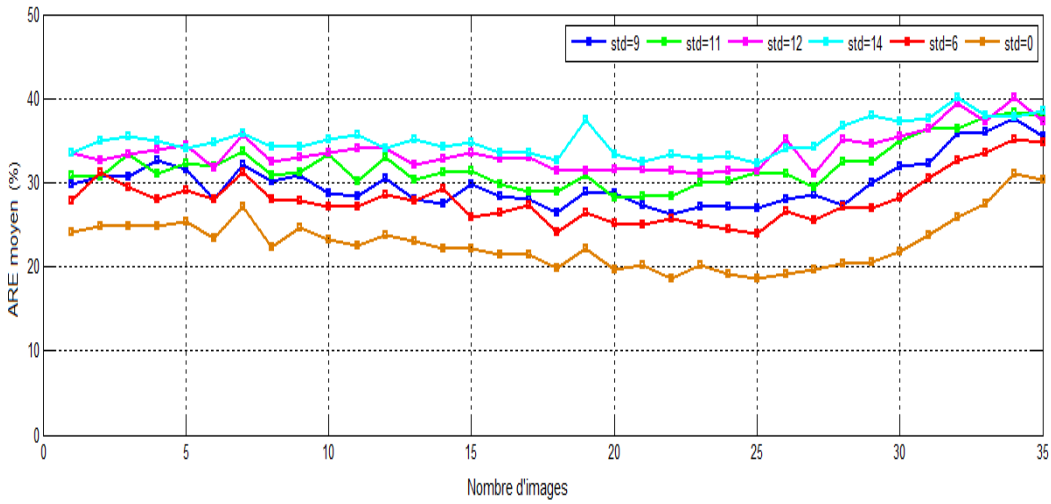


FIGURE 5.16 – L'erreur relative sur le TAC ($\overline{ARE}$) pour la séquence hyper-catadioptrique bruitée avec un bruit blanc de différents écarts-type $\sigma_n \in \{0, 6, 9, 11, 12, 14\}$.

orthographique et une caméra Sony Uye.

— Capteur hyper-catadioptrique composé d'un miroir hyperboloïde couplé à une caméra CCD Sony Uye.

Les caractéristiques des vidéos réelles acquises avec ces deux capteurs sont regroupées dans la table 5.6. La séquence seq11 est la même séquence para-catadioptrique utilisée dans les expérimentations de la MGG, alors que la seq20 est une séquence hyper-catadioptrique réelle. Les paramètres intrinsèques sont donnés dans les tables 5.1 et 5.2. Trois images

TABLE 5.6 – Séquences réelles utilisées dans les expérimentations de la méthode MLFO.

Séquence	miroir	obstacle	Angle $\lvert\widehat{\rho}\rvert$
seq 11	paraboloïde	toute la scène	0
seq 20	hyperboloïde	toute la scène	0

réelles para-catadioptriques, correspondant aux instants $t = 100$, $t = 200$ et $t = 300$, sont montrées dans la première ligne de la figure 5.17a et trois images réelles hyper-catadioptriques, correspondant aux instants $t = 100$, $t = 150$ et $t = 200$ sont affichées dans la première ligne de la figure 5.17b. De la même manière que dans le cas synthèse,

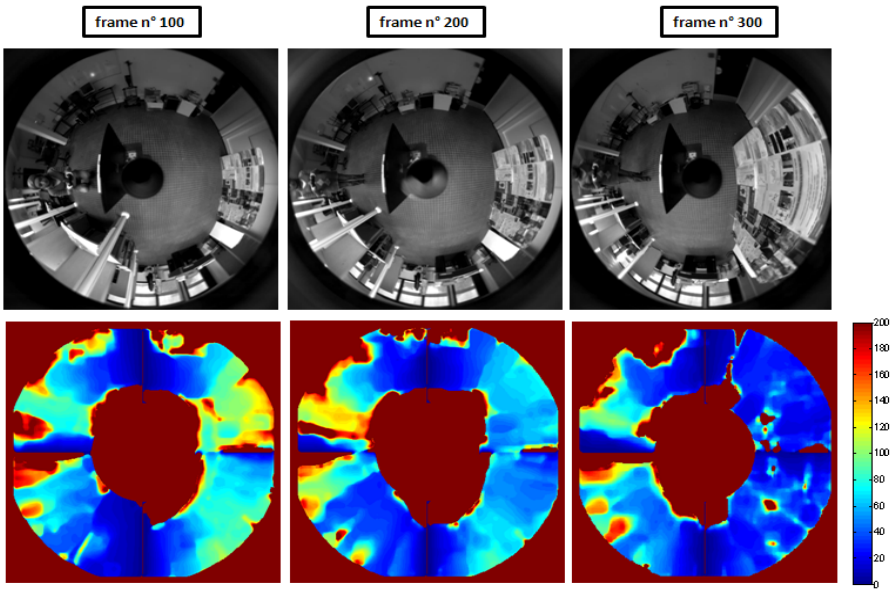
nous avons appliqué l'équation (4.38) avec le flot optique de Lucas-Kanade (deuxième rangée de la figure). A noter que sur cette figure, le bleu foncé indique une valeur faible du TAC (robot proche de l'obstacle), une rouge clair désigne une valeur élevée du TAC (robot loin de l'obstacle), le rouge foncé désigne les valeurs aberrantes alors que le bleu extrêmement foncé indique des TAC nuls.

Dans le cas des expérimentations réelles pour la MLFO, la vérité terrain n'est pas connue de manière exacte. Pour valider l'estimation locale du TAC, nous nous contentons seulement d'une analyse visuelle (qualitative) des cartes. Dans les différentes cartes du TAC de la figure 5.17, l'intérêt du TAC catadioptrique est justifié par le fait qu'il est capable de détecter la profondeur suivant l'axe **X** des objets en avant et en arrière du robot.

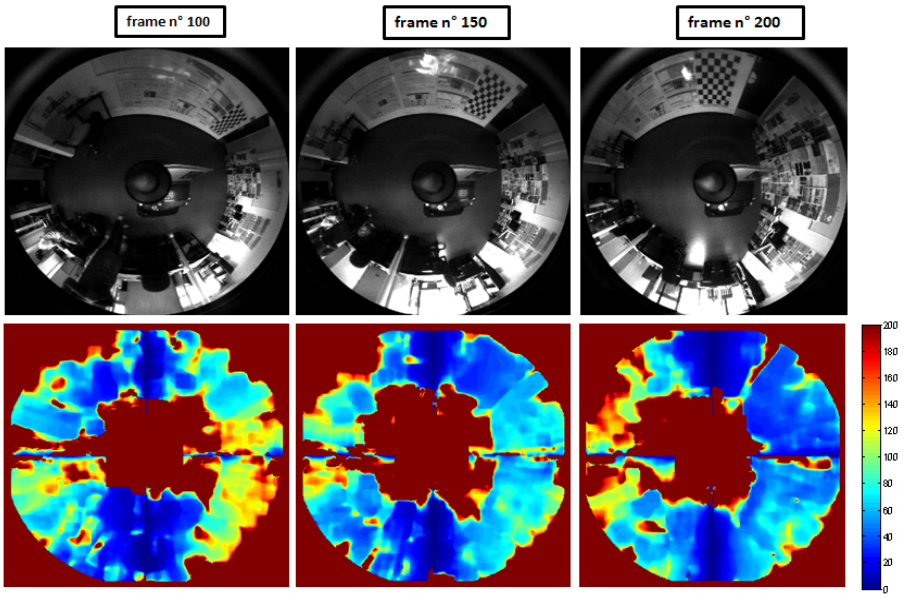
Néanmoins, quelques erreurs d'estimation du TAC sont présentes. Ces erreurs sont dues d'une part, au changement de luminosité dans la scène de l'image, ce qui provoque une violation de la contrainte de conservation lumineuse et ainsi des erreurs d'estimation du flot optique. D'autre part ces erreurs sont dues au foyer d'expansion comme nous l'avons expliqué dans la section précédente.

Pour vérifier l'intérêt de la relation (4.38) en robotique mobile, nous avons testé si celle-ci permettrait de détecter les obstacles. Pour ce faire, nous avons appliqué la méthode proposée sur une nouvelle séquence réelle qui représente un environnement inconnu avec deux obstacles fixes (voir la première rangée de la figure 5.18).

Les deux obstacles sont placés respectivement derrière et devant le robot (nous ne prenons pas en considération les objets latéraux). La deuxième rangée de la figure 5.18, contient quatre images différentes de la séquence (de gauche à droite : à $t = 150$, $t = 220$, $t = 300$ et $t = 320$). La troisième rangée représente les cartes de TAC pour les quatre images de la séquence. Le TAC de l'obstacle derrière le robot (obstacle $n°1$) croît de l'image $n°150$ à l'image $n°220$, ceci est illustré par la modification de la couleur, du bleu foncé (TAC faible) au rouge clair (TAC élevé), alors que pour le TAC de l'obstacle devant (obstacle $n°2$), sa couleur change du rouge clair (TAC élevé) au bleu foncé (TAC faible). La technique de détection des obstacles proposée ici est s'appuie sur une simple binarisation de la carte de TAC. Celle ci est effectuée en fixant le seuil sur le TAC à 30. Sur la carte binaire (figure 5.18, dernière rangée), une couleur blanche indique « obstacle détecté » alors que la couleur noire indique « obstacle non detecté ». Notons que, la binarisation n'est pas appliqué sur



(a)



(b)

FIGURE 5.17 – Estimation du TAC basée sur le flot optique TLK pour les deux séquences réelles (seq11 et seq20). La première rangée illustre la progression au cours du temps de trois images des deux séquences. La deuxième rangée donne les cartes de TAC calculées avec le flot optique classique TLK. (a) : Séquence para-catadioptrique. (b) : Séquence hyper-catadioptrique.

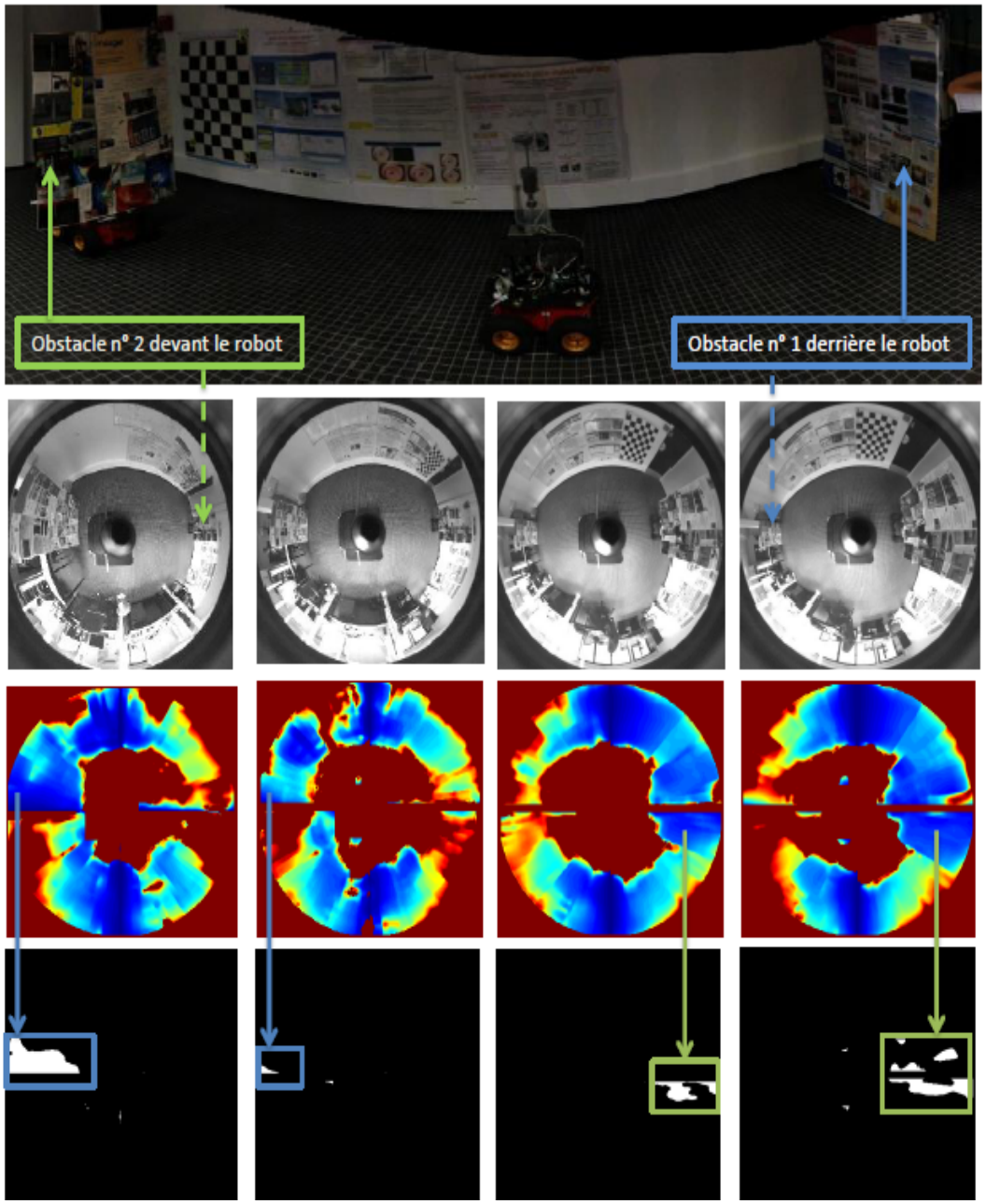


FIGURE 5.18 – Détection des obstacles basée sur la binarisation des cartes de TAC. La 1e rangée représente une vue panoramique du robot avec les deux obstacles immobiles. La 2e rangée représente l'évolution dans le temps de l'image para-catadioptrique. La 3e rangée, montre les cartes de TAC pour ces quatre images de la séquence et la 4e rangée contient les cartes binaires associées, le rectangle bleu encadre l'obstacle $n°1$ détecté et le rectangle vert encadre l'obstacle $n°2$ détecté.

les objets qui ne se trouvent pas sur l'axe de mouvement du robot. La quatrième rangée de la figure 5.18 montre l'obstacle $n°1$ encadré en bleu et l'obstacle $n°2$ encadré en vert. Au début de la séquence (image $n°150$), le robot peut avancer en avant ou s'arrêter. De l'image $n°220$ à l'image $n°300$, l'espace de navigation du robot ne présente pas de risque de collision. Cependant, après l'image $n°300$ la tache blanche de l'obstacle $n°2$ apparaît de plus en plus dans la carte binaire, ce qui montre que le robot s'approche peu à peu d'un obstacle.

5.2.3 Mouvement général du capteur catadioptrique : stratégie de ré-orientation de l'image

Dans cette sous-section, nous traitons le cas général de l'estimation du TAC pour les images catadioptriques. Nous montrons l'intérêt de la stratégie de ré-orientation dans l'estimation du TAC. Pour cela, nous proposons une généralisation de nos deux modèles (4.38) et (4.28) dans le cas d'un mouvement quelconque du capteur catadioptrique (voir la Fig. 5.19). Tout d'abord, nous appliquons les étapes de la stratégie de ré-orientation sur deux images réelles afin d'illustrer la procédure d'extraction du FOE et par suite l'estimation de l'angle $\widehat{\rho}$ de la ré-orientation. Ensuite, nous montrons expérimentalement l'amélioration apportée à l'estimation du TAC sur les images catadioptriques en faisant une comparaison entre l'estimation avec et sans la technique de ré-orientation.

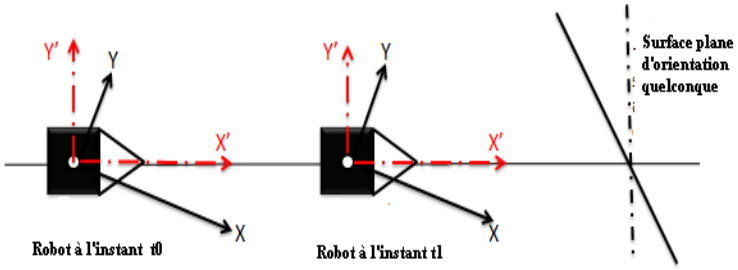
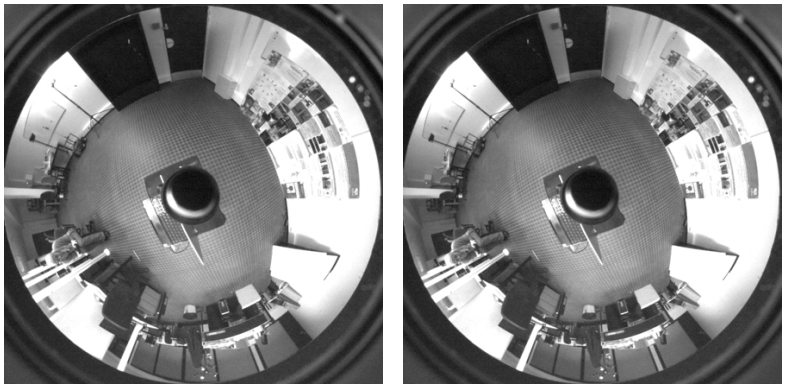


FIGURE 5.19 – Mouvement général du capteur catadioptrique embarqué sur le robot pioneer 3-AT. (O, X, Y) représente le système du repère caméra et (O, X', Y') représente le système du repère du robot.

5.2.3.1 Stratégie de ré-orientation : illustration pour l'extraction du FOE et le calcul de l'angle $\widehat{\rho}$ de la ré-orientation

Afin d'illustrer les étapes de la stratégie de ré-orientation décrite dans le chapitre 4 section 4.5, nous avons choisi de l'appliquer sur une séquence réelle. Notre objectif est d'estimer l'angle $\widehat{\rho}$ entre l'axe horizontal $\mathbf{u}$ de l'image et l'axe du FOE, afin d'effectuer une rotation de l'image d'angle $\widehat{\rho}$ et ainsi de pouvoir appliquer les relations (4.38) et (4.28) dans le cas général du mouvement du capteur catadioptrique.
La première étape de l'algorithme de ré-orientation consiste à utiliser deux images catadioptriques successives (voir la Fig. 5.20).



(a) (b)

FIGURE 5.20 – Deux images successives de la séquence réelle utilisée dans l'expérimentation de la stratégie de ré-orientation.

Par la suite, nous appliquons une transformation polaire et une interpolation bicubique pour passer aux images panoramiques (Fig. 5.21).
Dans la troisième étape, nous avons besoin d'estimer le flot optique sur les images panoramiques (Fig. 5.22).
Dans la quatrième étape de l'algorithme de Kim [Kim 2007], nous calculons la matrice d'orientation du flot optique. Les vecteurs du flot optique sont classés en deux groupes : le premier groupe avec des orientations faibles ($< \pi/2$) et le deuxième avec des orientations élevées ($\geq \pi/2$) (Fig. 5.23).
Dans la cinquième étape, la matrice d'orientation doit être binarisée et filtrée afin de ne garder que les orientations faibles (région au centre). Les techniques adoptées de binari-

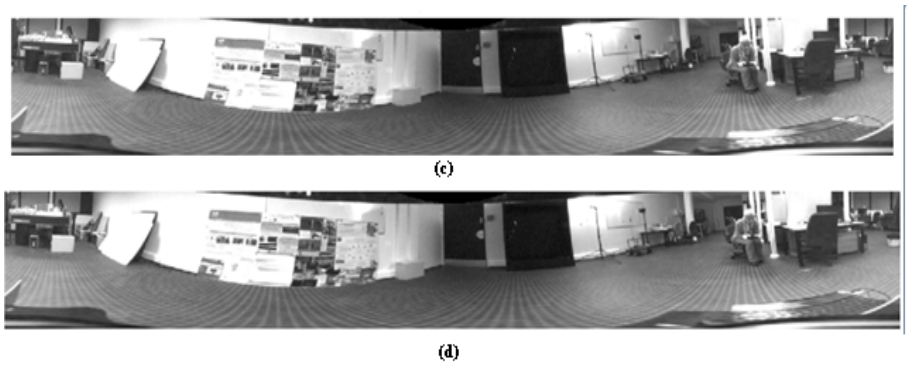


FIGURE 5.21 – Images panoramiques : (c) image panoramique de l'image omnidirectionnelle 5.20a. (d) image panoramique de l'image omnidirectionnelle 5.20b.

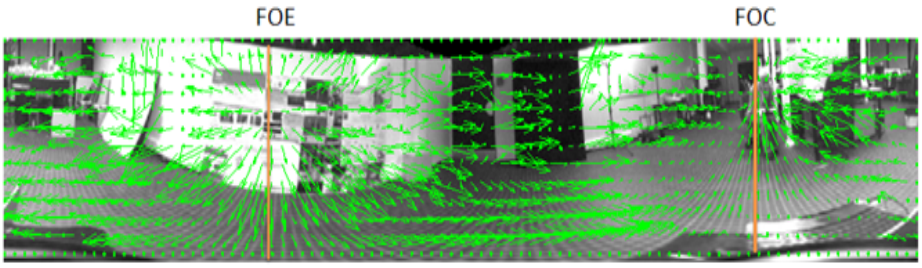


FIGURE 5.22 – Estimation du flot optique par la méthode de Lucas-Kanade à partir des deux images panoramiques de la figure 5.21.

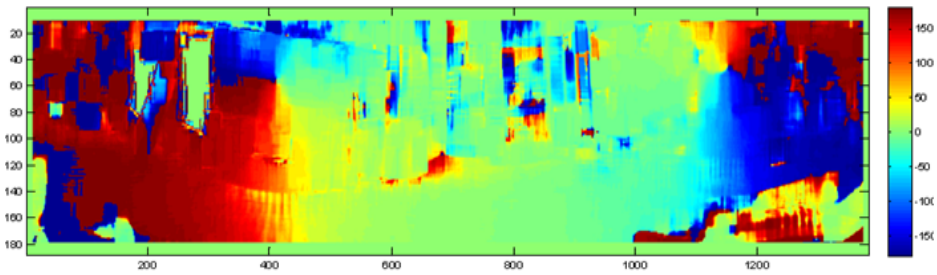


FIGURE 5.23 – Matrice d'orientation du flot optique calculée à partir de l'estimation du flot optique de la figure. 5.22.

sation et de filtrage sont présentées dans la figure 4.5, l'équation (4.40) et le tableau 4.1. La matrice d'orientation du flot optique (5.23) après l'étape de binarisation et de filtrage est montrée à 5.24.

La sixième étape consiste à extraire le FOE à partir de l'histogramme des orientations du

FIGURE 5.24 – Matrice d'orientation après avoir effectué biniarisation et filtrage.

flot optique (HOF) illustré dans la figure 5.25. Les valeurs maximale et le minimale des bins de l'histogramme des angles deviennent des candidats pour les angles des vecteurs (FOE) et (FOC). La taille de bin d'angles choisie est de 6°, le seuil est choisi à 300 et l'angle du vecteur (FOE) est de 88°.

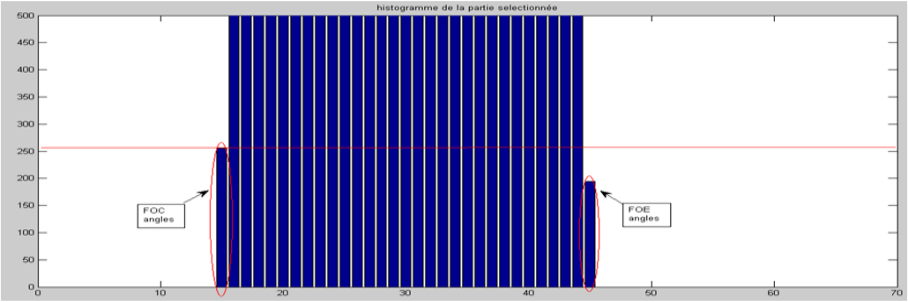


FIGURE 5.25 – Histograme des orientations du flot optique (HOF) sur la région segmentée 5.24.

Nous retrouvons les coordonnées du point FOE dans l'image panoramique comme il est montré dans la figure 5.26. Son correspondant dans l'image omnidirectionnelle est calculé à partir de la transformation polaire inverse (voir la Fig. 5.27).

Nous venons d'illustrer les étapes de l'algorithme de la stratégie de ré-orientation. L'estimation de l'angle $\widehat{\rho}$ dans ce cas est faite sur deux images de la séquence réelle puisque la direction de la translation est suposée stable durant toute la séquence, la position du FOE est constante. En cas de présence de rotations instantanées au cours du mouvement, nous devons suivre le FOE le long de la séquence et estimer au fur et à mesure les angles $\widehat{\rho}_k$ des ré-orientations instantanées.

Dans les deux sous-sections suivantes, nous appliquons la stratégie de la ré-orientation sur des séquences de synthèse et réelles afin de généraliser les deux méthodes d'estimation du

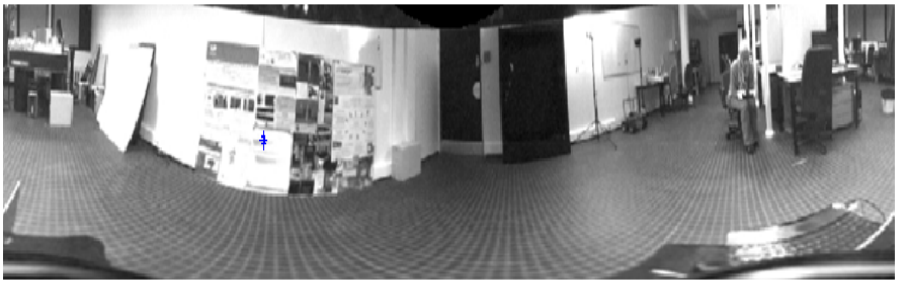


FIGURE 5.26 – Le point FOE estimé dans l'image panoramique. Le point FOE est matérialisé par l'étoile bleue.

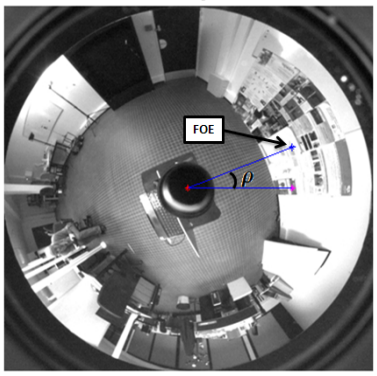


FIGURE 5.27 – Le point FOE dans l'image omnidirectionnelle et l'angle $\widehat{\rho}$ estimé pour la ré-orientation de l'image.

TAC sur les images catadioptriques.

5.2.3.2 Méthode d'estimation du TAC basée sur les gradients d'images (MGG) et la réorientation

L'objectif de cette sous-section est d'évaluer la méthode MGG avec la stratégie de ré-orientation. Il s'agit d'une généralisation de cette méthode au cas où le robot se déplace dans la scène suivant un axe dans le plan **(X,Y)**. Pour cela, nous proposons de prendre en considération la direction du FOE afin de corriger le désalignement des deux repères du capteur $(\mathbf{X}, \mathbf{Y})$ et du robot $(\mathbf{X'}, \mathbf{Y'})$ et ceci en appliquant la stratégie de ré-orientation (voir le diagramme 4.3). Nous présentons tout d'abord les résultats sur les images de synthèse et ensuite sur les images réelles.

A. Séquences de synthèse

Nous utilisons de nouvelles séquences de synthèse générées une nouvelle fois avec le logiciel

Povray et présentées dans la table ??. Dans ce cas, le mouvement du capteur virtuel para-catadioptrique s'effectue dans le plan **(X-Y)** tel que : $\dot{X} \neq 0$ et $\dot{Y} \neq 0$.

[Caractéristiques des séquences de synthèse utilisées dans le cas de la MGG pour un mouvement général du capteur para-catadioptrique.

Séquence	miroir	X_0	$(\dot{X}, \dot{Y}, \dot{Z})$	$(-1, a, b)$	Angle $\lvert\widehat{\rho}\rvert$	σ_n
seq 21	paraboloïde	50	(1,0.2,0)	$(-1, 0, 0)$	8°	0
seq 22	paraboloïde	50	(1,0.5,0)	$(-1, 0, 0)$	24°	0
seq 23	paraboloïde	50	(1,0.7,0)	$(-1, 0, 0)$	32°	0
seq 24	paraboloïde	50	(1,1,0)	$(-1, 0, 0)$	42°	0

Un exemple de séquence de synthèse généré avec une vitesse $(\dot{X}, \dot{Y}) = (1, 0.2)$ est illustré à la figure 5.28. L'extraction de la position du point FOE pour cette séquence est illustrée dans la figure 5.30 et par suite l'angle de ré-orientation estimé $\lvert\widehat{\rho}\rvert$ est de 8° degrés. Nous

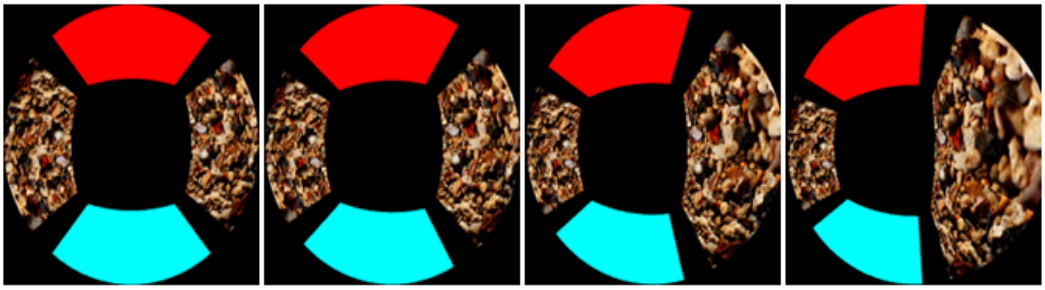


FIGURE 5.28 – Quatre images de la séquence de synthèse seq 21 issue d'un capteur para-catadioptrique virtuel en mouvement dans le plan **(X-Y)** . De gauche à droite : image $n°1$, image $n°11$, image $n°26$ et image $n°35$.

estimons le TAC avec la relation (4.28) et la technique de ré-orientation, appliquées aux séquences seq 21, seq 22, seq 23 et seq 24 de la table ??. Des améliorations significatives de l'estimation du TAC sont obtenues lorsque l'estimateur du TAC basé sur les gradients d'images est calculé avec la technique de ré-orientation (voir les Figs. 5.30 et 5.7).

B. Séquences vidéo réelles

Nous avons également traité des séquences d'images réelles pour tester l'extension, formée de la relation (4.28) et de la ré-orientation de la méthode proposée dans le cas du mou-

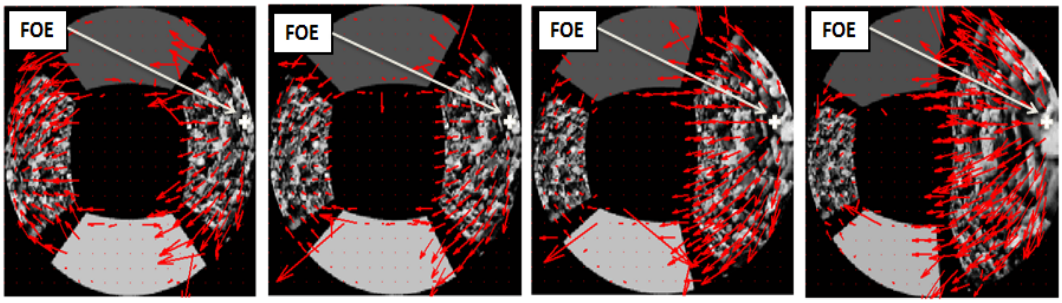


FIGURE 5.29 – Les vecteurs de flot optique estimés et la position du FOE sur les quatre images de la séquence de la figure 5.28.

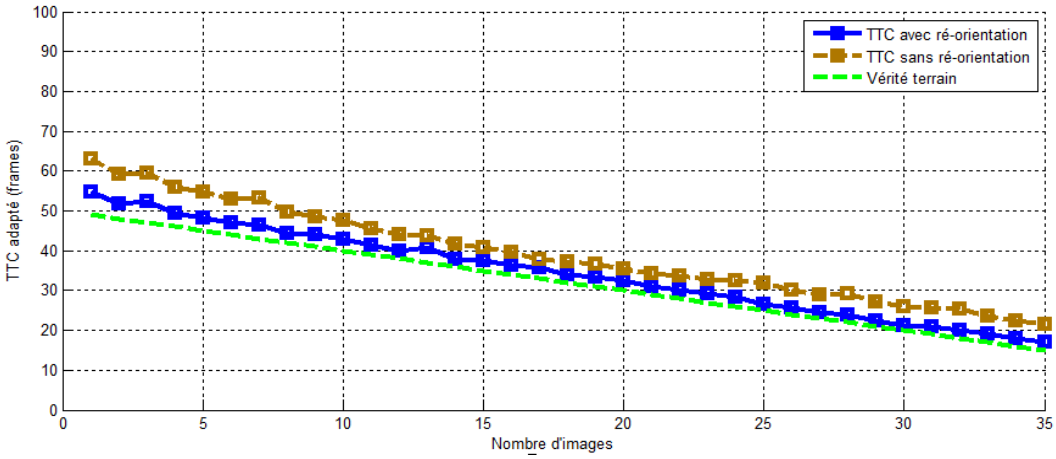


FIGURE 5.30 – Comparaison entre l'estimateur du TAC (4.28) couplé à la stratégie de ré-orientation (courbe bleue) et le même estimateur (4.28) sans la technique de ré-orientation (courbe marron). La vérité terrain est la courbe verte.

TABLE 5.7 – L'erreur ARE globale ($\widehat{ARE}$) de l'estimation du TAC, avec et sans ré-orientation pour la MGG.

ρ (en degré)	8°	24°	32°	42°
$\widehat{ARE}$ avec ré-orientation (%)	3	13	20	27
$\widehat{ARE}$ sans ré-orientation (%)	9	68	76	89

vement axial du capteur para-catadioptrique. La séquence réelle utilisée dans ce cas est décrite dans la table 5.8.

TABLE 5.8 – Caractéristiques de la séquence réelle utilisée dans le cas de la MGG avec mouvement général du capteur para-catadioptrique.

Séquence	miroir	Angle $\|\rho\|$	σ_n
seq R25	paraboloïde	46°	0

La figure 5.31 représente une comparaison entre la courbe du TAC avec et sans ré-orientation. Comme attendu, l'estimation du TAC avec la technique de ré-orientation est satisfaisante.

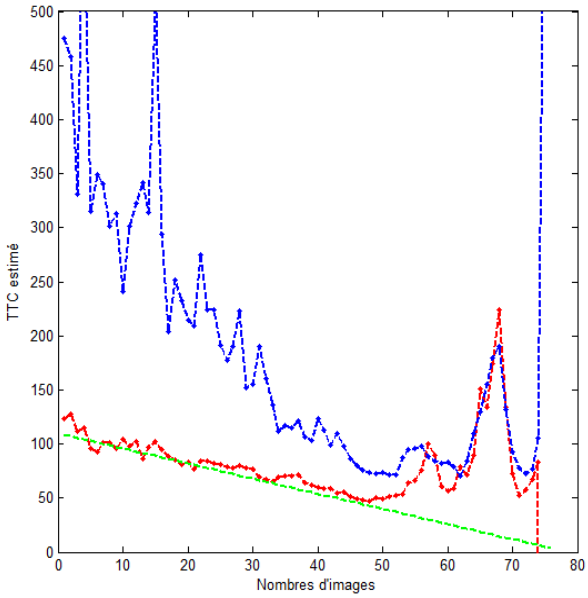


FIGURE 5.31 – Comparaison dans le cas réel entre l'estimateur du TAC de la MGG (4.28) et la stratégie de ré-orientation (courbe rouge), et le même estimateur sans la stratégie de ré-orientation (courbe bleue). L'angle de ré-orientation estimée est $|\widehat{\rho}| = 46°$. La vérité terrain est la courbe verte.

Dans ce qui suit, nous reportons les expérimentations réalisées sur des images de synthèse et des images réelles pour valider la méthode MLFO (4.38) couplé avec la ré-orientation.

5.2.3.3 Méthode basée sur le flot optique (MLFO) et ré-orientation

Dans la section 5.2.2.2 nous avons validé la relation (4.38) en utilisant le flot optique de la méthode Lucas-Kanade et le flot adapté. Dans cette section, nous allons validé la

relation (4.38) dans le cas où l'axe du capteur catadioptrique n'est pas aligné avec l'axe de déplacement du robot.

A. Séquences de synthèse

Les séquences de synthèse utilisées dans ce cas sont présentées dans la table 5.9. Les paramètres intrinsèques de ces séquences restent identiques (table 5.1).

TABLE 5.9 – Caractéristiques des séquences de synthèse utilisées dans le cas de la MLFO avec mouvement général du capteur hyper-catadioptrique.

Séquence	miroir	X_0	$(\dot{X}, \dot{Y}, \dot{Z})$	Angle $\lvert\rho\rvert$	σ_n
seq 25	hyperboloïde	50	(1,0.2,0)	8°	0
seq 26	hyperboloïde	50	(1,0.5,0)	24°	0
seq 27	hyperboloïde	50	(1,0.7,0)	32°	0
seq 28	hyperboloïde	50	(1,1,0)	42°	0

Le capteur hyper-catadioptrique virtuel se déplace avec une vitesse de composantes $\dot{X} = 1$ et $\dot{Y} = 0.2, 0.5, 0.7$ et 1. Dans ce cas, le calcul du TAC directement avec la relation (4.38) n'est pas valable car la containte $\dot{Y} = 0$ n'est pas vérifiée. Dans la discussion précédente, nous avons montré que le calcul du TAC est approprié aussi bien avec le flot optique adapté TALK qu'avec le flot classique TLK. Dans cette section, nous validons la technique de ré-orientation avec la seule méthode de flot optique de Lucas-Kanade. Nous comparons les deux estimations du TAC à partir de TLK, avec et sans ré-orientation. Les résultats sont montrés dans la deuxième et la troisième lignes de la figure 5.32. A noter que, sur cette figure, le bleu foncé indique une valeur faible du TAC (capteur proche de l'obstacle), un rouge clair désigne une valeur élevée du TAC (capteur loin de l'obstacle), alors que le rouge extrêmement foncé désigne les valeurs aberrantes. Comme prévu, le calcul du TAC avec TLK combiné avec la ré-orientation donne de meilleurs résultats et est plus précis que le calcul du TAC avec TLK appliqué directement sur la séquence. Ceci est confirmé par le calcul de l'erreur relative moyenne $\overline{ARE}$ présenté dans la figure 5.33 et par le calcul de l'erreur relative globale $\widehat{ARE}$ (5.3) reportée dans la table 5.10.

B. Séquences vidéo réelles

De la même manière que dans la section précédente, nous testons notre algorithme de

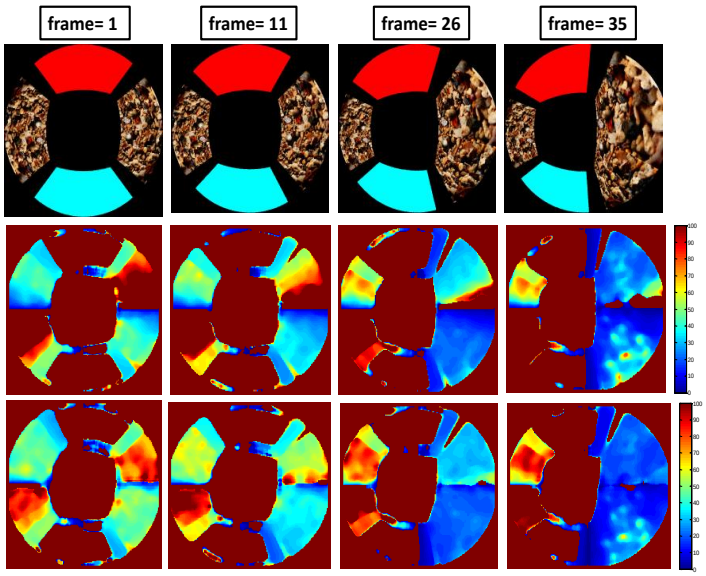


FIGURE 5.32 – Estimation du TAC sur la séquence de synthèse (seq25) : La première rangée montre l'évolution dans le temps de l'image de syntèse au cours du temps. La deuxième ligne donne les résultats de calcul du TAC avec TLK et la troisième ligne représente les résultats du calcul du TAC à partir de TLK avec la ré-orientation. L'angle ρ estimé est de 8°.

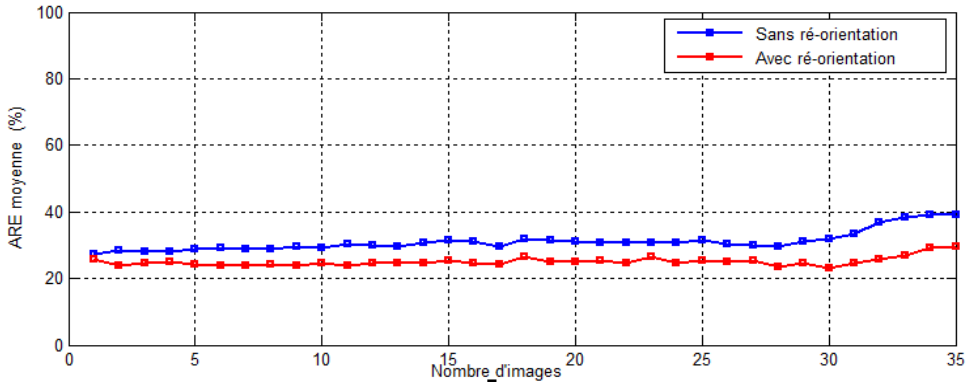


FIGURE 5.33 – Comparaison des deux estimateurs ($\overline{ARE}$) du TLK avec et sans la technique de la ré-orientation (seq25). La courbe bleue désigne le $\overline{ARE}$ du TLK sans la ré-orientation. La courbe rouge montre le $\overline{ARE}$ du TLK avec la ré-orientation.

calcul du TAC à partir de TLK couplé avec la ré-orientation sur la même séquence réelle para-catadioptrique (voir table 5.8) utilisée dans le cas de la première méthode. La figure 5.34 présente une comparaison entre la méthode du calcul du TAC à parir de TLK avec

TABLE 5.10 – L'erreur ARE globale ($\widehat{ARE}$) de l'estimation du TAC avec et sans ré-orientation.

$\lvert\rho\rvert$ (en degré)	8°	24°	32°	42°
$\widehat{ARE}$ avec ré-orientation (%)	25	29.5	33	41
$\widehat{ARE}$ sans ré-orientation (%)	31	41.5	44	45

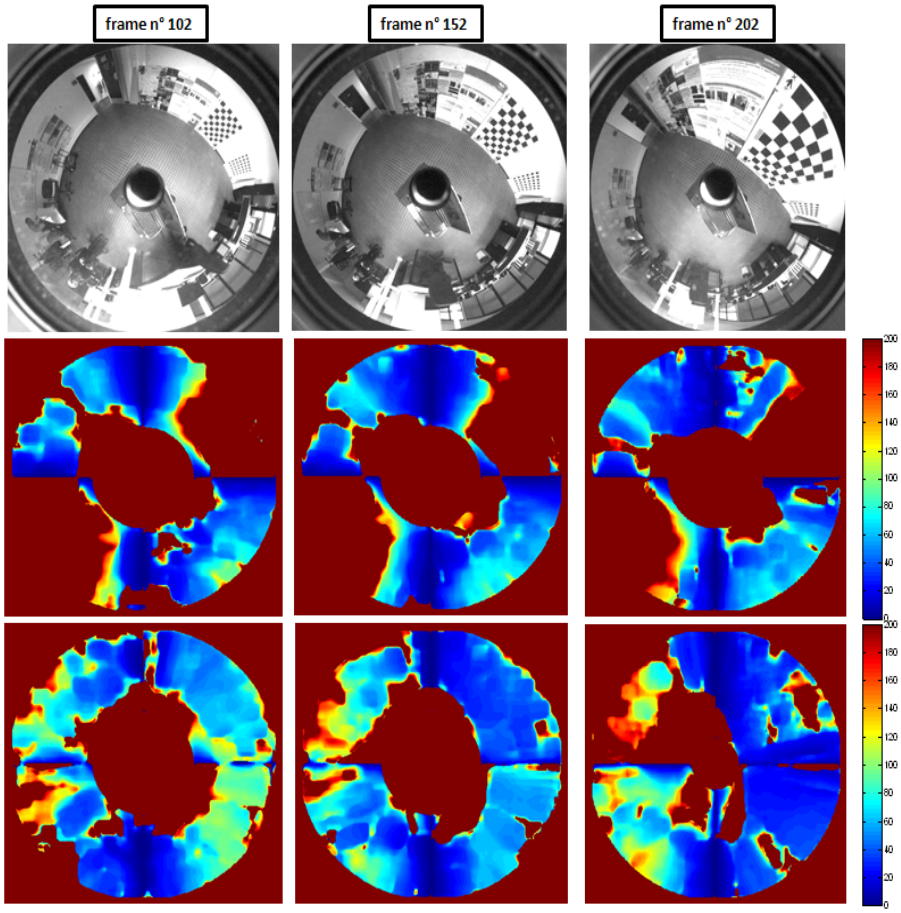


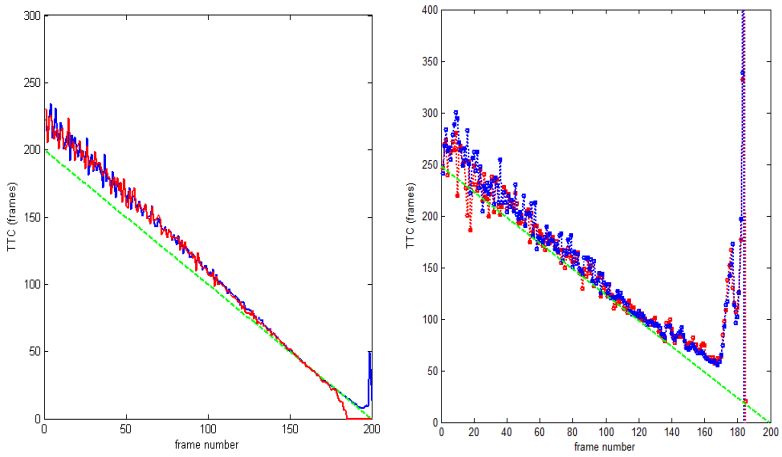
FIGURE 5.34 – Estimation du TAC sur la séquence réelle : La première rangée montre l'évolution dans le temps de l'image réelle. La deuxième rangée donne les résultats de calcul du TAC avec TLK et la troisième rangée représente les résultats du calcul du TAC à partir de TLK avec la stratégie de ré-orientation. L'angle ρ estimé est de 46°.

et sans ré-orientation appliquée sur une séquence d'images réelles résultant d'un mouvement quelconque dans le plan $(X - Y)$ du robot mobile sur lequel est montée la caméra

catadioptrique. Nous remarquons clairement que la ré-orientation améliore les résultats d'estimation du TAC dans le cas général du mouvement du capteur para-catadioptrique. Dans ce qui suit, nous présentons une comparaison des deux estimateurs du TAC proposés dans ce travail. Cette comparaison est réalisée sur une séquence de synthèse et une séquence réelle.

5.2.4 Comparaison des deux méthodes MGG et MLFO

Dans cette section, nous comparons les deux estimateurs adaptés du TAC : MGG et MLFO. Le premier estimateur est local, calculé en chaque pixel, c'est la raison pour laquelle nous utilisons la médiane sur tous les TAC dans la région d'intérêt pour chaque image afin de pouvoir le comparer avec le deuxième estimateur. La Figure 5.35 montre les résultats obtenus.



(a) (b)

FIGURE 5.35 – Comparaison entre les deux méthodes MGG et MLFO de calcul du TAC. La Courbe rouge montre l'estimateur du TAC calculé à partir de la MLFO et la courbe bleue l'estimateur du TAC calculé à partir de la MGG. (a) comparaison dans le cas de la séquence de synthèse. (b) Comparaison dans le cas de la séquence réelle.

Il est clair que les deux méthodes sont en mesure d'estimer correctement le TAC. L'avantage de la première méthode est qu'elle est valable pour tout type d'obstacles. En revanche, la deuxième méthode n'est valable que pour des surfaces planes. Ainsi, en terme d'application, la méthode basée sur le flot optique est plus générale : elle est applicable pour

n'importe quel obstacle, valable pour tout capteur à PVU. La méthode basée sur les gradients d'image est destinée seulement aux surfaces planes.
En terme de complexité, la méthode basée sur les gradients d'image est la plus simple ; elle est directe et ne nécessite pas une estimation du flot optique au préalable. La méthode basée sur le flot optique est une méthode qui dépend fortement de la qualité d'estimation du flot optique, pour pouvoir fournir une bonne précision dans l'estimation du TAC.
En terme de temps de calcul, nous avons montré pour la méthode basée sur le flot optique que le TLK donne de bons résultats sur les images de synthèse et les images réelles. Cette méthode estime le TAC en utilisant le flot optique de Lucas-Kanade. Elle est moins rapide que la méthode basée sur les gradients. En effet, pour une surface de synthèse de taille 5194 pixels, le temps de calcul (de 2 images de la séquence de la syhtnèse) du MGG est de $0.07s$ alors qu il faut $1.21s$ pour la MLFO, sachant que seul le calcul du flot optique de Lucas-kanade a nécessité $1.170095s$ avec le logiciel Matlab 7.1.2.0 a2011 sur une machine de 8Go de RAM et de processeur core(i5). Nous synthétisons les avantages et les inconvénients de chaque méthode dans le tableau suivant :

TABLE 5.11 – Les avantages et les inconvénients des deux méthodes d'estimation du TAC : méthode basée sur le flot optique (locale) et méthode basée sur les gradients (globale).

	Avantages	*Inconvénients*
Méthode basée sur le flot optique (MLFO)	- Valable pour tout type d'obstacle -Modèle général de TAC aux images catadioptrique	- N'est pas directe : nécessite tout d'abord l'estimation du flot optique -Problème du modèle sur les deux axes d'expansion .
Méthode basée gradients (MGG)	- Simple - Rapide	- Gradient d'image utilisé est non adapté - Modèle de TAC valable uniquement pour les images para-catadioptriques.

5.3 Conclusion

Nous avons présenté dans ce chapitre un ensemble d'expérimentations pour évaluer les contributions apportées dans cette thèse. Nous avons appliqué les deux méthodes de calcul du TTC dans le chapitre 4 à différentes séquences de synthèse et réelles, différents obstacles, différents types de mouvement.

Ainsi, nous avons comparé les deux méthodes sur les mêmes séquences de synthèse et réelles afin d'en déduire les avantages et les inconvénients. Dans tous les résultats obtenus, nous remarquons un bon comportement des deux estimateurs du TAC dans le cas du mouvement axial. Ainsi, nous avons montré que le suivi du FOE et la compensation de l'angle par la ré-orientation améliorent énormément l'estimation du TAC.

Chapitre

6 Conclusion générale et perspectives

La rage de vouloir conclure
est une des manies les plus funestes
et les plus stériles qui
appartiennent à l'humanité....Je vois, au contraire,
que les plus grands génies et
les plus oeuvres n'ont jamais conclu

Gustave Flaubert

Sommaire

6.1 Bilan des contributions. **136**

6.2 Perspectives **137**

Ce mémoire de thèse a abordé les questions d'estimation du temps à collision (TAC) dans le cadre de la vision omnidirectionnelle. De nombreuses ambiguités en termes d'analyse du mouvement, détection d'obstacle,.. peuvent être levées grâce au large champ de vue fourni par les capteurs omnidirectionnels. Néanmoins, les images omnidirectionnelles contiennent des distorsions et offrent une résolution non uniforme ainsi un traitement classique comme celui destiné aux images perspectives n'est pas approprié. C'est pourquoi l'adaptation des outils de traitement pour les images omnidirectionnelles est primordiale. C'est dans ce contexte que s'inscrit notre travail où nous avons proposé des solutions pour l'estimation

du TAC en vision catadioptrique.

6.1 Bilan des contributions.

Les travaux présentés dans ce manuscrit de thèse ont permis d'adapter deux méthodes d'estimation du TAC dédiées aux caméras perspectives. Nous avons démontré tout d'abord que les méthodes classiques d'estimation du TAC ne peuvent être appliquées directement sur les images omnidirectionnelles. En effet, le TAC est une quantité dérivée du modèle géométrique de projection (des points 3D sur le plan image). La formation des images classiques est réalisée selon la loi perspective, alors que les images omnidirectionnelles se forment suivant une loi de projection différente. C'est pour cette raison que les expressions classiques du TAC ne sont pas utilisables dans le cas omnidirectionnel. Le premier travail a consisté en une nouvelle formulation de l'estimation du TAC utilisant seulement les gradients spatiaux d'intensité et la dérivée temporelle. Nous avons également exprimé le mouvement apparent en fonction du TAC, ce qui peut être naturellement employé pour un traitement inverse, c'est à dire, estimer le flot optique en utilisant le TAC. Nous avons aussi exprimé l'équation de la surface plane en fonction des coordonnées image en utilisant le modèle de projection parabolique. Ceci nous a amené à déduire une nouvelle expression d'estimation du TAC en fonction de l'inverse du TAC et le vecteur normal de la surface plane.

Cette première méthode est une méthode simple, rapide et permet de fournir une valeur globale du TAC entre le capteur et l'obstacle. Néanmoins, cette méthode nécessite une segmentation préalable de l'obstacle, elle est valable seulement pour les capteurs para-catadioptriques et elle est formalisée seulement pour les surfaces planes.

Cela nous a naturellement amené à nous poser deux questions abordées dans le chapitre 4, sections 4.3 et 4.4 : Peut-on définir de manière générale, une expression du TAC valable pour tout capteur central à PVU ? Peut-on estimer le TAC pour n'importe quelle forme géométrique d'obstacle ?

A ces deux questions, la réponse est oui. Il est possible de formaliser une expression du TAC générale valable pour tous les capteurs centraux à PVU et pour tout type d'obstacle. Cette méthode est la méthode locale basée sur le flot optique, elle calcule le TAC pour chaque pixel de l'image et nous permet de former des cartes à (360°) de TAC. C'est

une expression simple qui dépend seulement du flot optique, des coordonnées du point principal et des coordonnées images. Par conséquent, elle ne nécessite pas de calibrage complet comme la première méthode.
Les deux équations (4.28) et (4.38) ont été developpées théoriquement avec la contrainte de l'alignement de l'axe $\mathbf{u}$ de la caméra avec l'axe du mouvement $\mathbf{X}$ du robot. Ceci nous a amené à nous poser une troisième question : Peut-on estimer le TAC dans le cas d'un mouvement quelconque du robot ? Malheuresement, il n'a pas été possible d'introduire un modèle général d'estimation du TAC appliqué directement sur l'image car les calculs deviennent compliqués et difficiles à résoudre. Cependant, nous avons proposé une solution à cette problématique : c'est la stratégie de ré-orientation. Cette technique nécessite d'effectuer une rotation sur l'image catadioptrique afin de revenir aux cas d'un mouvement axial et ainsi pouvoir appliquer les deux relations (4.28) et (4.38) pour estimer le TAC.

6.2 Perspectives

La levée de certains verrous scientifiques qui n'ont pas été traités dans ce travail pourrait permettre d'améliorer les résultats et devra être étudié par la suite. Pour conclure ce manuscrit, nous allons lister ces quelques points :

— Le calcul des gradients spatiaux d'intensité a été réalisé par des filtres discrets classiques développés pour les images perspectives. Or il est clair que, pour des raisons semblables à celles-ci qui nous ont poussé à adapter les méthodes de TAC classiques, ces filtres ne sont pas adaptés aux images omnidirectionnelles. Ainsi, les filtres discrets utilisés dans nos algorithmes aboutissent en théorie à des erreurs numériques pour toutes les équations qui exploient ces gradients. C'est pourquoi, il serait important de calculer les gradients spatiaux d'intensité en tenant compte des particularités des images catadioptriques.

— La méthode globale bien que simple et rapide, reste encore fortement limitée en terme d'application puisque celle-ci n'est utilisable que pour des capteurs para-catadioptriques. Il serait intéressant d'approfondir nos recherches pour que cette méthode puisse fonctionner pour tous les capteurs catadioptriques à point de vue unique. Nous pourrions par exemple travailler sur la sphère d'équivalence qui modélise l'ensemble de ces capteurs comme nous l'avons fait pour la méthode

locale.

— Nous avons remarqué que nos approches sont sensibles aux phénomènes d'aliasing. En effet, lorsque notre robot s'approche de l'obstacle, le mouvement apparent dans l'image s'accroît et l'approximation des dérivées temporelles de l'intensité par des différences finies entre les images devient imprécise. Nous devrions élaborer une approche multi-échelle pour pallier ce phénomène. Néanmoins une approche pyramidale, comme le propose différents travaux dans le cas perspectif, n'est pas directement applicable et il nous faudra donc développer de nouvelles techniques.

— Les deux méthodes développées dans ce travail ne sont applicables que pour des robots mobiles en déplacement sur une surface plane. En effet, dans les deux approches, nous supposons que $\dot{Z} = 0$. Ainsi, nous ne pouvons calculer le TAC pour un robot volant puisque dans ce cas, cette hypothèse n'est pas respectée. C'est pourquoi il serait intéressant de lever cette hypothèse pour que nos méthodes soient utilisables pour des robots à 6 degrés de libertés.

— Tous les résultats présentés dans ce manuscrit ont été implémentés hors ligne. En effet, pour valider nos méthodes, les algorithmes ont été testés sur des séquences d'images et non directement pendant le déplacement du robot. Ainsi, il sera important pour nos applications en robotique mobile d'implémenter nos algorithmes en temps réel sur nos robots.

Ces points seront étudiés par la suite pour améliorer nos approches et ainsi calculer le temps à collision (TAC) d'un robot en déplacement quelconque (6 ddl) de manière encore plus précise à l'aide de tous types de capteurs à point de vue unique.

Bibliographie

[Adelson 1985] H. Adelson et R. Bergen. *Spatiotemporal Energy Models for the Perception of Motion.* Journal of Optical Society of America, A., vol. 2, no. 2, pages 284–299, 1985.

[Alibouch 2012] B. Alibouch, A. Radgui, M. Rziza et D. Aboutajdine. *Optical Flow Estimation on Omnidirectional Images : An Adapted Phase Based Method.* In Image and Signal Processing, volume 7340 of *Lecture Notes in Computer Science*, pages 468–475. Springer Berlin Heidelberg, 2012.

[Aloimonos 1990] J. Aloimonos. *Perspective approximations.* Image and Vision Computing, vol. 8, no. 3, pages 179–192, 1990.

[Anandan 1989] P. Anandan. *A computational framework and an algorithm for the measurement of visual motion.* International Journal of Computer Vision, vol. 2, no. 3, pages 283–310, 1989.

[Ancona 1993] N. Ancona et T. Poggio. *Optical flow from 1D correlation : Application to a simple time-to-crash detector.* In Computer Vision, 1993. Proceedings., Fourth International Conference on, pages 209–214, 1993.

[Aris 1965] R. Aris. Vectors, tensors, and the basic equations of fluid mechanics. Prentice-Hall International Series in the Physical and Chemical Engineering Sciences. Prentice Hall, 1965.

[Bak 2011] A. Bak, S. Bouchafa et D. Aubert. *Focus of Expansion Localization through Inverse C-Velocity.* In Giuseppe Maino et GianLuca Foresti, editeurs, Image Ana-

lysis and Processing - ICIAP 2011, volume 6978 of *Lecture Notes in Computer Science*, pages 484–493. Springer Berlin Heidelberg, 2011.

[Baker 1998] S. Baker et S. K. Nayar. *A Theory of Catadioptric Image Formation.* In Proceedings of the 6th International Conference on Computer Vision, pages 35 – 42, January 1998.

[Barron 1994] J. L. Barron, D. J. Fleet et S. S. Beauchemin. *Performance of optical flow techniques.* International Journal of Computer Vision, vol. 12, pages 43–77, 1994.

[Barth 1996] M. Barth et C. Barrows. *A fast panoramic imaging system and intelligent imaging technique for mobile robots.* In Intelligent Robots and Systems '96, IROS 96, Proceedings of the 1996 IEEE/RSJ International Conference on, volume 2, pages 626–633, 1996.

[Benamar 2012] F.Z. Benamar, C. Demonceaux, S. EL Fkihi, E. Mouaddib et D. Aboutajdine. *Time to Contact Estimation on Paracatadioptric Cameras.* In Pattern Recognition (ICPR), 2012 21th International Conference on, pages 3602 – 3605, November 2012.

[Benamar 2013] F.Z. Benamar, S. EL Fkihi, C. Demonceaux, E. Mouaddib et D. Abotajdine. *Gradient based time to contact on paracatadioptric camera.* In IEEE Int. Conf. on Image Processing, ICIP'13, Melbourne, Australia, September 2013.

[Bernard 1998] C.P. Bernard. *Discrete wavelet analysis : A new framework for fast optic flow computation.* In Hans Burkhardt et Bernd Neumann, editeurs, Computer Vision - ECCV'98, volume 1407 of *Lecture Notes in Computer Science*, pages 354–368. Springer Berlin Heidelberg, 1998.

[Bruno 2002] E. Bruno et D. Pellerin. *Robust motion estimation using spatial Gabor-like filters.* Signal Processing, vol. 82, no. 2, pages 297 – 309, 2002.

[Camus 1995] T. Camus. *Calculating Time-to-Contact Using Real-Time Quantized Optical Flow.* In National Institute of Standards and Technology NISTIR 5609, 1995.

[Caron 2011] G. Caron et D. Eynard. *Multiple Camera Types Simultaneous Stereo Calibration.* In IEEE Int. Conf. on Robotics and Automation, ICRA'11, Shanghai, China, May 2011.

[Chaudhry 2009] R. Chaudhry, A. Ravichandran, G. Hager et R. Vidal. *Histograms of oriented optical flow and Binet-Cauchy kernels on nonlinear dynamical systems for*

the recognition of human actions. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 1932–1939, 2009.

[Cipolla 1997] R. Cipolla et A. Blake. *Image Divergence and Deformation from Closed Curves.* International Journal of Robotics Research, vol. 16, pages 77–96, 1997.

[Colombo 2000] C. Colombo. *Time to collision from first-order spherical image motion.* Robotics and Autonomous Systems, vol. 31, no. 1-2, pages 5 – 15, 2000.

[Coombs 1998] D. Coombs, M. Herman, T. Hong et M. Nashman. *Real-time obstacle avoidance using central flow divergence, and peripheral flow.* Robotics and Automation, IEEE Transactions on, vol. 14, no. 1, pages 49–59, 1998.

[Cutler 2002] R. Cutler, Y. Rui, A. Gupta, J. J. Cadiz, I. Tashev, L. wei He, A. Colburn, Z. Zhang, Z. Liu et S. Silverberg. *Distributed meetings : a meeting capture and broadcasting system.* In Lawrence A. Rowe, Bernard Mérialdo, Max Muhlhauser, Keith W. Ross et Nevenka Dimitrova, editeurs, ACM Multimedia, pages 503–512. ACM, 2002.

[Daniilidis 2002] K. Daniilidis, A. Makadia et T. Bulow. *Image processing in catadioptric planes : spatiotemporal derivatives and optical flow computation.* In Omnidirectional Vision, 2002. Proceedings. Third Workshop on, pages 3–10, 2002.

[David 1990] J.F. David et A.D. Jepson. *Computation of component image velocity from local phase information.* International Journal of Computer Vision, vol. 5, no. 1, pages 77–104, 1990.

[Demonceaux 2006] C. Demonceaux et P. Vasseur. *Markov random fields for catadioptric image processing.* Pattern Recognition Letters, vol. 27, no. 16, pages 1957–1967, 2006.

[Di Marco 2003] M. Di Marco, A. Garulli, D. Prattichizzo et A. Vicino. *A set theoretic approach for time-to-contact estimation in dynamic vision.* Automatica, vol. 39, no. 6, pages 1037 – 1044, 2003.

[Duric 1999] Z. Duric, A. Rosenfeld et J. Duncan. *The Applicability of Green's Theorem to Computation of Rate of Approach.* International Journal of Computer Vision, vol. 31, pages 83–98, 1999.

[Fabrizio 1994] M. Fabrizio et F. Nicolas. *Visual Control of Altitude and Speed in a Flying Agent.* In Proceedings of the Third International Conference on Simulation

of Adaptive Behavior : From Animals to Animats 3 : From Animals to Animats 3, SAB94, pages 91–99, Cambridge, MA, USA, 1994. MIT Press.

[Faugeras 1993] O. Faugeras. Three dimensional computer vision : A geometric viewpoint. the MIT Press, 1993.

[Galbraith 2005] J. Galbraith, G. Kenyon et R. Ziolkowski. *Time-to-Collision Estimation from Motion Based on Primate Visual Processing.* IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 8, pages 1279–1291, 2005.

[Gautama 2002] T. Gautama et M. Van Hulle. *A phase-based approach to the estimation of the optical flow field using spatial filtering.* Neural Networks, IEEE Transactions on, vol. 13, no. 5, pages 1127–1136, 2002.

[Geyer 2001] C. Geyer et K. Daniilidis. *Catadioptric Projective Geometry.* IJCV, vol. 45, no. 3, pages 223–243, December 2001.

[Giachetti 2000] A. Giachetti. *Matching techniques to compute image motion.* Image and Vision Computing, vol. 18, no. 3, pages 247 – 260, 2000.

[Gluckman 1998] J. Gluckman et S.K. Nayar. *Ego-motion and omnidirectional cameras.* In Computer Vision, 1998. Sixth International Conference on, pages 999–1005, 1998.

[Green 2004] W.E. Green, P.Y. Oh et G. Barrows. *Flying insect inspired vision for autonomous aerial robot maneuvers in near-earth environments.* In Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on, volume 3, pages 2347–2352 Vol.3, 2004.

[Hecht 1974] E. Hecht et A. Zajac. Optics. Reading, MA : Addison-Wesley, 1974.

[Heeger 1988] D.J. Heeger. *Optical flow using spatiotemporal filters.* International Journal of Computer Vision, vol. 1, no. 4, pages 279–302, jan 1988.

[Horn 1981] B.K.P. Horn et G.S. Brian. *Determining Optical Flow.* Artifical Intelligence, vol. 17, pages 185–203, 1981.

[Horn 2007] B.K.P. Horn, F. Yajun et I. Masaki. *Time to Contact Relative to a Planar Surface.* In Intelligent Vehicles Symposium, 2007 IEEE, pages 68–74, 2007.

[Horn 2009] B.K.P. Horn, Y. Fang et I. Masaki. *Hierarchical framework for direct gradient-based time-to-contact estimation.* In In IEEE Intelligent Vehicles Symposium, pages 1394–1400, Shaanxi, China, June 2009.

[Ibarra 2013] E.J. Ibarra, R. Medina et M. Garreau. *2-D segmentation of left ventricle in magnetic resonance images based on an optical flow algorithm.* In Health Care Exchanges (PAHCE), 2013 Pan American, pages 1–6, 2013.

[Ishiguro 1992] H. Ishiguro, M. Yamamoto et S. Tsuji. *Omni-directional stereo.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 14, no. 2, pages 257–262, 1992.

[Jain 1983] R. Jain. *Direct Computation of the Focus of Expansion.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. PAMI-5, no. 1, pages 58–64, 1983.

[Jiang 1990] Z. Jiang et S. Tsuji. *From anorthoscope perception to dynamic vision.* In Robotics and Automation, 1990. Proceedings., 1990 IEEE International Conference on, pages 1154–1160, 1990.

[Joarder 1994] K. Joarder et D. Raviv. *A new method to calculate looming for autonomous obstacle avoidance.* In Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on, pages 777–780, 1994.

[Kaplan 1991] W. Kaplan. Advanced calculus. 4th ed. Reading, MA : Addison-Wesley, 1991.

[Kim 2007] J. Kim et Y. Suga. *An omnidirectional vision-based moving obstacle detection in mobile robot.* International Journal of Control Automation and Systems, vol. 5, no. 6, pages 663–673, 2007.

[Koenderink 1975] J. J. Koenderink et A. J. Van Doorn. *Invariant Properties of the Motion Parallax Field due to the Movement of Rigid Bodies Relative to an Observer.* Optica Acta, vol. 22, no. 9, pages 773–791, 1975.

[Kondo 2005] K. Kondo, Y. Yagi et M. Yachida. *Non-isotropic omnidirectional imaging system for an autonomous mobile robot.* In Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on, pages 1228–1233, 2005.

[Lee 1976] D. N. Lee. *A theory of visual control of braking based on information about time-to-collision.* Perception, vol. 5, pages 437–459, 1976.

[Li 2006] J. Li, S.G. Nikolov, N.E. Scott-Samuel et C.P. Benton. *Reliable Real-Time Optical Flow Estimation for Surveillance Applications.* In Crime and Security, 2006. The Institution of Engineering and Technology Conference on, pages 402–407, 2006.

[Longuet-Higgins 1980] H. C. Longuet-Higgins et K. Prazdny. *The Interpretation of a Moving Retinal Image.* Proceedings of the Royal Society of London. Series B. Biological Sciences, vol. 208, no. 1173, pages 385–397, 1980.

[Lucas 1981] B. Lucas et T. Kanade. *An iterative image registration technique with an application to stereo vision.* Proc. DARPA IU Workshop, pages 121–130, 1981.

[Matsumoto 1996] Y. Matsumoto, M. Inaba et H. Inoue. *Visual navigation using view-sequenced route representation.* In Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on, volume 1, pages 83–88, 1996.

[McCarthy 2008] C. McCarthy, N. Barnes et R. Mahony. *A Robust Docking Strategy for a Mobile Robot Using Flow Field Divergence.* Trans. Rob., vol. 24, no. 4, pages 832–842, Août 2008.

[Meldrum 1993] R. Meldrum et G. J. Amanda. *Collision avoidance of flying locusts : steering torques and behaviour.* Journal of Experimental Biology, vol. 183, pages 35–60, 1993.

[Menéndez 1999] J.M. Menéndez, N. Garcia, L. Salgado et E. Rendon. *Model-based analytical {FOE} determination.* Signal Processing : Image Communication, vol. 14, no. 10, pages 785 – 798, 1999.

[Meyer 1992] F. Meyer et P. Bouthemy. *Estimation of time-to-collision maps from first order motion models and normal flows.* In Pattern Recognition, 1992. Vol.I. Conference A : Computer Vision and Applications, Proceedings., 11th IAPR International Conference on, pages 78–82, Aug 1992.

[Meyer 1994] F. G. Meyer. *Time-to-collision from first-order models of the motion field - Robotics and Automation.* IEEE Transactions on, vol. 10, no. 6, 1994.

[Mitiche 1996] A. Mitiche et P. Bouthemy. *Computation and analysis of image motion : A synopsis of current problems and methods.* International Journal of Computer Vision, vol. 19, no. 1, pages 29–55, 1996.

[Mouaddib 2005a] E. Mouaddib. *01-Introduction à la vision panoramique catadioptrique.* 2005.

[Mouaddib 2005b] E. Mouaddib. *Introduction à la Vision Panoramique Catadioptrique.* Traitement du Signal, vol. 22, September 2005.

[Mouaddib 2005c] E. Mouaddib et C. Pégard. Vision omnidirectionnelle, volume 22. Numéro spécial de la revue Traitement du Signal TS, December 2005.

[Mundy 1992] J. Mundy et A. Zisserman. *Introduction-towards a new framework for vision.* In Geometric invariance in computer vision, pages 1–39. MIT Press, 1992.

[Nalwa 1996] V. Nalwa. *A true Omnidirectional Viewer.* In Rapport technique, Bell Laboratoires, Holmdel NJ 07733, USA, February 1996.

[Negahdaripour 1989] S. Negahdaripour et B.K.P. Horn. *A direct method for locating the focus of expansion.* Computer Vision, Graphics, and Image Processing, vol. 46, no. 3, pages 303 – 326, 1989.

[Nelson 1989] R.C. Nelson et J. Aloimonos. *Obstacle avoidance using flow field divergence.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 11, no. 10, pages 1102–1106, 1989.

[Radgui 2008] A. Radgui, C. Demonceaux, E. M. Mouaddib, M. Rziza et D. Aboutajdine. *An adapted Lucas-Kanade's method for optical flow estimation in catadioptric images.* In The 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras-OMNIVIS, 2008.

[Radgui 2009] A. Radgui, C. Demonceaux, E. M. Mouaddib, M. Rziza et D. Aboutajdine. *Omnidirectional Egomotion Estimation from Adapted Motion Field.* In Signal-Image Technology Internet-Based Systems (SITIS), 2009 Fifth International Conference on, pages 8–15, 2009.

[Radgui 2011] A. Radgui, C. Demonceaux, EL. Mouaddib, M. Rziza et D. Aboutajdine. *Optical flow estimation from multichannel spherical image decomposition.* Computer Vision and Image Understanding, vol. 115, no. 9, pages 1263 – 1272, 2011.

[Sandini 1980] G. Sandini et V. Tagliasco. *An anthropomorphic retina-like structure for scene analysis.* Computer Graphics and Image Processing, vol. 14, no. 4, pages 365 – 372, 1980.

[Santos-Victor 1994] J. Santos-Victor et G. Sandini. *Visual Behaviors for Docking*, 1994.

[Santos-Victor 1997] J. Santos-Victor et G. Sandini. *Visual Behaviors for Docking.* Comput. Vis. Image Underst., vol. 67, no. 3, pages 223–238, Septembre 1997.

[Sarachik 1989] K.B. Sarachik. *Characterising an indoor environment with a mobile robot and uncalibrated stereo.* In Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on, pages 984–989 vol.2, 1989.

[Sazbon 2004] D. Sazbon, H. Rotstein et E. Rivlin. *Finding the focus of expansion and estimating range using optical flow images and a matched filter.* Machine Vision and Applications, vol. 15, no. 4, pages 229–236, 2004.

[Shanshan 2013] Z. Shanshan, C. Bauckhage, D.A. Klein et A.B. Cremers. *Moving pedestrian detection based on motion segmentation.* In Robot Vision (WORV), 2013 IEEE Workshop on, pages 102–107, 2013.

[Sharma 1996] R. Sharma et Y. Aloimonos. *Early detection of independent motion from active control of normal image flow patterns.* Systems, Man, and Cybernetics, Part B : Cybernetics, IEEE Transactions on, vol. 26, no. 1, pages 42–52, 1996.

[Srinivasan 2000] M. V. Srinivasan, S. W. Zhang, J. S. Chahl, E. Barth et S. Venkatesh. *How honeybees make grazing landings on flat surfaces.* Biological Cybernetics, vol. 83, no. 3, pages 171–183, 2000.

[Stratmann 2004] I. Stratmann et E. Solda. *Omnidirectional Vision and Inertial Clues for Robot Navigation.* J. Field Robotics, vol. 21, no. 1, pages 33–39, 2004.

[Subbarao 1990] M. Subbarao. *Bounds on time-to-collision and rotational component from first-order derivatives of image flow.* Computer Vision, Graphics, and Image Processing, vol. 50, no. 3, pages 329 – 341, 1990.

[Svoboda 2001] T. Svoboda et T. Pajdla. *Matching in Catadioptric Images with Appropriate Windows, and Outliers Removal.* In Proceedings of the 9th International Conference on Computer Analysis of Images and Patterns, CAIP '01, pages 733–740, London, UK, 2001. Springer-Verlag.

[Tistarelli 1992] M. Tistarelli et G. Sandini. *On the Advantages of Polar and Log-polar Mapping for Direct Estimation of Time-to-impact from Optical Flow.* IEEE Trans. on PAMI, vol. 15, pages 401–410, 1992.

[Tosic 2005] I. Tosic, I. Bogdanova, P. Frossard et P. Vandergheynst. *Multiresolution Motion Estimation for Omnidirectional Images.* In In Proceeding of European Signal Processing Conference, 2005.

[Van Leeuwen 2000] M.B. Van Leeuwen et F.C.A. Groen. *Motion estimation with a mobile camera for traffic applications.* In Intelligent Vehicles Symposium, 2000. IV 2000. Proceedings of the IEEE, pages 58–63, 2000.

[Van Leeuwen 2002] M.B. Van Leeuwen et F.C.A Groen. *Motion interpretation for in-car vision systems.* In Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on, volume 1, pages 135–140, 2002.

[Vassallo 2002] R.F. Vassallo, J. Santos-Victor et H.J. Schneebeli. *A general approach for egomotion estimation with omnidirectional images.* In Omnidirectional Vision, 2002. Proceedings. Third Workshop on, pages 97–103, 2002.

[Weber 1995] J. Weber et J. Malik. *Robust computation of optical flow in a multi-scale differential framework.* International Journal of Computer Vision, vol. 14, no. 1, pages 67–81, 1995.

[Yamazawa 1993] K. Yamazawa, Y. Yagi et M. Yachida. *Omnidirectional imaging with hyperboloidal projection.* In Intelligent Robots and Systems '93, IROS '93. Proceedings of the 1993 IEEE/RSJ International Conference on, volume 2, pages 1029–1034 vol.2, Jul 1993.

[Yang 2012] Y. Yang, Q. Liu, R. Ji et Y. Gao. *Dynamic 3D Scene Depth Reconstruction via Optical Flow Field Rectification.* PLoS ONE, vol. 7, no. 11, page 47041, 11 2012.

[Yoshizaki 2008] W. Yoshizaki, Y. Mochizuki, N. Ohnishi et A. Imiya. *Free Space Detection from Catadioptric Omnidirectional Images for Visual Navigation using Optical Flow.* In The 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras - OMNIVIS, Marseille, France, 2008. Rahul Swaminathan and Vincenzo Caglioti and Antonis Argyros.

[Zufferey 2006] J.C. Zufferey et D. Floreano. *Fly-inspired visual steering of an ultralight indoor aircraft.* Robotics, IEEE Transactions on, vol. 22, no. 1, pages 137–146, 2006.