



HAL
open science

CONTRIBUTION A LA STEREOVISION OMNIDIRECTIONNELLE ET AU TRAITEMENT DES IMAGES CATADIOPTRIQUES: APPLICATION AUX SYSTEMES AUTONOMES

Omar El Kadmiri

► **To cite this version:**

Omar El Kadmiri. CONTRIBUTION A LA STEREOVISION OMNIDIRECTIONNELLE ET AU TRAITEMENT DES IMAGES CATADIOPTRIQUES: APPLICATION AUX SYSTEMES AUTONOMES. Traitement des images [eess.IV]. UNIVERSITE MOHAMMED V, FACULTE DES SCIENCES de Rabat, 2014. Français. NNT: 2725 . tel-01071502v1

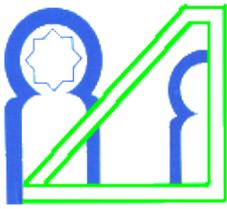
HAL Id: tel-01071502

<https://theses.hal.science/tel-01071502v1>

Submitted on 14 Oct 2014 (v1), last revised 19 May 2015 (v2)

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.



UNIVERSITE MOHAMMED V-AGDAL
FACULTE DES SCIENCES
Rabat



N° d'ordre : 2725

THÈSE DE DOCTORAT

Présentée par

Omar El KADMIRI

Spécialité : Électronique Et Traitement Des Images

Titre

**CONTRIBUTION A LA STEREOVISION OMNIDIRECTIONNELLE ET
AU TRAITEMENT DES IMAGES CATADIOPTRIQUES :
APPLICATION AUX SYSTEMES AUTONOMES**

Soutenue le 8 Septembre 2014 devant le jury composé de :

Présidente :

Mme. Rajaa CHERKAOUI EL MOURSLI PES Faculté des Sciences de Rabat,
Maroc

Examineurs :

Mr. Lhoussaine MASMOUDI PES Faculté des Sciences de Rabat, Maroc

Mr. Noureddine ZAHID PES Faculté des Sciences de Rabat, Maroc

Mr. Fatah JARACHI PES Ecole Mohammadia d'ingénieurs, Maroc

Mr. Javier González Jiménez PES Universidad de Málaga, Espagne

Faculté des Sciences, 4 Avenue Ibn Battouta B.P. 1014 RP, Rabat – Maroc

Tel +212 (0) 37 77 18 34/35/38, Fax : +212 (0) 37 77 42 61, <http://www.fsr.ac.ma>

*A mon père;
A ma mère;
A ma femme Kaoutar ;
A ma sœur Nadia;
A mon frère Zakaria;
A mon frère Samir;*



Ali EL KADMIRI & Fatima SERROUKH

Remerciements

Les travaux de recherche de cette thèse ont été réalisés dans le Laboratoire d'Electronique et de Traitement du Signal/Géomatique, sous la direction de Mr Lhoussaine MASMOUDI.

Sans doute la section la plus juteuse d'une thèse de Doctorat, est celle des remerciements. Dans un futur lointain, ce mémoire sera feuilleté par quelqu'un, dans une bibliothèque ou par curiosité, probablement son contenu pourra être expiré, mais la passion qui a réuni tant d'humains à accomplir ce travail ne le sera jamais.

Tout d'abord je tiens à remercier Pr. Lhoussaine MASMOUDI, mon directeur de thèse, qui m'a fait découvrir la vision par ordinateur, depuis mon projet de fin d'études du cycle licence, et m'a initié à l'art de l'écriture d'articles. Sans son soutien, à la fois scientifique, moral et matériel, ce travail n'aurait pas été possible.

Je suis très reconnaissant à Madame Rajaa CHERKAOUI EL MOURSLI qui a accepté de présider le jury.

Je remercie le Professeur Nouredine ZAHID d'avoir accepté de rapporter ma thèse, d'autant que les délais étaient très courts, et d'avoir assisté à ma soutenance de thèse autant que membre du jury.

Je tiens à remercier Pr. Fatah JARACHI, de l'Ecole Mohammadia d'Ingénieurs, autant que rapporteur, ainsi que membre du jury.

Mes remerciements vont également vers Mr. Javier González Jiménez qui s'est déplacé d'Espagne pour assister à la soutenance autant que rapporteur et membre du jury.

Je remercie chaleureusement tous les membres du Laboratoire LETS/Geomat, spécialement mon frère Zakaria EL KADMIRI doctorant dans la même structure de recherche pour son soutien. Je ne peux oublier les jeunes membres de l'équipe "LETS juniors", Saleh EL JOUMANI, Salah MOUGHYT, Said OUKACHA, et Zakaria EL JOUAHARI, qui ont partagé avec moi la passion de l'électronique et de la robotique.

Abstract

Computer vision and digital image processing are two disciplines aiming to endow computers with a sense of perception and image analysis, similar to that of humans. Artificial visual perception can be greatly enhanced when a large field of view is available.

This thesis deals with the use of omnidirectional cameras as a mean of expanding the field of view of computer vision systems. The visual perception of depth (3D) by means of omnistereoscopic configurations, and special processing algorithms adapted to catadioptric images, are the main subjects studied in this thesis.

Firstly a survey on 3D omnidirectional vision systems is conducted. It highlights the main approaches for obtaining depth information, and provides valuable indications for the choice of the configuration according to the application requirements. Then the design of an omnistereoscopic sensor is addressed, we present a new configuration of the proposed sensor formed by a unique catadioptric camera, dedicated to robotic applications. An experimental investigation of depth estimation accuracy was conducted to validate the new configuration.

Digital images acquired by catadioptric cameras present various special geometrical properties, such as non-uniform resolution and severe radial distortions. The application of conventional algorithms to process such images is limited in terms of performance. For that, new algorithms adapted to the spherical geometry of catadioptric images have been developed.

Gathered omnidirectional computer vision techniques were finally used in two real applications. The first concerns the integration of catadioptric cameras to a mobile robot. The second focuses on the design of a solar tracker, based on a catadioptric camera.

The results confirm that the adoption of such sensors for autonomous systems offer more performance and flexibility in regards to conventional sensors.

Keywords: omnidirectional stereo vision, catadioptric camera, image processing, autonomous systems.

Résumé

La vision par ordinateur est une discipline qui vise doter les ordinateurs d'un sens de perception et d'analyse d'image semblable à celui de l'homme. La perception visuelle artificielle peut être grandement améliorée quand un grand champ de vision est disponible.

Cette thèse traite de l'utilisation des caméras omnidirectionnelles comme un moyen d'élargir le champ de vision des systèmes de vision artificielle. La perception visuelle de la profondeur (3D) par le biais de configurations omnistéréo, et les algorithmes de traitement adaptés aux images catadioptriques, sont les principaux sujets étudiés.

Tout d'abord une étude des systèmes de vision omnidirectionnelle 3D est menée. Elle met en évidence les principales approches pour obtenir l'information sur la profondeur et fournit des indications précieuses sur le choix de la configuration en fonction des besoins de l'application. Ensuite, la conception d'un capteur omnistéréo est adressée ; nous présentons une nouvelle configuration du capteur proposé basé une caméra catadioptrique unique, et dédié à la robotique mobile. Des expérimentations sur la précision d'estimation de la profondeur ont été menées pour valider la nouvelle configuration.

Les images catadioptriques présentent diverses propriétés géométriques particulières, telles que la résolution non-uniforme et de fortes distorsions radiales. L'application des algorithmes de traitement classiques à ce type d'images se trouve limité en termes de performances. Dans ce sens, de nouveaux algorithmes adaptés à la géométrie sphérique de ces images ont été développés.

Les techniques de vision omnidirectionnelle artificielle recueillies ont été finalement exploitées dans deux applications réelles. La première concerne l'intégration des caméras catadioptriques à un robot mobile. La seconde porte sur la conception d'un suiveur solaire, à base d'une caméra catadioptrique.

Les résultats obtenus confirment que l'adoption de tels capteurs pour les systèmes autonomes offre plus de performances et de flexibilité en regards aux capteurs classiques.

Mots-clefs : stéréovision omnidirectionnelle, caméra catadioptrique, traitement d'images, systèmes autonomes

Table des matières

| | |
|---|------------|
| Remerciements | i |
| Abstract | ii |
| Résumé | iii |
| 1. Introduction, contexte et problématique | 1 |
| 1.1 Introduction générale | 1 |
| 1.1.1 La vision à large champ dans la nature | 1 |
| 1.1.2 Les capteurs de vision omnidirectionnelle | 4 |
| 1.1.2.1 Les lentilles spéciales | 5 |
| 1.1.2.2 Les systèmes d'acquisition d'images multiples..... | 6 |
| 1.1.2.3 Les systèmes catadioptriques..... | 7 |
| 1.2 Contexte de la thèse..... | 9 |
| 1.3 Problématique | 9 |
| 1.4 Contributions et organisation du mémoire..... | 11 |
| 2. Les systèmes d'acquisition 3D omnidirectionnels | 13 |
| 2.1 Introduction..... | 13 |
| 2.2 Les systèmes d'acquisition 3D actifs | 13 |
| 2.2.1 La télémétrie..... | 14 |
| 2.2.1.1 La télémétrie par mesure du temps de vol | 14 |
| 2.2.1.2 La télémétrie par mesure du déphasage..... | 18 |
| 2.2.1.3 La télémétrie par mesure de la fréquence | 18 |
| 2.2.2 La triangulation active..... | 19 |
| 2.2.3 Notes sur des systèmes d'acquisition actifs..... | 23 |
| 2.3 Les systèmes d'acquisition 3D passifs..... | 26 |
| 2.3.1 Les techniques <i>shape-from-X</i> | 26 |
| 2.3.2 Contribution à la stéréovision omnidirectionnelle..... | 34 |
| 2.3.2.1 Choix du miroir | 40 |
| 2.3.2.2 Choix de la configuration..... | 46 |
| 2.3.2.3 Notes sur la stéréovision omnidirectionnelle..... | 53 |
| 2.4 Conception d'un système de stéréovision omnidirectionnel..... | 54 |
| 2.4.1 Architecture matérielle du capteur proposé..... | 55 |
| 2.4.2 Étalonnage du capteur | 56 |

| | | |
|--|---|-----------|
| 2.4.3 | Estimation de la profondeur par triangulation passive..... | 57 |
| 2.4.4 | Résultats expérimentaux | 59 |
| 2.4.5 | Discussion et synthèse | 62 |
| 2.5 | Conclusion | 62 |
| 3. Contribution au traitement des images | | |
| omnidirectionnelles..... | | 65 |
| 3.1 | Introduction..... | 65 |
| 3.2 | Application des traitements classiques aux images omnidirectionnelles | 66 |
| 3.3 | Déballage des images omnidirectionnelles | 67 |
| 3.4 | Adaptation des traitements..... | 71 |
| 3.4.1 | Un modèle de projection unifié | 73 |
| 3.4.2 | La projection stéréographique | 74 |
| 3.4.3 | Les opérateurs classiques de détection des contours | 75 |
| 3.4.3.1 | Détecteurs de contours du premier ordre | 76 |
| 3.4.3.2 | Détecteurs de contours du second ordre | 78 |
| 3.4.4 | Détecteur de contours dédié aux images omnidirectionnelles..... | 79 |
| 3.4.4.1 | Modèle sphérique de charges électriques virtuelles..... | 79 |
| 3.4.4.2 | Algorithme de détection des contours dans les images sphériques | 82 |
| 3.4.4.3 | La méthode d'évaluation de Fram & Deutsch | 84 |
| 3.4.4.4 | Résultats expérimentaux..... | 85 |
| 3.4.5 | Adaptation du détecteur de points d'intérêt de Harris..... | 90 |
| 3.4.5.1 | La détection des points d'intérêt | 90 |
| 3.4.5.2 | Les détecteurs de coins basés modèle | 91 |
| 3.4.5.3 | Les détecteurs de coins basés sur la géométrie | 91 |
| 3.4.5.4 | Le détecteur de coins de Harris..... | 92 |
| 3.4.5.5 | Détecteur de coins dédié aux images omnidirectionnelles..... | 93 |
| 3.4.5.6 | Résultats illustratifs et comparaisons | 96 |
| 3.5 | Conclusion | 103 |
| 4. La vision omnidirectionnelle pour les systèmes autonomes | | |
| 106 | | |
| 4.1 | Introduction..... | 106 |
| 4.2 | Conception et réalisation d'un robot mobile | 107 |
| 4.2.1 | L'électronique | 109 |

| | | |
|-------|---|------------|
| 4.2.2 | La mécanique..... | 113 |
| 4.2.3 | Application : suivi d'une cible en mouvement..... | 116 |
| 4.3 | Conception et réalisation d'un suiveur solaire | 121 |
| 4.3.1 | Le suiveur solaire omnidirectionnel : concept et contrôle | 125 |
| 4.3.2 | Validation expérimentale..... | 128 |
| 4.3.3 | Discussion des résultats..... | 130 |
| 4.4 | Conclusion | 132 |
| | Conclusions | 134 |
| | Références | 135 |

Table des figures

| | |
|---|----|
| Figure 1.1 : Champ de vision d'une mouche. | 2 |
| Figure 1.2 : Les yeux du Gigantocypris d'ostracodes..... | 3 |
| Figure 1.3 : (a) ViewPlus ASTRO Sensor (b) L'araignée sautante. (c) Le Scallop. | 3 |
| Figure 1.4 : (a) Champ de vision directionnel (b) Champ de vision panoramique (c) champ de vision omnidirectionnel..... | 5 |
| Figure 1.5 : (a) Lentille fisheye Samyang 8mm F3.5 (b) Une image obtenue avec une lentille fisheye. (c) Spherecam..... | 6 |
| Figure 1.6 (a) Système de vision panoramique par mosaïquage. (b) Système de vision omnidirectionnel à multi-caméras. (c) système de vision panoramique à quatre caméras..... | 7 |
| Figure 1.7 : (a) Le siège de Syracuse et les miroirs d'Archimède. (b) L'autoportrait en miroir sphérique (1935). | 8 |
| Figure 1.8 : Formation d'une image catadioptrique par projection orthographique. | 10 |
| Figure 2.1 : (a) Le robot Segway RMP. (b) Configuration du système laser-camera omnidirectionnelle..... | 16 |
| Figure 2.2 (a) Robot mobile dédié à la cartographie. (b) Exemple d'une scène. (c) La carte générée représentant l'espace libre..... | 17 |
| Figure 2.3 : (a) Une voiture de la flotte Google Street View équipée d'un système de camera R5. (b) système de camera R5. (c) système de camera R7. | 17 |
| Figure 2.4 : Représentation de la triangulation active par émission d'un faisceau laser. | 19 |
| Figure 2.5 : triangulation active par émission d'un plan laser. | 20 |
| Figure 2.6 : Acquisition 3D active par émission d'un motif de segments laser. ... | 20 |
| Figure 2.7 : (a) Le dispositif d'acquisition 3D KINECT (b) Motif de lumière structurée projeté par KINECT..... | 21 |
| Figure 2.8 : (a) Photographie du capteur de vision omnidirectionnelle. (b) Configuration du capteur de vision omnidirectionnelle. | 21 |
| Figure 2.9 : Système d'acquisition 3D omnidirectionnel actif..... | 22 |
| Figure 2.10 : (a) Illustration d'un système de télémétrie omnidirectionnel. (b) Triangulation active..... | 23 |
| Figure 2.11 : Illusion d'optique due au manque d'information sur la profondeur. | 23 |
| Figure 2.12 : (a) Système de balayage à miroir pour télémètre laser. (b) RIEGL VQ-450 scanner laser à large champ de vision. | 25 |
| Figure 2.13 : Modèle de lentille mince. | 28 |
| Figure 2.14 : Reconstruction de la forme à partir de la texture. (a) Image réelle : tissu à texture anisotrope non stationnaire (b) Diagramme à aiguille des Texels détectés (c) La profondeur de la surface illustrée en niveaux de gris (d) Maillage de la surface vue de profil..... | 29 |
| Figure 2.15 : Exemple de reconstruction à partir d'une image avec des ombres noires. | 31 |

| | |
|--|----|
| Figure 2.16 : Reconstruction 3D d'une boule (ping-pong) par technique shape-from-shading. | 32 |
| Figure 2.17 : Principe de la stéréovision. | 32 |
| Figure 2.18 : La géométrie épipolaire. (a) : Une scène observée sous différents points de vue, la géométrie épipolaire doit décrire la relation entre les deux images résultantes. (b) Schématisation de la géométrie épipolaire. | 33 |
| Figure 2.19 : le système de stéréovision omnidirectionnelle SPHEREO..... | 35 |
| Figure 2.20 : Système de vidéosurveillance à base de capteurs omnidirectionnels. (a) capteur catadioptrique utilisé. (b) Principe de la stéréovision omnidirectionnelle N-Oculaire. (c) Aperçu du système. | 36 |
| Figure 2.21 : Les différentes configurations d'un system omnistéréo monté sur un véhicule..... | 37 |
| Figure 2.22 : (a) Capteur d'image stéréo omnidirectionnel. (b) Système de projection 3D immersive CYLINDRA. | 37 |
| Figure 2.23 : Carte de profondeur panoramique générée à partir d'images stéréo. | 38 |
| Figure 2.24 : Miroirs convexes pour caméras catadioptriques. (a) Miroir conique (b) Miroir sphérique (c) Miroir hyperbolique (d) Miroir parabolique. | 39 |
| Figure 2.25 : Géométrie de formation de l'image acquise par caméra catadioptrique à miroir conique. | 40 |
| Figure 2.26 : Géométrie de formation de l'image acquise par caméra catadioptrique à miroir sphérique..... | 41 |
| Figure 2.27 : Système catadioptrique avec miroir sphérique à double lobe. (a) Configuration du système. (b) Exemple d'image. | 42 |
| Figure 2.28 : Géométrie de formation de l'image acquise par caméra catadioptrique à miroir hyperbolique..... | 43 |
| Figure 2.29 : système stéréo omnidirectionnel à miroirs hyperboliques. (a) configuration du système. (b) exemple d'image. | 44 |
| Figure 2.30: Géométrie de formation de l'image acquise par caméra catadioptrique à miroir parabolique..... | 45 |
| Figure 2.31 : Configuration omnistéréo horizontale..... | 46 |
| Figure 2.32 : (a) Triangulation omnistéréo horizontale. (b) Les lignes épipolaires. | 47 |
| Figure 2.33 : système de vision omnistéréo binoculaire-H..... | 48 |
| Figure 2.34 : système omnistéréo pour véhicule intelligent. (a) système hybride. (b) système catadioptrique. | 49 |
| Figure 2.35 : Configuration omnistéréo binoculaire verticale..... | 49 |
| Figure 2.36 : La géométrie épipolaire d'un système omnistéréo binoculaire-V. (a) Les lignes épipolaires de l'image catadioptrique. (b) les lignes épipolaires de l'image panoramique..... | 50 |
| Figure 2.37 : (a) Triangulation et calcul de la profondeur (b) Système omnistéréo binoculaire-V à deux cameras catadioptriques. | 51 |
| Figure 2.38: DaX, le robot mobile Scitos G5. | 51 |

| | |
|--|----|
| Figure 2.39 : Système omnistéréo à caméra unique et deux miroirs hyperboliques. (a) Apparence du système de vision. (b) Une image acquise avec ce système. (c) Triangulation. | 52 |
| Figure 2.40 : Système omnistéréo à caméra unique et multiple miroirs (sphériques). | 53 |
| Figure 2.41 : (a) Architecture du capteur catadioptrique proposé. (b) Configuration du système omnistéréo. | 55 |
| Figure 2.42 : Illustration du processus d'ajustement du capteur catadioptrique. (a) Capteur non ajusté. (b) Capteur ajusté. | 56 |
| Figure 2.43 : Triangulation dans le cas du système omnistéréo proposé. | 57 |
| Figure 2.44 : Procédure d'estimation de la profondeur. | 59 |
| Figure 2.45 : Une paire d'images omnistéréo. (a) Image acquise à position haute. (b) Image acquise à position basse. | 60 |
| Figure 2.46 : Courbe de variation de l'erreur, $D=2.18$ | 60 |
| Figure 2.47 : Courbe de variation de l'erreur, $D=3.18$ | 61 |
| Figure 2.48 : Courbe de variation de l'erreur, $D=4.18$ | 61 |
| Figure 3.1 : Déballage d'une image catadioptrique (a) image catadioptrique. (b) l'image panoramique correspondante. | 67 |
| Figure 3.2 : Principe du déballage des images omnidirectionnelles. | 68 |
| Figure 3.3: changement du repère de l'image catadioptrique. | 69 |
| Figure 3.4: Image panoramique non interpolée. | 70 |
| Figure 3.5: (a) Image omnidirectionnelle utilisée dans l'expérimentation du déballage. (b) Image panoramique correspondante, résolution : 122 X 749, $\alpha=1$, $\beta=0.5$ | 71 |
| Figure 3.6 : Problème du voisinage pixellique. (a) Voisinage pixellique incorrecte dans le cas d'un filtre planaire. (b) Voisinage correcte dans le cas d'un filtre sphérique. | 72 |
| Figure 3.7 : Projection d'un point 3D par le modèle unifié. | 73 |
| Figure 3.8 : La projection stéréographique. | 74 |
| Figure 3.9 : Modèles de contours. (a) Marche d'escalier. (b) Rampe. (c) Toit. (d) Ligne. | 75 |
| Figure 3.10 : Distribution sphérique de charges électrostatiques. | 80 |
| Figure 3.11 : (a) Image catadioptrique. (b) Image mappée sur la sphère. (c) L'image panoramique correspondante. | 82 |
| Figure 3.12 : Algorithme de détection de contours dans l'image sphérique. | 83 |
| Figure 3.13 : (a) L'image synthétique générée pour l'évaluation quantitative de la détection de contours, avec un contour verticale à 0° et un contour diagonale à 60° . (b) La région de contour considérée. | 84 |
| Figure 3.14 : (a) Valeurs de $P1$ avec bruit gaussien de variance croissante (contour vertical, un saut de 50 niveaux de gris). (b) Valeurs de $P2$ | 85 |
| Figure 3.15 : (a) Valeurs de $P1$ avec bruit gaussien de variance croissante (contour vertical, un saut de 100 niveaux de gris). (b) Valeurs de $P2$ | 86 |
| Figure 3.16 : (a) Valeurs de $P1$ avec bruit gaussien de variance croissante (contour diagonal, un saut de 150 niveaux de gris). (b) Valeurs de $P2$ | 86 |

| | |
|---|-----|
| Figure 3.17 : (a) Valeurs de $P1$ avec bruit gaussien de variance croissante (contour diagonal, un saut de 50 niveaux de gris). (b) Valeurs de $P2$ | 86 |
| Figure 3.18 : (a) Valeurs de $P1$ avec bruit gaussien de variance croissante (contour diagonal, un saut de 100 niveaux de gris). (b) Valeurs de $P2$ | 87 |
| Figure 3.19 : (a) Valeurs de $P1$ avec bruit gaussien de variance croissante (contour vertical, un saut de 150 niveaux de gris). (b) Valeurs de $P2$ | 87 |
| Figure 3.20 : Image réelle utilisée pour l'évaluation de la détection de contours. | 88 |
| Figure 3.21 : Image réelle filtrée avec l'opérateur de Sobel..... | 88 |
| Figure 3.22 : Image réelle filtrée avec l'opérateur de Prewitt..... | 89 |
| Figure 3.23 : Image réelle filtrée avec l'opérateur PMCEV..... | 89 |
| Figure 3.24 : Image réelle filtrée avec l'opérateur SMCEV..... | 89 |
| Figure 3.25 : Exemple d'extraction des points d'intérêt. (a) Image catadioptrique acquise avec un miroir parabolique. (b) Image panoramique correspondante. (c) Extraction de points d'intérêt..... | 90 |
| Figure 3.26 : Algorithme de détection de coins dans l'image sphérique..... | 94 |
| Figure 3.27: L'algorithme proposé de détection des coins..... | 95 |
| Figure 3.28 : (a) La première image synthétique utilisée pour l'évaluation de la détection des coins. (b) La seconde image synthétique..... | 96 |
| Figure 3.29 : Résultat de détection de coins dans la première image synthétique. (a) L'algorithme proposé. (b) Harris standard..... | 97 |
| Figure 3.30 : Résultat de détection de coins dans la seconde image synthétique originale. (a) L'algorithme proposé. (b) Harris standard..... | 98 |
| Figure 3.31 : Résultat de détection de coins dans la seconde image synthétique bruitée. (a) L'algorithme proposé. (b) Harris standard..... | 99 |
| Figure 3.32 : Résultat de détection de coins dans la seconde image synthétique floue. (a) L'algorithme proposé. (b) Harris standard..... | 100 |
| Figure 3.33 : Résultats d'évaluation de la localisation des coins détectés. (a) L'approche proposée. (b) Harris standard..... | 101 |
| Figure 3.34 : Résultats de détection de coins sur une image réelle en utilisant l'algorithme proposé..... | 102 |
| Figure 3.35 : Résultats de détection de coins sur une image réelle en utilisant l'algorithme standard de Harris..... | 103 |
| Figure 4.1 : Le robot ESCALADE360 v1.0..... | 107 |
| Figure 4.2 : Différentes configurations pour les robots à chenilles..... | 108 |
| Figure 4.3 : ESCALADE360 v2.0..... | 108 |
| Figure 4.4 : ESCALADE360 v3.0..... | 109 |
| Figure 4.5 : (a) Capteur de vision omnistéréo à baseline variable. (b) Capteur omnistéréo à baseline fixe..... | 110 |
| Figure 4.6 : Capteur de vision omnidirectionnelle monoculaire..... | 111 |
| Figure 4.7 : Unité de traitement du robot ESCALADE360..... | 112 |
| Figure 4.8 : Circuit de commande du robot ESCALADE360..... | 112 |
| Figure 4.9 : Système de contrôle du robot ESCALADE360..... | 113 |
| Figure 4.10 : Moteur électrique DC 12 V..... | 114 |
| Figure 4.11 : Les moteurs primaires du robot..... | 114 |
| Figure 4.12 : Mécanisme des flippers du robot ESCALADE360..... | 115 |

| | |
|--|-----|
| Figure 4.13 : Les quatre positions principales des flippers. (a) Quatre points de contact, $\alpha \in 0^\circ, 180^\circ$. (b) contact allongé, $\alpha = 180^\circ$. (c) Contact condensé $\alpha = 0^\circ$. (d) Contact réduit. $\alpha = 180^\circ, 360^\circ$ | 115 |
| Figure 4.14 : Les dimensions du robot ESCALADE360..... | 116 |
| Figure 4.15 : La cible suivie. (a) Image catadioptrique HSV à faible éclairage. (b) Image catadioptrique HSV à fort éclairage..... | 118 |
| Figure 4.16 : Détection et orientation du robot vers la cible suivie. | 120 |
| Figure 4.17 : Expérience de suivi autonome d'une cible en mouvement | 121 |
| Figure 4.18 : Photo-capteur à balance de luminosité | 123 |
| Figure 4.19 : Carte de la trajectoire solaire de Rabat (Maroc)..... | 124 |
| Figure 4.20 : Illustration de la localisation du soleil avec une caméra catadioptrique. | 124 |
| Figure 4.21 : Suiveur solaire Helianthus360..... | 125 |
| Figure 4.22 : L'architecture du suiveur solaire proposé..... | 126 |
| Figure 4.23 : Unité de contrôle du suiveur solaire Helianthus360..... | 126 |
| Figure 4.24 (a) Image omnidirectionnelle échantillon d'un ciel partiel. (b) L'image omnidirectionnelle après seuillage. (c) Schématisation de la procédure de suivi du soleil..... | 127 |
| Figure 4.25 : Les mesures de puissance du 17 Juillet 2013..... | 129 |
| Figure 4.26 : Les mesures de puissance du 20 Septembre 2013..... | 129 |
| Figure 4.27 : (a) Première expérience 17 Juillet 2013. L'ensoleillement total journalier: 25.7 MJm^2 (b) Deuxième expérience 20 septembre 2013. L'ensoleillement total journalier: 21 MJm^2 | 130 |
| Figure 4.28 : Robot mobile équipé d'un suiveur solaire..... | 132 |

Liste des tableaux

| | |
|--|-----|
| Tableau 2-1 : Les techniques passives d'acquisition 3D..... | 26 |
| Tableau 2-2 : Tableau récapitulatif des miroirs principaux..... | 46 |
| Tableau 2-3 : Tableau récapitulatif des principales configurations omnistéréo. . | 54 |
| Tableau 2-4 : Variation de l'erreur sur la profondeur en fonction de h pour D=2.18m | 60 |
| Tableau 2-5 : Variation de l'erreur sur la profondeur en fonction de h pour D=3.18m | 61 |
| Tableau 2-6 : Variation de l'erreur sur la profondeur en fonction de h pour D=4.18m | 61 |
| Tableau 3-1 : Caractéristiques de la plateforme d'expérimentation..... | 70 |
| Tableau 3-2 : Temps d'exécution de l'algorithme du déballage..... | 70 |
| Tableau 3-3 : Taux d'évaluation de la première image synthétique..... | 98 |
| Tableau 3-4 : Taux d'évaluation de la seconde image synthétique..... | 98 |
| Tableau 3-5 : Taux d'évaluation de la seconde image synthétique (bruit gaussien) | 99 |
| Tableau 3-6 : Taux d'évaluation de la seconde image synthétique (floue). | 100 |
| Tableau 3-7 : Evaluation de la localisation des coins détectés par mesure de la distance euclidienne..... | 101 |
| Tableau 4-1 : Variation de l'erreur en fonction de D, h=0.425m..... | 121 |

1. Introduction, contexte et problématique

1.1 Introduction générale

Le traitement d'images et la vision artificielle sont des disciplines relativement récentes (~années 60) et qui ont connu une évolution rapide. Elles sont en plein développement et donnent naissance chaque année à une profusion de progrès académiques, technologiques, et industriels. L'engouement pour ces disciplines s'explique par l'émergence permanente d'applications et d'enjeux industriels dans des domaines assez variés. Aujourd'hui, grâce aux technologies toujours en engouement, les images ont envahi notre quotidien, que ce soit dans le divertissement (Photos numériques, jeux vidéo...) ou dans un cadre plus utilitaire. Citons à titre d'exemple l'utilisation croissante de la vision dans des domaines variés tels que la médecine (IRM), la vidéosurveillance (sécurité, analyse routière...), la robotique, la défense (missiles, détection, véhicules autonomes...), l'astronomie, etc...

Par traitement d'images, on désigne l'ensemble des opérations sur les images numériques, qui transforment une image en une autre image, ou en une autre primitive formelle. Tant qu'à la vision par ordinateur, concerne la compréhension d'une scène ou d'un phénomène à partir des informations contenues dans l'image. Liant implicitement la perception, le comportement et le contrôle. Les domaines liés à ces deux disciplines vont du traitement du signal à l'intelligence artificielle, on ne saurait donc prétendre à l'exhaustivité, mais on s'intéresse plutôt à l'exploration d'un certain nombre de techniques importantes et actuelles.

Tout naturellement, l'idée d'augmenter le champ de vision s'est donc imposée avec comme espoir d'obtenir une vision omnidirectionnelle, c'est-à-dire un champ de vision à 360°, vue que certains types d'applications, la robotique comme exemple, exige un champ de vision plus large que celui des caméras standards.

Pour étendre le champ de vision, l'homme s'est inspiré du système de vision animale. En effet, dans leur milieu naturel, la connaissance de l'environnement est indispensable à la survie des espèces, afin de détecter efficacement le mouvement ou de repérer les prédateurs. La géométrie oculaire s'est adaptée à l'animal considéré et à son environnement.

1.1.1 La vision à large champ dans la nature

La vision est certainement la plus importante de nos cinq sens, C'est un processus complexe qui nécessite la collaboration de nombreux composants de l'œil humain et du cerveau. Un composant essentiel du système de vision, soit humain, animal ou synthétique, est le capteur de vision qui collecte les données de l'environnement. À cause de nos capteurs, les yeux, la vision humaine a un champ de vision (FOV) très limité par rapport à d'autres systèmes visuels trouvés dans la nature. La puissance de calcul massif offert par le cortex cérébral humain permet de réaliser de façon régulière des tâches de navigation et de reconnaissance, malgré le manque

d'informations visuelles. Au cours de leur vie, d'autres animaux ont aussi besoin de trouver leur chemin dans des environnements imprévisibles, de reconnaître la nourriture ou d'éviter les prédateurs, entre autres tâches vitales. Néanmoins, le cerveau des insectes, par exemple, dispose de 10^5 à 10^6 de neurones par rapport à 10^{11} que compte un cerveau humain, donc ils disposent de beaucoup moins de puissance de traitement [1]. Il est plausible que la mémoire et la capacité de calcul d'un tel petit cerveau sont inférieures à celles fournies par les puissants ordinateurs de nos jours. Cependant, les humains ne sont pas encore capables de construire des systèmes de vision similaires. Par conséquent, il est logique de penser que les performances de ces systèmes volants parfaits sont améliorées par le développement particulier de leurs yeux.

La majorité des arthropodes bénéficient d'un large champ de vision donné par leurs yeux composés ¹(figure 1.1). Un œil composé est un organe visuel qui comprend jusqu'à 30 000 ommatidies qui sont de minuscules capteurs qui distinguent entre la luminosité et l'obscurité, et parfois peuvent détecter la couleur. L'image perçue par l'arthropode est une combinaison d'entrées à partir de nombreuses ommatidies, qui sont orientées vers un point dans des directions légèrement différentes afin de couvrir un large champ de vision. Contrairement à d'autres types de yeux, il n'y a pas de lentille ou de rétine centrale, ce qui entraîne une faible résolution d'image ; cependant, il peut détecter un mouvement rapide et, dans certains cas, la polarisation de la lumière.

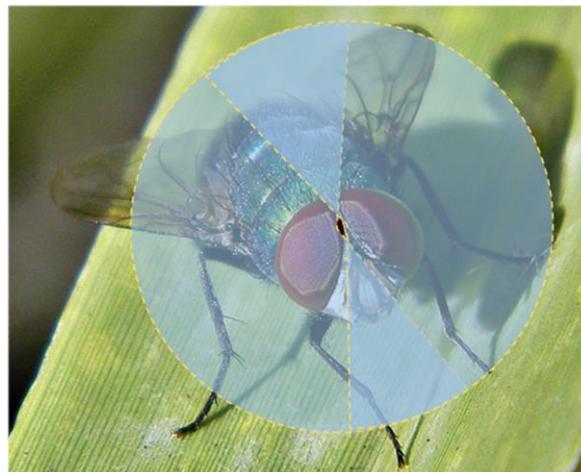


Figure 1.1 : Champ de vision d'une mouche.

Les yeux des crustacés sont constitués de combinaisons de miroirs et de surfaces de réfraction qui concentrent la lumière à différents points sur la rétine. Le Gigantocypris d'ostracodes ²(figure 1.2) est un crustacé avec de grands yeux réfléchissants, ayant une structure similaire à l'optique utilisée pour télescopes.

¹ Source de l'image de fond : [140]

² Source de figure : [142]

Ses yeux sont adaptés à des conditions de lumière très faible trouvés dans les profondeurs des océans, grâce aux miroirs paraboliques qui les composent.



*Figure 1.2 : Les yeux du
Gigantocypris d'ostracodes*

La conception de nouveaux capteurs de vision a été amplement inspirée par la nature. Un système de stéréovision omnidirectionnelle comme la caméra panoptique [2] *ViewPlus ASTRO Sensor*³(figure 1.3.a), nous fera certainement penser aux yeux de l'araignée sautante⁴(Figure 1.3.b) ou le *Scallop*⁵(Figure 1.3.c), deux espèces avec un système de vision constitué de multiple paire d'yeux, chacune étant orientée vers une direction.



*Figure 1.3 : (a) ViewPlus ASTRO Sensor (b) L'araignée sautante. (c) Le
Scallop.*

³ Source de figure : [143]

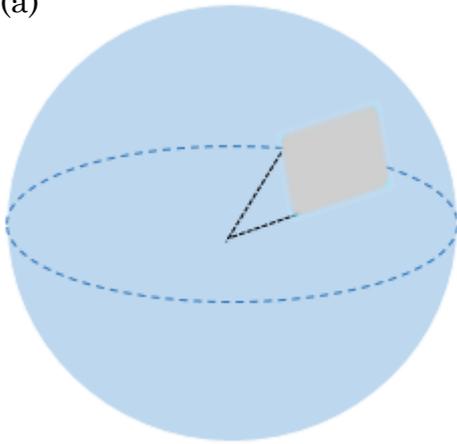
⁴ Source de figure : [144]

⁵ Source de figure : [145]

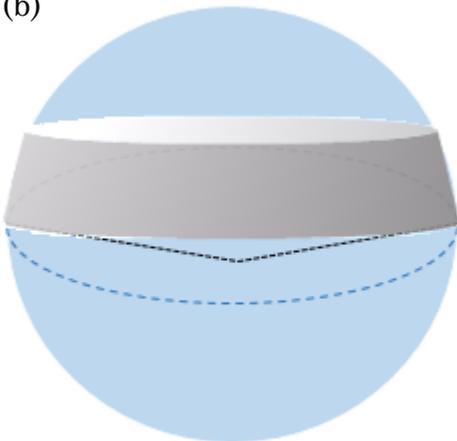
1.1.2 Les capteurs de vision omnidirectionnelle

La ⁶Figure 1.4 représente un classement des capteurs de vision selon le champ de vision qu'ils couvrent. Une caméra directionnelle a un champ de vision qui couvre moins d'un hémisphère. Elles sont aussi connues sous le nom de caméras standards car elles sont les plus couramment utilisées. Nous appelons capteur omnidirectionnel ou caméra omnidirectionnelle un capteur de vision qui a un champ de vision quasi sphérique atteignant 360 degrés en azimut. Les capteurs panoramiques ont aussi un champ de vision azimutale de 360° mais ne couvrent pas toute la sphère.

(a)



(b)



⁶ Source de la photo : [146]

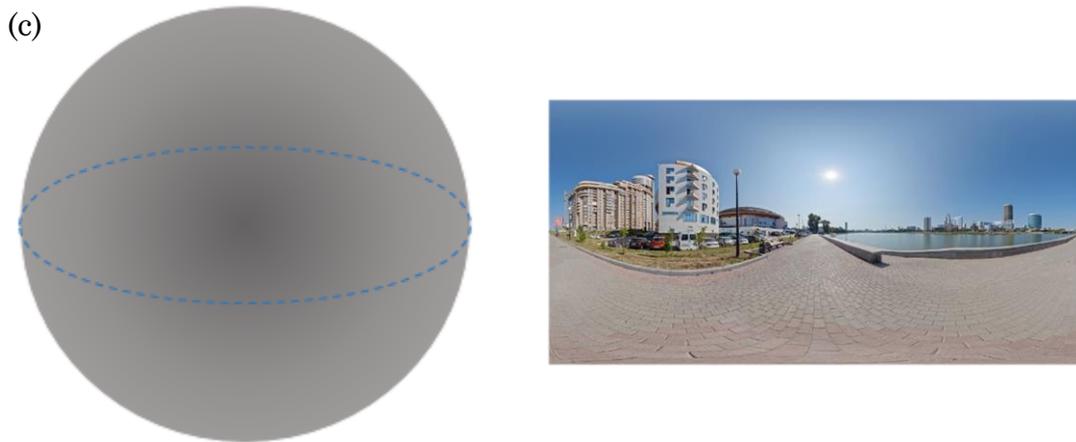


Figure 1.4 : (a) Champ de vision directionnel (b) Champ de vision panoramique
(c) champ de vision omnidirectionnel

La classification courante des capteurs omnidirectionnels existants est souvent liée aux technologies utilisées pour leur fabrication. Ainsi, pour améliorer le champ de vision de caméras conventionnelles, les chercheurs ont utilisé :

- ✓ Les lentilles spéciales.
- ✓ Plusieurs systèmes d'acquisition d'images au moyen de rotation des caméras ou en utilisant des structures de plusieurs caméras avec des champs de vue complémentaires.
- ✓ Les systèmes catadioptriques qui sont des combinaisons de caméras et de miroirs.

Si la scène est observée à partir d'un point dans l'espace le capteur possède un centre unique de projection soit (*Single View Point, SVP*).

1.1.2.1 Les lentilles spéciales

Parmi les moyens d'élargissement du champ de vision d'une caméra, la possibilité d'utiliser un ensemble complexe de plusieurs lentilles. Il s'agit de systèmes dioptriques. L'utilisation de ces optiques génère de fortes distorsions radiales au niveau de l'image acquise.

Un objectif hypergone ou lentille *fish-eye* ⁷(figure 1.5.a) est un objectif spécial à lentille très convexe avec une focale très courte qui donne un champ de vision hémisphérique ⁸(figure 1.5.b). En dépit des inconvénients mentionnés ci-dessus, le grand champ de vision offert par l'objectif fish-eye fait de ce type d'éléments dioptriques un choix attrayant pour plusieurs chercheurs. En 1993, Basu et al. [3] ont publié une étude fournissant deux solutions mathématiques pour compenser la distorsion fish-eye. La transformation fish-eye est définie comme étant une fonction de mappage à résolution variable qui génère une image à haute résolution dans la zone centrale et diminue de façon non linéaire la résolution vers la

⁷ Sources des figures : [147]

⁸ Source de figure : [148]

périphérie. Ils ont également souligné que la résolution variable des images peut être transformée en une fonction utile, car elle réduit la quantité d'informations à traiter et maintient au même temps un grand champ de vision.

En 1992 Dan Slater [4], a conçu un système de vision appelé *Spherecam* avec deux objectifs fish-eye montés dos-à-dos, offrant un champ de vision complètement omnidirectionnel avec la possibilité de prise de vue instantanée ⁹(Figure 1.5.c).



Figure 1.5 : (a) Lentille fisheye Samyang 8mm F3.5 (b) Une image obtenue avec une lentille fisheye. (c) Spherecam

Compte tenu des principaux points évoqués ci-dessus, nous voyons que les caméras à lentilles fish-eye peuvent fournir des images appropriées pour un certain nombre d'applications. Cependant leur champ de vision reste assez limité pour d'autres.

1.1.2.2 Les systèmes d'acquisition d'images multiples

Des vues panoramiques en haute résolution peuvent être obtenues en utilisant une mosaïque formée par plusieurs images fournies par une caméra rotative. Par exemple, Dans [5] on propose une méthode pour générer des images panoramiques par mosaïquage d'une multitude d'images perspectives, acquises par une caméra rotative ¹⁰(figure 1.6.a). Deux images panoramiques prises peuvent servir comme une paire stéréo panoramique. Le système décrit génère une image panoramique stéréo à l'aide de projections circulaires des images ou vidéos prises par une seule caméra. Une image omnidirectionnelle complète est obtenue en rassemblant les images acquises au moyen d'une procédure coûteuse en termes de temps de traitement, ce qui limite la convenance de telles méthodes dans des applications en temps réel.

Les yeux composés des insectes ont été une source d'inspiration pour un nouveau type de systèmes omnidirectionnels. En 2000, Fermuller et al. [6] ont développé un système de vision omnidirectionnelle à partir de plusieurs caméras orientées vers des directions différentes ¹¹(figure 1.6.b).

⁹ Source de figure : [149]

¹⁰ Source de figure : [5]

¹¹ Source de figure : [6]

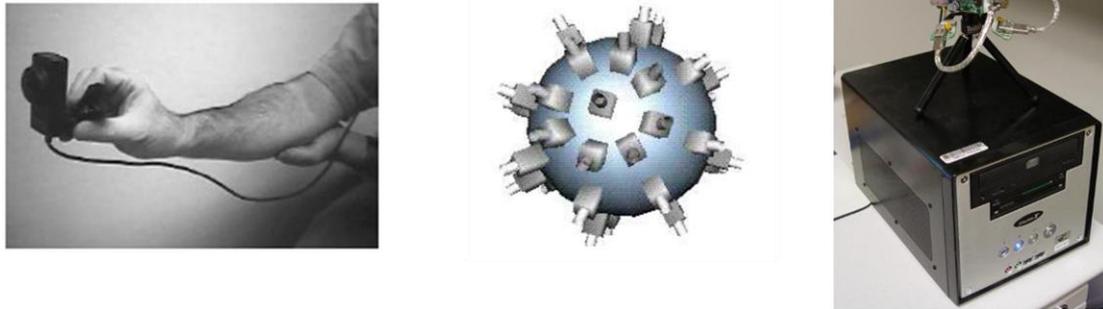


Figure 1.6 (a) Système de vision panoramique par mosaïquage. (b) Système de vision omnidirectionnel à multi-caméras. (c) système de vision panoramique à quatre caméras.

Un champ de vue omnidirectionnel complet n'est pas toujours nécessaire. Des applications telles que la vidéo-conférence n'ont besoin que d'un angle vertical de vue limité. Patil et Al., ont présenté dans [7] un système de vision à haute résolution pour la détection et le suivi des personnes, qui se compose d'un système à multi-caméras omnidirectionnelles destinées à être utilisées dans des environnements de réunion. Ce système ¹²(figure 1.6.c) appelé CAMEO est conçu pour surveiller les activités des gens dans les réunions de sorte qu'il peut générer un résumé sémantiquement indexé de ce qui s'est passé lors de la réunion. Ce genre d'arrangement n'offre pas un champ de vision quasi omnidirectionnel, mais permet de réduire le nombre de caméras et, par conséquent, les coûts de fabrication. Les capteurs de vision avec des champs de vue complémentaires ont surmonté quelques problèmes de caméras mobiles, mais le capteur devient plus difficile à calibrer car il implique plusieurs caméras, ainsi que les cartes d'acquisition d'images correspondantes.

En résumé, l'avantage d'utiliser des caméras rotatives ou des configurations de plusieurs caméras couvrant la scène offre une très bonne qualité d'image en haute résolution. Cependant, cette fonction très utile est entachée par un ensemble d'inconvénients. Dans le cas d'une caméra rotative, la partie de la scène derrière le plan d'image est toujours cachée, donc ce type d'appareils ne peut pas être utilisé dans des applications en temps réel. La vue omnidirectionnelle obtenue à partir de plusieurs images partielles est difficile à exploiter, coûteuse et peut contenir des discontinuités.

1.1.2.3 Les systèmes catadioptriques

Une autre façon d'obtenir des images omnidirectionnelles est en pointant une caméra vers un miroir spécialement conçu. La combinaison de miroirs (catoptriques) et les lentilles (dioptriques) est connu comme catadioptriques. Dans l'histoire de la Grèce antique les propriétés des miroirs courbés de focalisation de

¹² Source de figure : [7]

la lumière ont été exploitées pour brûler les navires ennemis ¹³(figure 1.7.a) [8]. Les catadioptriques ont également été utilisées au XVII^{ème} siècle par Isaac Newton qui a développé un télescope réflecteur (réflecteur newtonien) qui utilise des miroirs au lieu des lentilles pour corriger l'aberration chromatique (Halo). Observant la scène environnante au moyen de surfaces réfléchissantes n'est pas une idée nouvelle et un exemple classique, est le dessin de l'artiste graphique néerlandais Escher (1898 - 1972), illustré à la ¹⁴figure 1.7.b. Cependant, ce qui a réellement déclenché la recherche sur les systèmes catadioptriques était l'utilisation de caméras modernes qui ont fourni une image de bonne qualité de la scène réfléchi.

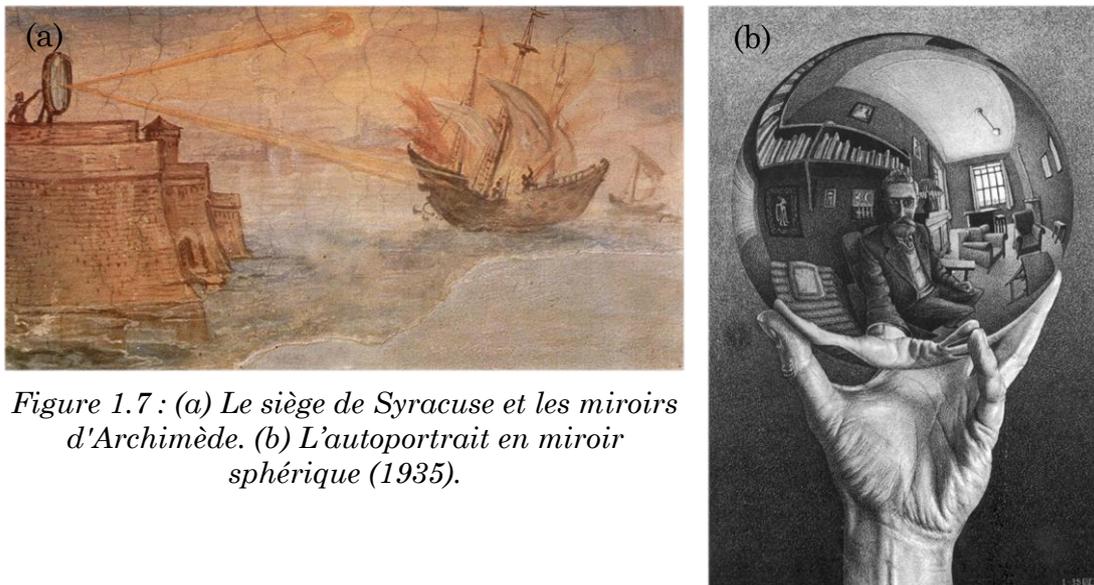


Figure 1.7 : (a) Le siège de Syracuse et les miroirs d'Archimède. (b) L'autoportrait en miroir sphérique (1935).

Une utilisation précoce des catadioptriques pour une application réelle a été proposée par Rees en 1970 [9]. Il a inventé une caméra panoramique de télévision basée sur un miroir de forme hyperbolique convexe. Vingt ans plus tard, les chercheurs concentrent à nouveau leur attention sur les possibilités offertes par les systèmes catadioptriques, principalement dans le domaine de la robotique. En 1990, l'équipe japonaise de Mitsubishi Electric Corporation dirigée par Y. Yagi [10] a étudié les scènes panoramiques générées en utilisant un capteur à base de miroir conique. Le capteur, appelé COPIS, a été utilisé pour générer la carte de l'environnement à l'intérieur d'une scène à l'aide d'un robot mobile.

À partir des années quatre-vingt-dix, l'attention accordée à la vision omnidirectionnelle et sa base de connaissances n'a cessé de croître comme le nombre de chercheurs impliqués dans l'étude de caméras omnidirectionnelles a augmenté. De nouveaux modèles mathématiques pour projection catadioptrique ont été établis, par conséquent, de meilleurs capteurs catadioptriques performants

¹³ Source de figure : [150]

¹⁴ Source de figure : [162]

sont apparus. En résumé, le grand intérêt suscité par les catadioptriques est due à leurs propres avantages par rapport à d'autres systèmes omnidirectionnels. Parmi ces avantages : la prise de vue omnidirectionnelle instantanée, une bonne résolution, et en particulier le prix et la compacité. Pour une discussion détaillée sur la théorie et la pratique de réalisation des catadioptriques, le lecteur est renvoyé au chapitre 2 de cette thèse.

1.2 Contexte de la thèse

Les travaux de recherche de cette thèse ont commencés et ont été réalisés principalement dans le Laboratoire d'Electronique et de Traitement du Signal/Géomatique, connu à l'époque sous le nom "LETS" avant sa fusion avec le Laboratoire de Géomatique. Il fait partie des structures de recherche affiliées à la Faculté des Sciences de Rabat, Université Mohamed V-Agdal. Les axes de recherche du laboratoire couvrent des domaines tels que la vision par ordinateur, le traitement du signal notamment de l'image, la télédétection et la géomatique. Avec des applications en robotique, environnement, agriculture et les domaines de l'imagerie médicale.

Dans le contexte des travaux menés dans le domaine de la vision artificielle et du traitement des images numériques, le large champ de vision offert par les caméras omnidirectionnelles est apparu comme une solution possible aux problèmes spécifiques introduits par l'utilisation de caméras traditionnelles. De plus, l'utilité de ces techniques et le défi de leur application aux systèmes autonomes nous ont motivés pour lancer une recherche dans le domaine des capteurs de vision catadioptriques. La première étape était une phase de recherche d'un an en 2011 où les principes fondamentaux de la vision omnidirectionnelle ont été étudiés. Les algorithmes et techniques de base ainsi élaborés nous ont permis d'étendre l'étude vers une application dans le domaine de la robotique. Et aussi dans le domaine de l'énergie photovoltaïque en collaboration avec le Laboratoire d'Energie Solaire et Environnement représenté par le Professeur Mohammed Najib BARGACH.

1.3 Problématique

L'utilisation des systèmes autonomes est une option attrayante dans les situations où l'intervention humaine est trop coûteuse, dangereuse, ou tout simplement dans le cas où la tâche peut être bien gérée par une machine que par un humain. Les systèmes autonomes doivent interagir avec l'environnement en utilisant une combinaison de capteurs embarqués, et éventuellement traiter les données obtenues et de les transformer en informations utiles pour la prise de décision ou pour l'interprétation humaine. Par conséquent, il est essentiel de fournir à ces systèmes un modèle de la scène réelle ou la possibilité de construire un tel modèle de façon autonome. La construction des systèmes de vision artificielle non intrusive de l'environnement, ainsi que le développement des algorithmes de traitement de l'information visuelle avaient motivé notre recherche.

Toutefois, les capteurs catadioptriques de vision omnidirectionnelle présentent deux inconvénients majeurs par rapport aux caméras perspectives, qui sont la résolution non uniforme, ainsi que de fortes distorsions radiales introduites dans les images omnidirectionnelles. Ces lacunes affectent l'applicabilité des outils classiques de traitement à ce type d'images. Il convient aussi de préciser que la géométrie de formation des images catadioptriques est totalement différente de celle des images perspectives. Par conséquent, la redéfinition des outils mathématiques de modélisation, calibrage, stéréovision est fermement nécessaire dans ce cas.

La résolution non uniforme et les distorsions sont dues à la forme du miroir utilisé. La figure 1.8 illustre comment deux surfaces égales seront projetées orthographiquement sur le plan image selon deux résolutions différentes en fonction de leurs positions sur le quadrique de révolution utilisé.

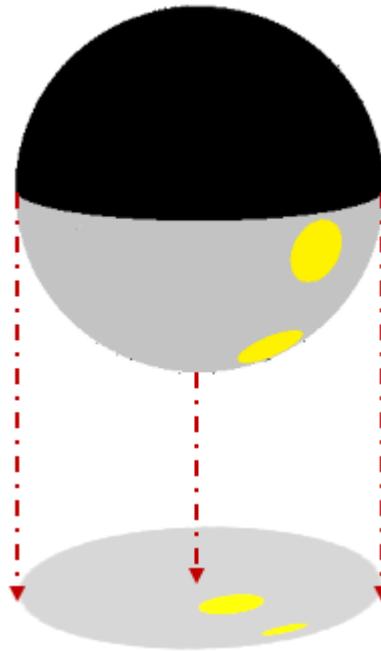


Figure 1.8 : Formation d'une image catadioptrique par projection orthographique.

Un autre problème qui surgit lors du traitement des images catadioptriques par des outils classiques conçus pour les images perspectives, est celui du voisinage pixellique. Ce problème est bien décrit dans la section 13.4.

Compte tenu de la nature spéciale des images omnidirectionnelles, ainsi que de leurs systèmes d'acquisition, nous nous sommes intéressés au développement d'outils techniques et théoriques permettant l'exploitation efficace des systèmes de stéréovision omnidirectionnelle, comme premier axe de recherche. Le deuxième axe a concerné l'adaptation des algorithmes et des outils classiques de traitement

aux images omnidirectionnelles afin d'en tirer profit tout en respectant leurs propriétés géométriques.

1.4 Contributions et organisation du mémoire

Ce document est constitué de quatre chapitres relatant les principaux travaux effectués dans le domaine de la vision omnidirectionnelle :

- Le premier chapitre présente un aperçu général de la vision omnidirectionnelle. Plusieurs qualités importantes des capteurs de vision à large champ de vue sont discutées, et un classement général est introduit avec des exemples des capteurs omnidirectionnels les plus courants.
- Le chapitre 2, dans lequel les systèmes d'acquisition 3D omnidirectionnels sont classés selon la technologie appliquée dans leur fabrication et leurs propriétés optiques. Le chapitre commence par une présentation des systèmes d'acquisition 3D actifs. Ensuite, le chapitre présente une étude sur les systèmes d'acquisition 3D passifs, notamment les systèmes de stéréovision omnidirectionnelle catadioptriques, qui sont comparés en termes de configuration, résolution, isotropie et leur capacité à fournir des mesures 3D. La géométrie épipolaire pour les systèmes catadioptriques est également présentée ici. Nous présentons par la suite une solution matérielle qui consiste en un capteur omnidirectionnel stéréo (omnistéréo) à caméra unique. Son architecture, étalonnage, formules de triangulation, ainsi que les résultats d'estimation de la profondeur obtenus sont également décrits.
- Dans le troisième chapitre nous présentons les différentes approches adoptées dans le traitement des images omnidirectionnelles. Un premier algorithme permettant le déballage de ces images est proposé. Par la suite, le modèle de projection unifié, et la projection stéréographiques sont introduits. Ces deux outils ont été utilisés pour développer deux algorithmes d'extractions de primitives, adaptés à la géométrie de formation des images sphériques. Il s'agit d'un opérateur de détection de contours, basé sur un modèle sphérique de charges électriques virtuelles. Ainsi qu'une adaptation de l'algorithme de détection des coins de Harris aux images catadioptriques.
- Dans le chapitre 4 nous proposons l'utilisation des systèmes de vision omnidirectionnels pour améliorer les procédures de détection et de suivi des cibles mobiles. Nous présentons d'abord un robot mobile, doté d'un système de vision artificielle omnidirectionnelle, capable de détecter et de suivre une cible en mouvement de façon autonome. Ensuite nous présentons un suiveur solaire équipé d'une caméra catadioptrique. Lequel est un dispositif qui permet l'orientation optimale de panneaux photovoltaïques, dans le but d'augmenter leur rendement énergétique.
- Finalement la conclusion générale reprend les points essentiels soulevés.

2. Les systèmes d'acquisition 3D omnidirectionnels

2.1 Introduction

La vision est un mode de perception privilégiée du point de vue biologique. Elle est une pierre angulaire des systèmes de perception artificielle. Le développement de la technologie des caméras a été primordial dans l'évolution de la vision par ordinateur. La haute fréquence d'acquisition, la miniaturisation, et la haute résolution sont autant des atouts qui ont facilité l'intégration de ces capteurs dans plusieurs domaines notamment en robotique. Les caméras sont de plus en plus performantes en termes de qualité, précision et variété des informations acquises. Ainsi elles ont permis aux machines d'avoir la capacité de restituer une image sensorielle de l'environnement proche de celle perçue par l'être humain.

Le progrès technologique qu'a connu le domaine d'industrie des caméras a permis de pallier de nombreux défauts, tels que la sensibilité aux conditions d'éclairage ou le flux volumineux des données à traiter. En revanche, l'inconvénient majeur de tels capteurs se situe d'abord au niveau de leur champ de vision restreint. Les caméras *directionnelles* ont un champ de vision qui couvre moins d'un hémisphère, elles sont les plus couramment utilisées et commercialisées, c'est pour cette raison que le terme *cameras standards* est fréquemment utilisé pour les désigner. Pour augmenter le champ de vision de ces caméras, plusieurs solutions ont été proposées dans la littérature, telles que les objectifs spéciaux à distance focale très courte, les systèmes rotatifs associés à une caméra. Il s'agit des techniques de vision omnidirectionnelle. Généralement une caméra possédant un champ de vision de 360° est dite *panoramique*. Le terme *omnidirectionnelle* est réservé aux caméras dont le champ de vision dépasse l'hémisphère.

La robotique mobile est l'un des domaines qui exploite très avantageusement la vision omnidirectionnelle. Un grand nombre de travaux traitent de son utilisation dans ce cadre, soit pour la navigation [11], la localisation ou le SLAM (*Simultaneous Localization and Mapping*), qui a pour but la localisation et la cartographie de l'environnement en temps réel [12].

Une grande part des robots opèrent dans un contexte de scène bidimensionnelle (2D). Néanmoins dans certains types d'application, il est nécessaire au robot de percevoir la scène environnante en relief (3D). Ce besoin a donné naissance à de nombreuses études et théories.

2.2 Les systèmes d'acquisition 3D actifs

Les systèmes d'acquisition 3D actifs acquièrent l'information tridimensionnelle sur un objet ou une scène par émission-réception d'un rayonnement. Différents types d'ondes sont utilisées : Ultrasons, Lumière, Rayons X. La profondeur de l'objet peut être mesurée de deux façons : par **télémétrie** ou par **triangulation active**.

2.2.1 La télémétrie

Les télémètres sont des appareils permettant de mesurer une distance en exploitant la rétrodiffusion d'une onde par la cible visée. Les télémètres fournissent généralement la profondeur d'un unique point dans la direction de l'onde émise. Pour scanner une scène entière il est nécessaire d'utiliser un système de balayage par mécanisme rotatif ou à système de miroirs rotatifs.

2.2.1.1 La télémétrie par mesure du temps de vol

Cette technique est basée sur l'émission d'un signal impulsionnel, la mesure du temps de parcours de l'impulsion entre l'émission et la réception permet l'estimation de la distance entre le télémètre et la cible : Une partie de l'onde émise est rétrodiffusée dans une direction approximativement quasi-parallèle à l'onde incidente. La vitesse de propagation de l'onde v étant connue, si τ est la durée d'aller-retour de l'impulsion, la distance D à mesurer est donnée par la relation suivante: $D = \frac{1}{2}v\tau$.

Les télémètres à ultrasons se basent sur la mesure du temps de vol de l'onde ultrasonore acoustique pour estimer la distance. Ils sont simple électroniquement grâce à la faible vitesse de propagation des ondes sonores en comparaison avec celle de la lumière, et par conséquent ils sont moins chères. Toutefois ce type de télémètres offrent une résolution azimutal très faible d'environ 30° , d'une portée réduite, sont sensibles à la température, à la pression et également aux interférences.

Ce même principe est utilisé par les **radars** à impulsions « *radio detection and ranging* », qui opèrent dans une plage de fréquences électromagnétiques variant de la bande HF (3-30 MHz) jusqu'à la bande W (75–110 GHz) [13]. Les **lidars** « *light detection and ranging* » sont aussi des appareils qui utilisent le principe de mesure du temps de vol pour mesurer la distance. A la différence des radars, les lidars opèrent dans le domaine du visible, l'ultraviolet (UV), ou l'infrarouge (IR), mais la technologie la plus répandue utilisée par les lidars est basé sur l'émission d'un faisceau laser à impulsion.

Les **télémètres laser** (*laser rangefinders*) sont largement utilisés dans un vaste éventail de domaines liés à la vision par ordinateur, tels que la modélisation 3D d'objet, ou la reconnaissance d'objet en 3D. Cette technologie constitue le noyau de nombreux systèmes d'acquisition 3D notamment dans le cas de la vision à large champ.

De nombreux concepts combinant la télémétrie laser et la vision omnidirectionnelle ont été proposés récemment dans la littérature. Dans [14] les auteurs proposent deux méthodes de suivi et de poursuite de personne en utilisant une caméra omnidirectionnelle et un télémètre laser, embarqués sur un robot mobile. Un tel système fournit des capacités importantes d'interaction humain-robot et d'assistance de l'homme dans différents milieux. C'est pourquoi le suivi avec robots mobiles a constitué un domaine de recherche actif avec le succès de

nombreux systèmes développés, comme pour le guidage dans un musée [15], l'assistance dans un hôpital [16], ou le suivi de piétons [17]. L'utilisation d'une caméra omnidirectionnelle avec un champ de vision de 360° apporte plus de fiabilité au suivi des personnes mobiles, vu que l'occlusion dans ce cas est quasi-improbable. La première méthode présentée dans ce travail a été basée sur le suivi visuel. L'approche proposée utilise la compensation par ego-mouvement après transformation des images omnidirectionnelles en images perspectives, et la différence frame par frame pour détecter le mouvement. Un filtre à particules est ensuite utilisé pour suivre le mouvement et un algorithme EM (*Expectation–Maximization*) pour trouver le centre de gravité de l'objet suivi. La télémétrie laser a été utilisée comme outil secondaire pour l'estimation de la profondeur. Cette méthode est valable pour les faibles vitesses et dans le cas d'un environnement contrôlé. Sa performance se dégrade quand les conditions deviennent plus complexe. Tandis que la deuxième méthode combine la télémétrie laser comme outil principal et la vision omnidirectionnelle pour effectuer le suivi d'une personne dans un environnement dynamique non contrôlé tel que le cas d'une scène en plein air. Cette approche se sert du télémètre laser pour extraire la position 3D relative des silhouettes qui pouvaient désigner une personne et exploite ces mesures avec un filtre probabiliste d'association des données (PDAF). Ce système d'acquisition 3D hybride qui a été embarqué sur le robot *Segway RMP*¹⁵(figure 2.1.a), est configuré comme l'illustre la figure 2.1.b de telle façon à ce que chaque point de l'espace environnant soit repérable par sa profondeur d et son orientation β acquises par le télémètre laser. Ainsi que par ses coordonnées sphériques (θ, ϕ) par rapport au repère miroir de la camera catadioptrique. Considérant ${}^A T$ la transformation correspondante au repère robot, et ${}^L T$ ${}^C T$ sont respectivement les transformations correspondantes au télémètre laser et à la caméra. Les mesures d et β correspondent à un point 3D dont les coordonnées homogènes sont données par :

$$P = {}^A T {}^L T \begin{pmatrix} d \cos \beta \\ d \sin \beta \\ 0 \\ 1 \end{pmatrix} \quad 2.1$$

Ce même point est repéré par la camera par les coordonnées suivantes :

$${}^C P = \begin{pmatrix} {}^C x \\ {}^C y \\ {}^C z \\ 1 \end{pmatrix} = ({}^C T^{-1}) ({}^A T^{-1}) P \quad 2.2$$

Les coordonnées (u, v) du point P sur le plan image sont données par :

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} x_c + r \cos \theta \\ y_c + r \sin \theta \end{pmatrix}, \quad r = f(\sqrt{\tan^2 \phi + 1} - \tan \phi), \quad \theta = \text{atan2}({}^C x, {}^C y), \quad \phi = \text{atan2}({}^C z, \sqrt{{}^C x^2 + {}^C y^2}) \quad 2.3$$

¹⁵ Source de figure : [14]

f est la distance focale de la projection combinée, x_c et y_c sont les coordonnées du centre de l'image catadioptriques.

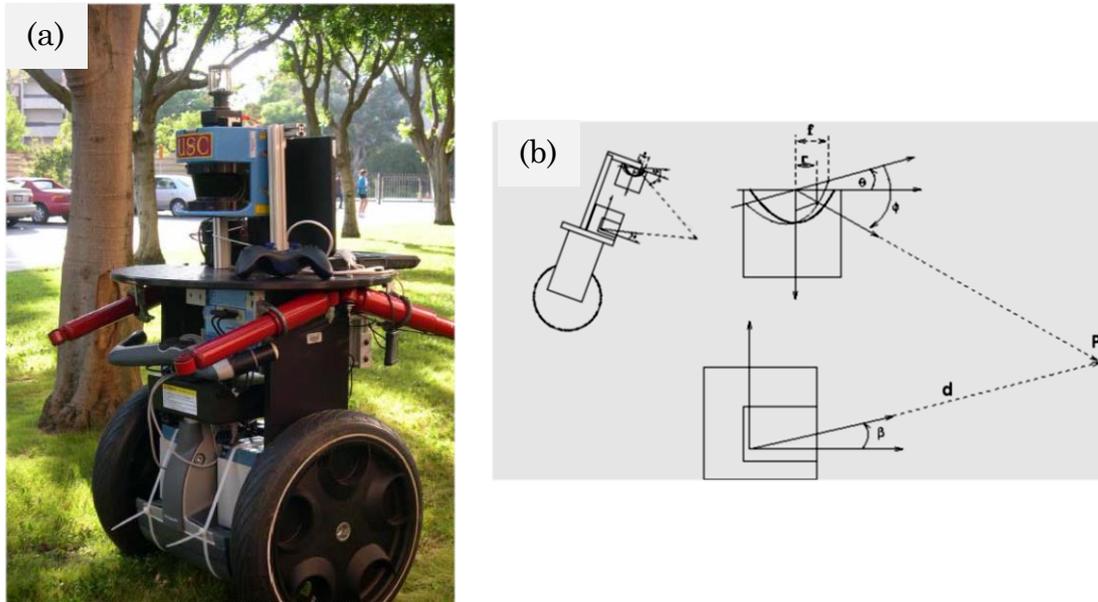


Figure 2.1 : (a) Le robot Segway RMP. (b) Configuration du système laser-camera omnidirectionnelle.

La cartographie est l'une des applications qui profite du large champ de vision offert par les cameras omnidirectionnelle, qui permettent d'économiser considérablement le temps d'acquisition. Dans [18] un système de cartographie 3D basé sur le couplage vision omnidirectionnelle et la télémétrie laser a été proposé. Ce système est constitué d'un capteur de vision omnidirectionnelle stéréo et d'un télémètre laser, le tout embarqué sur un robot mobile (figure 2.2.a). D'après les auteurs, un système d'acquisition omnidirectionnelle stéréo ne peut pas à lui seul garantir une fiabilité satisfaisante. D'où le besoin de lui adjoindre la télémétrie active par laser, ce qui permet de générer des cartes 3D avec plus de précision (figure 2.2.c). Puisque les deux capteurs peuvent détecter différentes parties d'un objet, les cartes de grille probabilistes sont d'abord générées distinctement par intégration temporelle des données provenant de chaque capteur. Les deux cartes qui en résultent sont ensuite intégrées en utilisant une règle d'intégration logique. Une méthode d'estimation ego-mouvement est également décrite, qui est nécessaire à l'intégration des données acquises par les capteurs à différentes positions.

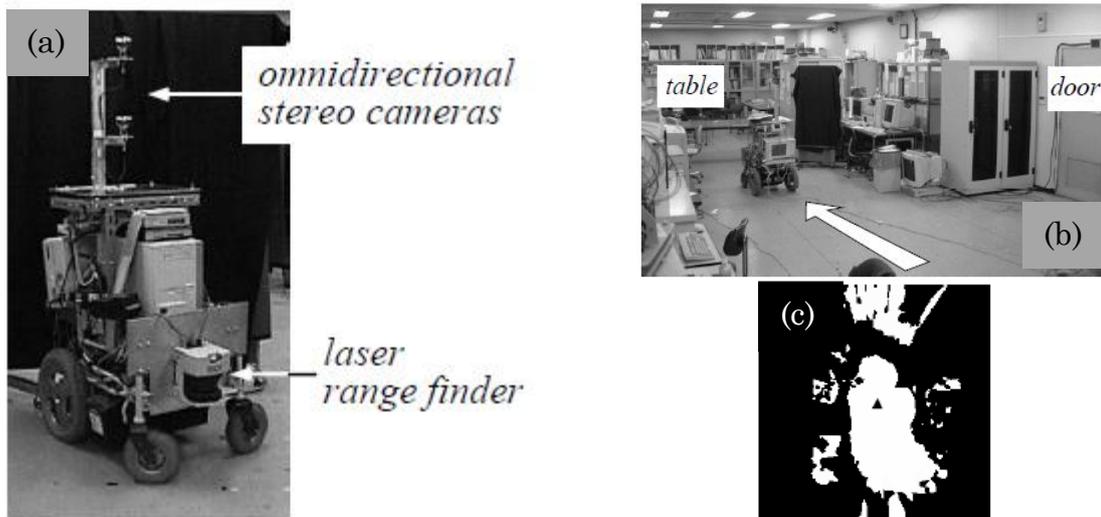


Figure 2.2 (a) Robot mobile dédié à la cartographie. (b) Exemple d'une scène. (c) La carte générée représentant l'espace libre.

Google Street View [19] est l'une des applications qui symbolise le triomphe des systèmes d'acquisition 3D actifs. Ce service lancé en 2007 dans le but d'être intégré dans *Google Maps* et *Google Earth* permet de naviguer virtuellement dans les rues de différentes villes à travers le monde. Les images sont enregistrées par une flotte de véhicules ¹⁶(figure 2.3.a) équipées de systèmes de vision artificielle, qui parcourent les différents emplacements à cartographier.

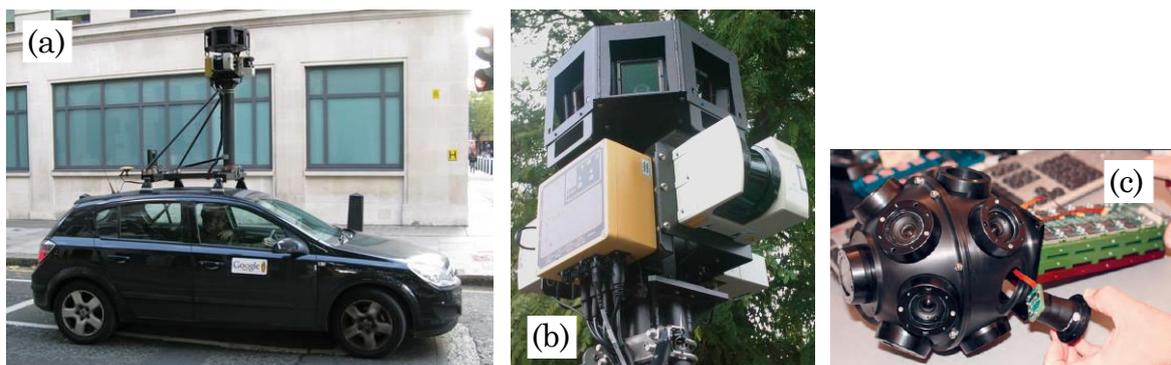


Figure 2.3 : (a) Une voiture de la flotte Google Street View équipée d'un système de camera R5. (b) système de camera R5. (c) système de camera R7.

Google a utilisé plusieurs générations de caméras pour capter les images *Street View*. La plupart des images récentes ont été acquises par le système de caméra de cinquième génération développé par *Google*, connu sous le nom *R5* ¹⁷(Figure 2.3.b). Le système *R5* se compose d'une rosette de huit caméras, chacune a une résolution

¹⁶ Source de figure : [151]

¹⁷ Source de figure : [152]

de 5 millions de pixels, et d'un objectif fish-eye orienté verticalement vers le haut (figure 2.3.b). Le successeur de la *R5*, connu sous le nom *R7*¹⁸(figure 2.3.c), se compose d'une rosette de 15 appareils du même type que la *R5* et est destiné à capturer des images à haute résolution avec un champ de vision plus large que la *R5*. Les systèmes d'acquisition des véhicules récents de *Google* sont aussi équipés de télémètres laser qui mesurent avec précision la profondeur d'une fenêtre verticale de points sur les deux côtés et à l'avant du véhicule. Les systèmes anciens se basaient sur le calcul du flux optique entre les images successives pour extraire l'information tridimensionnelle.

La caméra à temps de vol (caméra *TOF*) est une solution matérielle plus simple, compacte et plus accessible au grand public. Il s'agit d'un système de caméra qui résout la distance entre la caméra et la scène pour chaque point de l'image. Le télémètre incorporé est de type LIDAR à balayage. Avec une portée maximale d'environ 60m et une résolution en distance d'environ 1cm, ces caméras ont émergées à partir de 2000 et ont été utilisées dans diverses applications civiles telles que : l'aide à la conduite, Les interfaces homme-machine, les jeux vidéo, et la robotique.

2.2.1.2 La télémétrie par mesure du déphasage

La technique de télémétrie par décalage de phase se base sur l'émission d'un faisceau d'onde, généralement de type laser, modulé en amplitude par un signal sinusoïdal ou carré de fréquence constante f_m . Le déphasage $\Delta\varphi$ entre l'onde rétrodiffusée par l'objet et l'onde émise permet d'estimer la distance séparant la source de la surface de réflexion. Le décalage de phase est donné par la relation [20] suivante :

$$\Delta\varphi = 4\pi f_m \frac{D}{v} \text{mod}(2\pi) \quad 2.4$$

2.2.1.3 La télémétrie par mesure de la fréquence

Une autre méthode d'estimation de la distance par mesure de la fréquence, repose sur la modulation en amplitude de l'onde émise par un signal lui-même modulé en fréquence. La distance est mesurée par la valeur d'une fréquence de battement issue du décalage entre l'onde émise et l'onde rétrodiffusé. La relation [21] reliant la distance et la fréquence intermédiaire du signal s'écrit sous la forme :

$$D = \frac{vTf_i}{4\Delta F} \quad 2.5$$

T étant la période du signal de modulation f_i : la fréquence intermédiaire ou la fréquence de battement, et ΔF est l'excursion en fréquence du signal de modulation.

¹⁸ Source de figure : [19]

2.2.2 La triangulation active

La télémétrie par triangulation active est une technique qui se base également sur l'émission de la lumière structurée pour sonder la profondeur d'un objet ou d'une

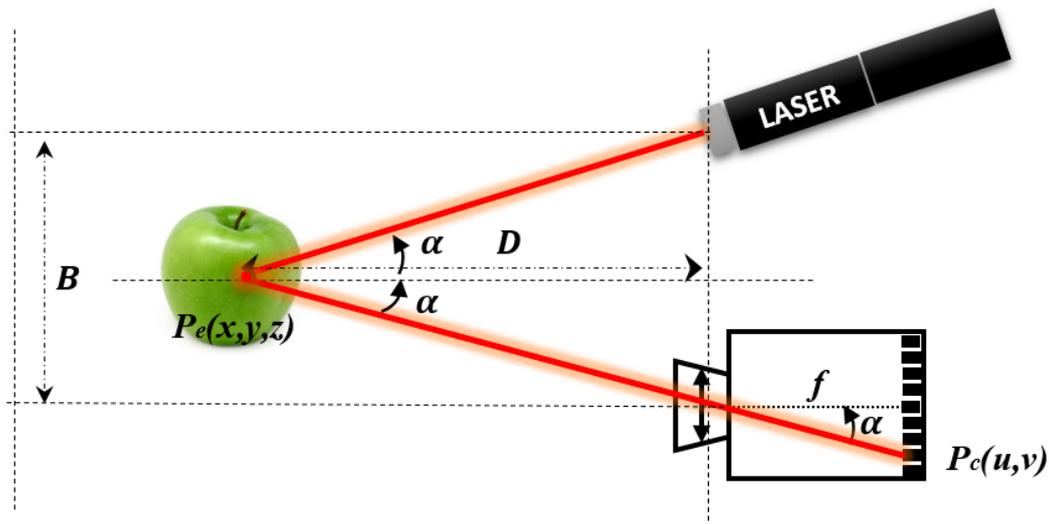


Figure 2.4 : Représentation de la triangulation active par émission d'un faisceau laser.

scène. A la différence des autres techniques de télémétrie active, cette approche se sert de la caméra pour intercepter la lumière rétrodiffusée. La lumière structurée dans sa forme la plus simple est **un rayon** qui se manifeste par un point de surbrillance sur l'objet visé. La direction du rayon étant connue, et la caméra étant calibrée, la profondeur peut alors être déterminée par une simple triangulation (figure 2.4).

Généralement, pour effectuer l'acquisition tridimensionnelle d'un objet ou d'une scène, le balayage se fait par **projection d'un plan de lumière structurée** (figure 2.5) au lieu d'un faisceau ce qui permet d'accélérer le processus d'acquisition. Dans ce cas un segment de surbrillance est intercepté par la caméra, ce segment appartient à un plan de projection d'équation connue. La profondeur des pixels associés à des directions dans l'espace 3D est estimée par triangulation.

Pour approcher l'acquisition 3D en temps réel, il est nécessaire de s'émanciper des systèmes à balayage. L'une des alternatives est la projection d'un **motif bidirectionnel** au lieu d'un faisceau ou d'un plan de **lumière structurée**. Ce motif peut être sous la forme d'une grille de segments, ou une matrice de points. Dans ce cas, l'estimation de la profondeur se fait par l'étude des déformations du motif projeté (figure 2.6).

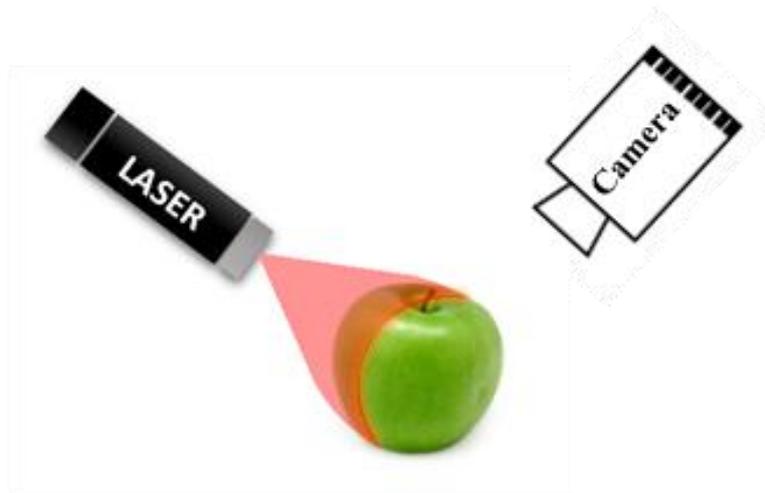


Figure 2.5 : triangulation active par émission d'un plan laser.

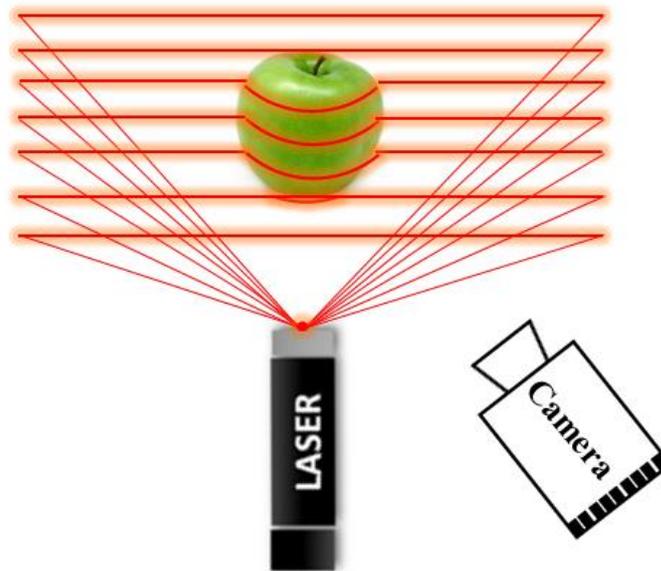


Figure 2.6 : Acquisition 3D active par émission d'un motif de segments laser.

Un exemple d'application de cette technique dans le domaine commercial, est le fameux contrôleur de jeu *KINECT™*¹⁹(figure 2.7.a) lancé par *Microsoft* en novembre 2010. Ce dispositif a connu un grand succès, et il a été utilisé dans une multitude d'applications, de la robotique à la réalité augmentée. *KINECT* [22] interprète les scènes en 3D en projetant en permanence un motif de points

¹⁹ Source de figure : [153]

infrarouge ²⁰(figure 2.7.b). L'étude de déformation de ce dernier permet d'extraire l'information sur la profondeur même en conditions de faible éclairage.

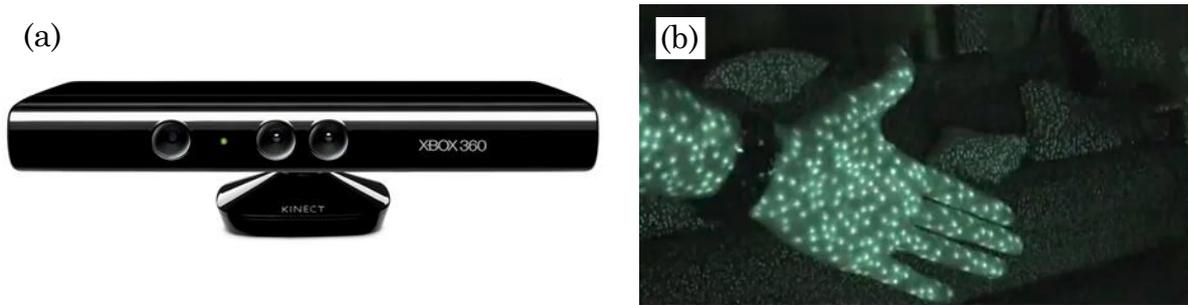


Figure 2.7 : (a) Le dispositif d'acquisition 3D KINECT (b) Motif de lumière structurée projeté par KINECT.

En vision à large champ, la triangulation active n'est pas privilégiée. Car concevoir un procédé qui permet la projection omnidirectionnelle d'un motif de lumière structurée est une tâche assez complexe, le calibrer en est une autre.

Quelques travaux de recherche ont abordé ce thème, dans [23] les auteurs proposent un dispositif omnidirectionnel de reconstruction 3D. Ce capteur ²¹(figure 2.8.a) est constitué de deux cameras CCD standards monochromatiques,

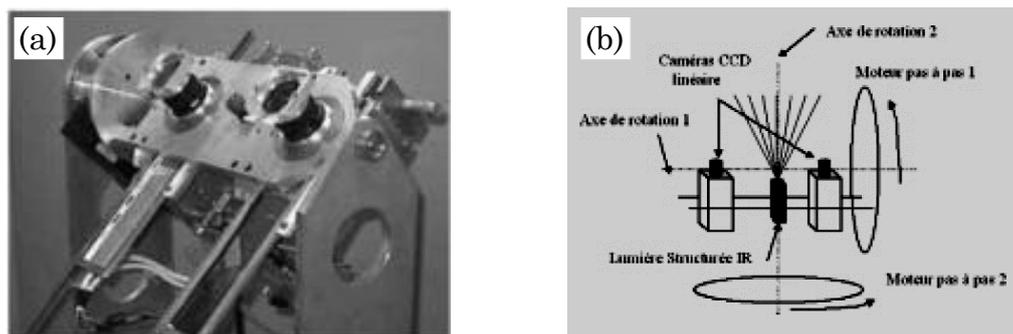


Figure 2.8 : (a) Photographie du capteur de vision omnidirectionnelle. (b) Configuration du capteur de vision omnidirectionnelle.

ainsi qu'un projecteur de lumière structurée infrarouge. L'architecture mécanique qui a été adoptée pour étendre le champ de vision de ces capteurs, repose sur l'utilisation de deux axes de rotation ; un horizontal et l'autre vertical. Cette configuration (figure 2.8.b) permet de couvrir une sphère d'observation. Tant qu'à la reconstruction tridimensionnelle, elle a été basée sur un algorithme d'appariement (mise en correspondance), lié discrètement au principe de l'acquisition multi spectral des capteurs CCD utilisés. Les informations géométriques de la scène, sont extraites à partir des primitives contenues dans les images des deux canaux infrarouges, et non pas par étude de la déformation du motif projeté. Les primitives utilisées sont les indices pixels des points lasers projetés. L'inconvénient majeur d'un tel concept provient du temps d'acquisition qui est assez long.

²⁰ Source de figure : [154]

²¹ Source de figure : [23]

D'autres solutions alternatives ont été présentées dans la littérature, telle que celle proposée dans [24]. Il s'agit d'un concept qui consiste à coupler une caméra catadioptrique est un projecteur de lumière structurée omnidirectionnel. Ce projecteur est disposé coaxialement sous la camera catadioptrique, le faisceau laser émis est diffusé par un miroir conique, ce qui permet de couvrir un large champ de projection. Cette configuration est bien illustrée sur la ²²figure 2.9.

La troisième dimension de la scène est perçue par le croisement des rayons lumineux émis par le laser avec ceux observés par la caméra, en d'autres termes, par exécution d'une triangulation. Les points du motif projeté par le laser sont réfléchis par le miroir conique et deviennent des tâches de lumière sur la scène. Le miroir parabolique reflète la scène entière vers la camera. Tous les points éclairés par le laser peuvent être identifiés. Enfin avec une modélisation exacte des deux composantes principales de ce dispositif, une estimation précise de la profondeur peut être réalisée. Une formulation mathématique pour la calibration de ce système d'acquisition a été proposée par les mêmes auteurs dans [25].

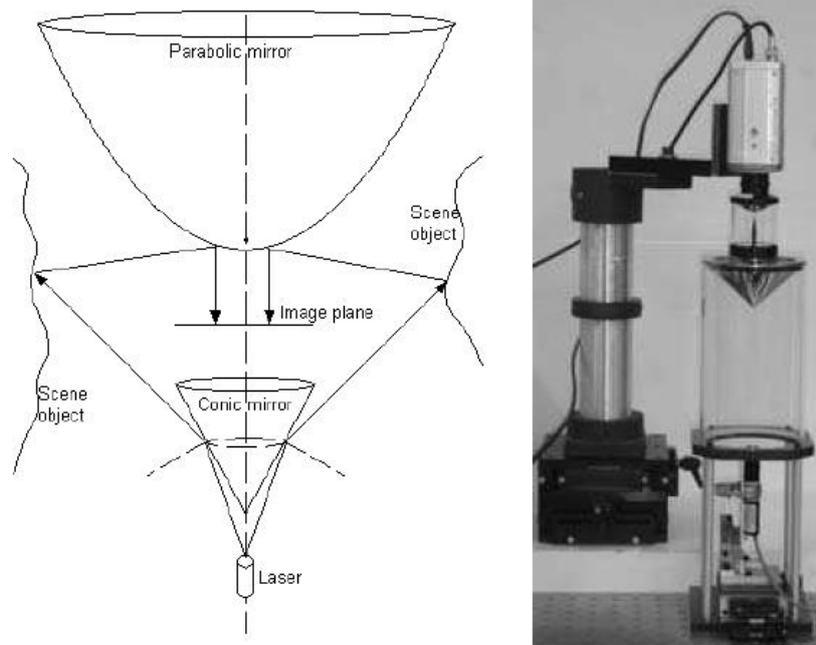


Figure 2.9 : Système d'acquisition 3D omnidirectionnel actif.

Les auteurs de [26] présentent un concept similaire ²³(figure 2.10.a), sauf qu'ils suggèrent l'utilisation d'un système de projection laser à quatre sources pour éviter l'utilisation d'un second miroir convexe. La caméra catadioptrique utilisée incorpore un miroir parabolique qui exige l'association d'une lentille télécentrique. Le système d'imagerie résultant est alors orthographique, ce qui facilite le calcul de la profondeur par triangulation comme illustré sur la figure 2.10.b.

²² Source de figure : [24]

²³ Source de figure : [26]

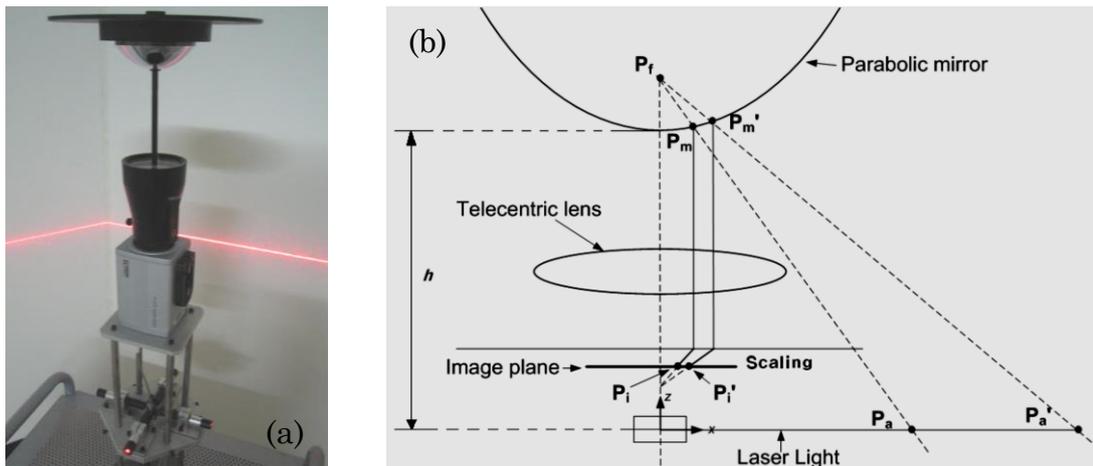


Figure 2.10 : (a) Illustration d'un système de télémétrie omnidirectionnel. (b) Triangulation active.

2.2.3 Notes sur des systèmes d'acquisition actifs

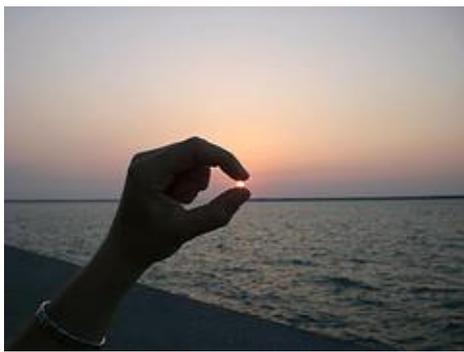


Figure 2.11 : Illusion d'optique due au manque d'information sur la profondeur.

L'acquisition par une caméra qu'elle soit à large champ de vue ou directionnelle, est fondamentalement une projection d'une scène 3D sur un plan 2D. Cela implique implicitement un écrasement de l'information sur la profondeur de la scène ²⁴(figure 2.11). L'approche active propose de restituer cette information en la réinjectant lors du processus d'acquisition. Cette opération se fait par émission-réception d'une onde radiofréquence, lumineuse ou ultrasonore. Dans le cas de la télémétrie un dispositif à part est chargé de sonder la profondeur, tandis qu'avec la

triangulation active, la camera est utilisée pour faire l'acquisition de la scène y compris l'onde lumineuse projetée.

Il est manifeste qu'un système d'acquisition actif n'est pas le meilleur choix lorsque la discrétion est requise par l'application visée. L'émission d'un signal quel qu'il soit peut s'avérer nuisible dans certaines situations.

Les techniques de télémétrie active se différencient essentiellement par leur portée, leur précision ainsi que leur sensibilité au bruit. Les télémètres à ultrasons sont relativement simples d'utilisation, peu coûteux. L'utilisation de la télémétrie à ultrasons est réservée généralement aux espaces clos, vides ou comme simple technique de détection d'obstacle, vu que la précision des mesures peut être

²⁴ Source de figure : [155]

perturbée par des interférences dus aux surfaces irrégulières des objets sondés [27], par le vent, les changements de températures, ou par la présence d'objets qui absorbent les ondes acoustiques.

Les radars et lidars à temps de vol ont une précision irréprochable sur les longues portées, mais ils sont moins adaptés aux applications opérant sur des portées plus courtes, puisque les longueurs d'onde utilisées ne le permettent pas. Les télémètres à décalage de phase sont généralement plus rapides est plus précis que ceux à temps de vol mais leur portée est réduite. Pour assurer une bonne acquisition le télémètre et le sujet à sonder doivent être stables, puisque chaque point de la scène est échantillonné à un instant différent, les vibrations à faible amplitude peuvent induire des distorsions dans les données collectées. Quelques solutions ont été proposées dans la littérature pour compenser les distorsions induites [28].

Les télémètres sont des dispositifs qui numérisent la scène environnante point par point, et se caractérisent par leur résolution, quantifiée en nombre de points numérisés par unité de surface, ainsi que par leur vitesse de balayage, mesurée en nombre de points par unité de temps. Plus la résolution est grande moins l'est la vitesse d'acquisition. Comme dans le cas des caméras, un compromis doit alors être fait entre les deux.

La triangulation active est une alternative qui permet de palier aux contraintes de la résolution limitée des techniques de télémétrie. Grâce à l'intégration de la camera comme capteur passif, cette technologie permet d'acquérir des quantités d'informations considérables de façon instantanée et avec moins de sensibilité aux variations d'éclairage.

La majorité des systèmes d'acquisition 3D actifs se partage une incommodité majeure, qui est la résolution azimutale restreinte. En vision directionnelle, cela ne représente pas un vrai handicap, mais en vision omnidirectionnelle, le champ de vue est précieux. Une multitude de solutions avaient été proposées dans la littérature et en industrie pour remédier à ce problème. Les Lidars à titre d'exemple utilisent des systèmes de balayage à miroirs rotatifs ; le rayon laser émis par la source est redirigé par des miroirs dans le but d'effectuer un balayage bidimensionnel de la scène. Un seul miroir à deux degrés de liberté est utilisé dans le cas des lidars à faible vitesse de balayage, alors que deux miroirs sont nécessaires lorsque la rapidité est envisagée. La façon la plus courante pour mouvoir les miroirs est l'utilisation d'un galvanomètre résonant ou d'un servomoteur (figure 2.12.a).

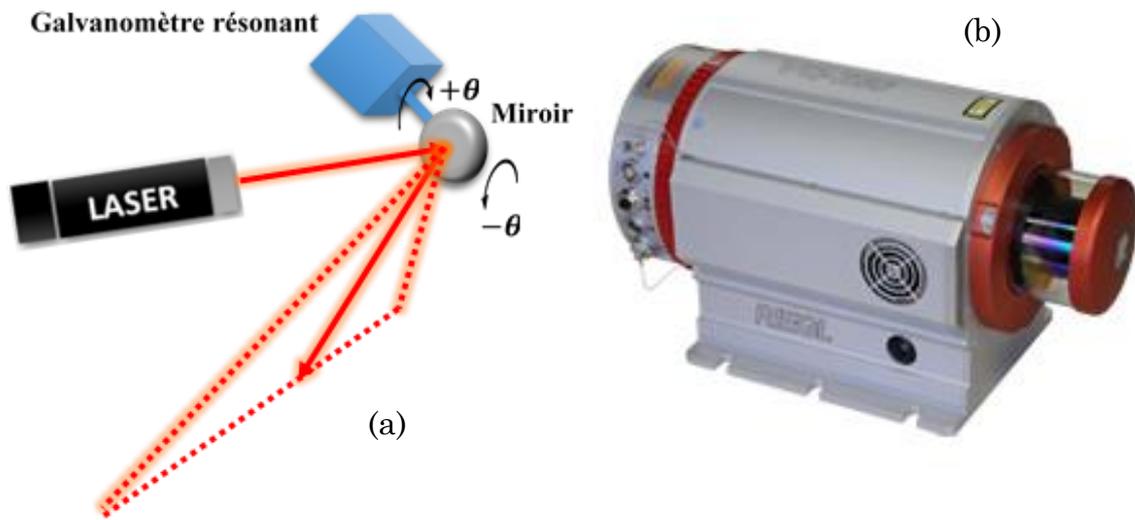


Figure 2.12 : (a) Système de balayage à miroir pour télémètre laser. (b) RIEGL VQ-450 scanner laser à large champ de vision.

Le progrès technologique qu'a connu le domaine de l'électronique a permis de concevoir des systèmes de balayage avec des champs de vision quasi-omnidirectionnels, et des vitesses de balayage assez importantes, des modèles commerciaux tels que le *RIEGL VQ-450*²⁵ (figure 2.12.b) ou le *HDS7000* ont atteint une vitesse d'un million de points échantillonnés par seconde, avec un champ de vision de $360^\circ \times 320^\circ$ dans le cas du *HDS7000* [29]. Cependant ces systèmes d'acquisition à balayage requièrent des post-traitements assez lourds sur le flux volumineux des données générées, le *HDS7000* comme exemple nécessite 13 secondes pour effectuer la numérisation de la scène environnante à faible résolution, et 6:44 min à haute résolution. On peut en déduire que de tels dispositifs ne sont pas compatibles avec des applications temps réel.

Ces mêmes contraintes sont rencontrées avec les systèmes à triangulation active ; le champ de vision des caméras peut être étendu aisément par les techniques de la vision omnidirectionnelle, mais celui des projecteurs de lumière structuré demeure moins extensible. Les systèmes à balayage ont une vitesse d'acquisition assez réduite pour qu'ils soient adoptés par les applications qui nécessitent des traitements en temps réel. L'évaluation des performances temporelles du prototype proposé dans [23] et illustré dans la figure 14, indique que l'acquisition d'une image panoramique monochromatique dure environ trois minutes, celle d'une image panoramique multi spectrale prend 12 minutes, alors que l'acquisition d'une scène entière peut durer plus qu'une heure. Vu toutes ces contraintes, l'acquisition 3D omnidirectionnelle par des capteurs passifs peut être considérée comme une alternative rivale à l'approche active, en fournissant une quantité d'informations appréciables de façon quasi instantanée.

²⁵ Source de figure : [161]

2.3 Les systèmes d'acquisition 3D passifs

2.3.1 Les techniques *shape-from-X*

Les systèmes d'acquisition tridimensionnelle passifs sont fondamentalement des systèmes de vision artificielle qui permettent une perception en relief de la scène sans émission d'onde. Ils utilisent des caméras à base de capteurs CCD ou CMOS pour resituer l'information 3D à partir des images acquises.

La vision 3D passive est l'une des disciplines les plus vieilles de la vision par ordinateur, une multitude de techniques connues sous le nom *shape-from-X* ont été développées afin d'inférer les formes des objets à partir des images. Selon le nombre d'images d'entrée, nous pouvons classer ces techniques en deux catégories : les méthodes basées sur deux ou plusieurs images et les méthodes basées sur une seule image. Dans le premier cas, les deux ou plusieurs images d'entrée peuvent être prises soit par multiples caméras fixes situées à différents angles de vue, ou par une seule caméra avec des objets en mouvement dans les scènes. Nous appelons les indices de profondeur utilisés par la première catégorie les indices de profondeur multioculaires. La deuxième catégorie d'indices de profondeur basés sur une seule image fixe, sont des indices de profondeur monoculaires. Le tableau 2-1 résume les techniques majeures d'acquisition 3D passives.

Tableau 2-1 : Les techniques passives d'acquisition 3D.

| Catégorie | Technique d'acquisition 3D | Principe |
|------------------------------|---------------------------------|--|
| Binoculaire ou Multioculaire | <i>Shape-from-motion</i> | Flux optique, factorisation, filtre de Kalman. |
| | <i>Shape-from-focus/defocus</i> | Un ensemble d'images à différents niveaux de mise au point et de l'estimation de la netteté, Décomposition locale de l'image en utilisant la base des polynômes d'Hermite, filtrage inverse, Transformée en S. |
| | <i>Shape-from-silhouette</i> | Modèle de maillage voxel déformable. |
| | <i>Shape-from-stereo</i> | Corrélation, Mise en correspondance, Triangulation. |
| Monoculaire | <i>Shape-from-texture</i> | Texture frontale. |
| | <i>Shape-from-Scattering</i> | Modèle de diffusion de la lumière. |
| | <i>Shape-from-shading</i> | Minimisation d'énergie. |

La technique *shape-from-motion* [30] désigne le processus d'estimation des structures tridimensionnelles à partir des séquences d'images à deux dimensions qui peuvent être couplés avec des signaux de mouvement locaux. Les humains perçoivent beaucoup d'informations sur la structure en trois dimensions de leur environnement en se déplaçant à travers elle. Lorsque l'observateur se déplace dans un milieu statique, l'information est obtenue à partir d'images captées au fil du temps. En biologie, cette habilité est connue sous le nom *Structure from motion* (SfM). Les algorithmes SfM sont susceptibles de fournir des estimations robustes de mouvement et de la forme, lorsqu'ils sont étendus aux cameras omnidirectionnelles, grâce à leur large champ de vision. Christopher Geyer et Kostas Daniilidis ont publié en 2001 un algorithme [31] de type *structure-from-motion* à partir de mise en correspondance des points dans les images acquises par les caméras catadioptriques non calibrées avec des miroirs paraboliques. Les auteurs supposent que trois paramètres intrinsèques sont à déterminer : la distance focale combinée du système catadioptrique miroir-lentille, et l'intersection de l'axe optique avec l'image. Une nouvelle représentation des images de points et de lignes dans les images catadioptriques a été introduite sous le nom de l'espace du cercle. Cet espace comprend des cercles virtuels, chacun d'entre eux est l'image de la conique absolue. La contrainte épipolaire a été formulée dans l'espace du cercle, et une matrice catadioptrique fondamentale de taille 4x4 avait été établit. Les auteurs ont démontré que l'image de la conique absolue appartient au noyau de cette matrice. Cela permet la reconstruction euclidienne à partir de deux prises de vues avec des paramètres constants, et de trois prises de vue avec des paramètres variables. Dans les deux cas, il y a toujours une prise de vue de moins, que celles nécessaires avec les caméras perspectives.

Des chercheurs de l'université *Carnegie Mellon* ont présenté dans [32] une extension de deux algorithmes SfM aux cameras catadioptriques non centrales. Le premier algorithme est un procédé de traitement par lots qui détermine la forme et estime le mouvement en exploitant toutes les observations de l'image simultanément. Généralement les algorithmes SfM se déroulent en deux étapes : l'extraction des attributs tels que les points d'intérêt, et leur suivi à travers les séquences d'images. Dans ce cas l'algorithme *Lucas-Kanade* [33] qui est une méthode itérative à résolution sous-pixellique a été adopté. Les caractéristiques résultantes sont ensuite utilisées pour estimer le mouvement de la caméra et les coordonnées tridimensionnelles des points suivis. La méthode *Levenberg-Marquard* utilisée permet de minimiser l'erreur entre les positions des points suivis et ceux prédits. La deuxième méthode présentée dans ce travail utilise un filtre de *Kalman* étendu pour estimer la forme, le mouvement et leurs covariances. Cette méthode présente l'avantage de pouvoir traiter chaque image dès qu'elle est disponible, et n'exige pas que toutes les images soient disponibles a priori. Ce qui la rend plus adaptée pour une utilisation sur des séquences d'images infinies (temps réel). Les deux méthodes proposées ont été testées sur des plateformes robotiques et ont prouvées leurs performances.

Les méthodes *shape-from-focus/defocus* génèrent une carte de profondeur à partir du degré de flou présent dans les images. Dans un système à lentille mince, les objets qui sont mis au point sont clairement représentés tandis que les objets à d'autres distances sont flous. La figure 2.13 [34] illustre un modèle de lentille mince d'un réel point objet hors mise au point et sa projection sur le plan de l'image.

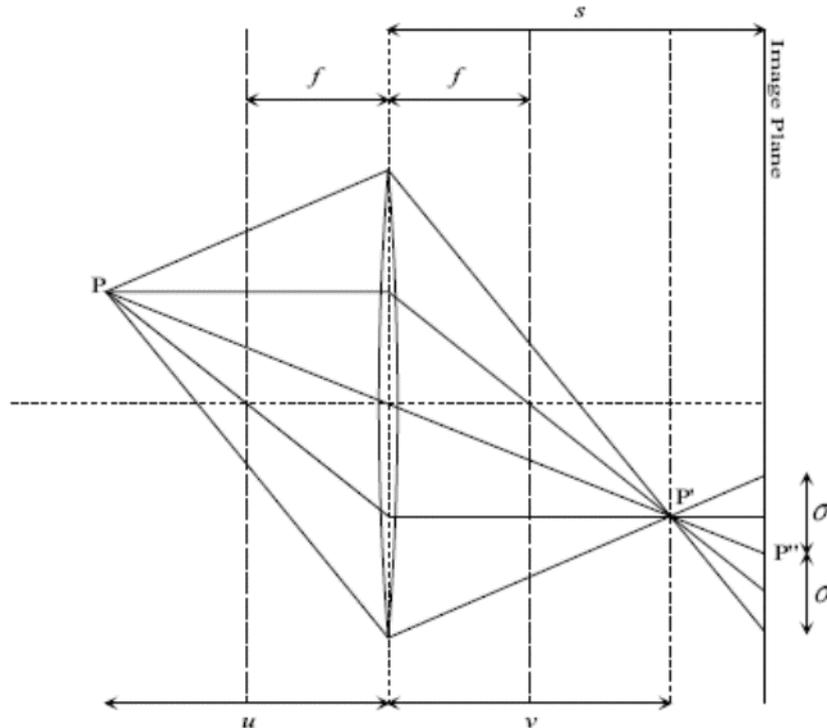


Figure 2.13 : Modèle de lentille mince.

Cette projection correspondante est une tâche de flou circulaire avec une luminosité constante, centrée autour de P'' avec un flou de rayon σ . Le flou est causée par la convolution de l'image idéale projetée par la fonction d'étalement ponctuel de la caméra (*point spread function, PSF*) $g(x, y, \sigma(x, y))$, où (x, y) sont les coordonnées du point image à P'' . On admet généralement que $\sigma(x, y) = \sigma$, ou σ est une constante pour une fenêtre donnée dans le but de simplifier ce modèle. Une fonction gaussienne est utilisée pour simuler la fonction PSF :

$$g_{\sigma}(x, y) = \frac{1}{\pi\sigma^2} e^{-\frac{x^2+y^2}{\sigma^2}} \quad 2.6$$

Afin d'estimer la profondeur u , nous avons besoin des deux équations suivantes ; L'équation fondamentale des lentilles minces qui décrit la relation entre u , v et f comme suit :

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f} \quad 2.7$$

Pentland dans [35] a déduit une relation entre la distance u (figure 2.13) et le flou σ dans l'équation suivante :

$$u = \begin{cases} \frac{fs}{s-f-kf\sigma} & \text{if } u > v \\ \frac{fs}{s-f+kf\sigma} & \text{if } u < v \end{cases} \quad 2.8$$

Où u est la profondeur, v est la distance entre la lentille et la position de la mise au point parfaite, s est la distance entre la lentille et le plan d'image, f est la distance focale de la lentille, et k est une constante déterminée par le système de lentilles. Lorsque les paramètres de la caméra sont obtenus après calibration, la profondeur u peut être calculée à partir de l'équation (2.8) une fois que le paramètre de flou σ est connu.

La silhouette d'un objet dans une image correspond au contour séparant l'objet de l'arrière-plan. Les méthodes **Shape-from-silhouette** nécessitent de multiples prises de vue de la scène par des caméras à différentes positions. Un tel processus peut générer un modèle complet en 3D des objets de la scène. Shape-from-silhouette nécessite un étalonnage précis de la caméra. Pour chaque image acquise, la silhouette des objets cible est segmentée en utilisant la soustraction du fond. Les silhouettes sont ensuite extraites à partir des images binaires, et sont de retour projetés sur un espace 3D commun avec des centres de projection correspondant à l'emplacement de la caméra. Un volume de forme conique résulte de la projection inverse de chaque silhouette. L'intersection de tous les cônes formant l'enveloppe visuelle de l'objet 3D cible, qui est souvent traité par représentation en voxel.

Les textures à motif offrent une bonne empreinte 3D grâce à deux éléments essentiels : la distorsion des Texels individuels et le taux de variation de distorsion dans la région de texture. Ce dernier est également connu comme gradient de texture. La reconstruction de la forme²⁶(figure 2.14) exploite les distorsions telles que la distorsion perspective, qui fait apparaître les Texels distants de la caméra plus petits, et la distorsion de raccourcissement, qui rend les Texels qui ne sont pas parallèles au plan de l'image plus courts.

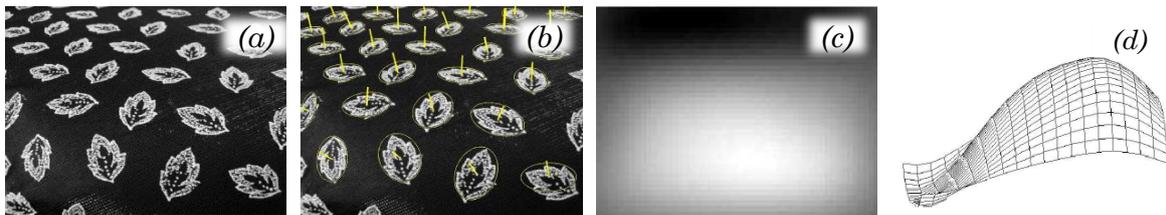


Figure 2.14 : Reconstruction de la forme à partir de la texture. (a) Image réelle : tissu à texture anisotrope non stationnaire (b) Diagramme à aiguille des Texels détectés (c) La profondeur de la surface illustrée en niveaux de gris (d) Maillage de la surface vue de profil.

En général, le résultat des algorithmes **shape-from-texture** (SfT) est une carte de densité des normales à la surface, ce qui permet de récupérer la forme 3D sous l'hypothèse d'une surface de texture lisse. En règle générale, la forme d'une surface

²⁶ Source de figure: [156]

en un point quelconque est complètement spécifiable par l'orientation et la courbure de la surface. Puisque l'estimation de la courbure s'avère être compliquée, les algorithmes *shape-from-texture* se concentrent sur la détermination des orientations de la surface texturée en termes des normales à la surface. Il est à noter que beaucoup d'images réelles contiennent des régions de texture différentes ou des zones texturées entourées d'autres zones non texturées. Ces différentes régions texturées doivent être segmentées avant que la plupart des algorithmes *shape-from-texture* ne puissent être appliqués. La figure 2.14 représente un processus typique de reconstruction d'une forme texturée. Cette classe de méthodes, qui nécessitent la segmentation des textures, est liée à l'approche basée sur les caractéristiques. Au cours des dernières années, il y a eu une évolution vers des méthodes *shape-from-texture* qui utilisent l'information spectrale et évitent la détection a priori des caractéristiques. Ces méthodes comparent la représentation spectrale des correctifs de l'image fenêtrée pour récupérer l'orientation. Les représentations spectrales couramment utilisées sont la transformée de Fourier, la décomposition en ondelettes et la transformée de Gabor.

La propagation de la lumière à travers l'atmosphère est affectée en termes de direction et d'intensité à cause de la diffusion du rayonnement par les petites particules de l'atmosphère. Cela conduit à un phénomène appelé la diffusion atmosphérique (*atmosphere scattering*), ce qui provoque divers effets visuels : des objets lointains apparaissent moins nets et plus bleutés que les objets proches, un faisceau de lumière d'une source intense est diffusé dans un environnement brumeux. Bien que la diffusion atmosphérique est l'un des sujets classiques de la physique, en psychologie elle est souvent considérée comme l'un des principaux indices de perception de la profondeur par le système visuel humain, peu de travaux dans le domaine de la vision par ordinateur ont pu établir un lien direct entre la diffusion atmosphérique et l'information sur la profondeur. Cozman et Krotkov [36] ont été les premiers à présenter une analyse de cette liaison en 1997. Leurs travaux se basaient sur le modèle de diffusion physique de Lord Rayleigh datant de 1871. L'algorithme est adapté pour estimer la profondeur des images extérieures contenant une partie du ciel. Après simplification du modèle physique complexe, la relation suivante est tirée entre la luminance d'une image et la distance entre l'objet et l'observateur :

$$\check{C} = C_0 e^{-\beta z} + S(1 - e^{-\beta z}) \quad 2.9$$

Où \check{C} est l'intensité mesurée d'un objet, alors que C_0 est l'intensité du même objet mais en absence de diffusion. β est le coefficient d'extinction, z est la profondeur de l'objet, et S est l'intensité du ciel. L'équation (2.9) comprend deux parties associées par addition. La première partie décrit le fait que l'intensité lumineuse est atténuée quand un faisceau lumineux est projeté sur un objet à travers un milieu diffusant. La deuxième partie reflète le phénomène inverse: un gain réel en intensité résulte de la diffusion. Dans la majorité des cas β et C_0 ne sont inconnus, alors que S peut être mesurée à partir de toutes les images qui contiennent une région du ciel. Pour les scènes d'intérieur, l'estimation de S implique la génération

artérielle de la vapeur d'eau, une tâche qui ne peut pas être automatisée. C'est l'une des limites de cet algorithme. Les résultats de l'algorithme sont des rapports de profondeur entre les différents objets de la scène, à partir desquels une carte de profondeur peut être générée.

La technique ***shape-from-shading*** (*SFS*) introduite par Horn [37] dans les années 70, permet elle aussi d'estimer la forme tridimensionnelle d'un objet uniforme, sous éclairage contrôlé, par l'étude des variations graduelles de son ombre à partir d'une seule image ²⁷(figure 2.15).

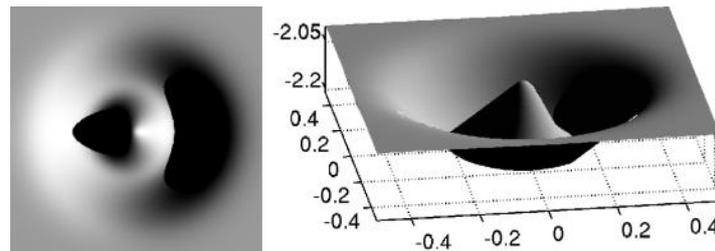


Figure 2.15 : Exemple de reconstruction à partir d'une image avec des ombres noires.

Les approches ***Shape-from-focus/defocus***, ***shape-from-texture***, ***shape-from-silhouette*** et ***shape-from-scattering*** sont peu envisageables dans le cas des systèmes de vision à large champ, vu les contraintes que présente leur implémentation, et les caractéristiques spéciales de ces systèmes de vision telles que les distorsions. Tandis que la technique ***shape-from-shading*** s'avère être utile en endoscopie ; une méthode d'exploration et d'imagerie médicale qui permet de visualiser l'intérieur des cavités du corps humain inaccessibles à l'œil. Un endoscope comporte généralement un objectif de distance focale courte à large champ de vue ($\sim 140^\circ$). Le problème majeur rencontré lors de l'extension des techniques ***shape-from-shading*** à ce genre de capteur est les distorsions radiales fortes qu'incorporent les images acquises. Deux approches sont alors envisageables : la première consiste à corriger les distorsions comme une étape préliminaire au traitement, la deuxième repose sur l'intégration du modèle des distorsions dans les équations de reconstruction 3D, ce qui permet d'éviter l'implication d'une étape d'interpolation et par conséquent de préserver les informations de l'image d'origine. Mais un second problème est souvent rencontré lors du traitement des images à large champ de vue ; la résolution de l'image n'est pas uniforme, plus on s'éloigne du centre de l'image, plus la résolution décroît. La compensation de cette déperdition de résolution doit être prise en considération lors de l'élaboration du modèle de reconstruction. La ²⁸figure 2.16 illustre un exemple de reconstruction d'une boule de calibration acquise par un endoscope à source d'éclairage intégrée.

²⁷ Source de figure: [157]

²⁸ Source de figure: [158]

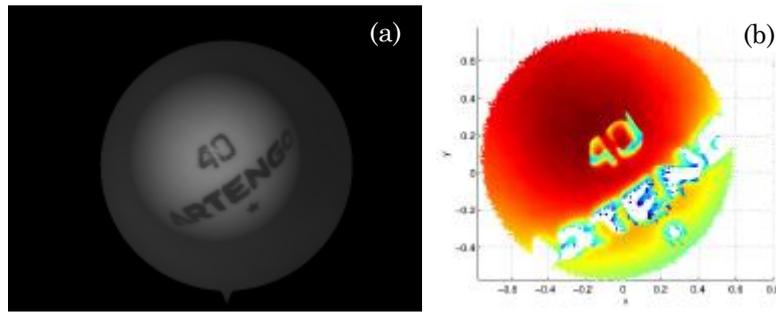


Figure 2.16 : Reconstruction 3D d'une boule (ping-pong) par technique shape-from-shading.

Avec deux images de la même scène prises à deux positions légèrement distantes, la disparité binoculaire peut être utilisée pour récupérer la profondeur d'un objet. Dans la nature, c'est le principal mécanisme de la perception de la profondeur. En vision par ordinateur cette technique est connue sous le nom *shape-from-stereo* ou stéréovision. Initialement un ensemble de points mis en correspondance dans la paire d'images doit être désigné. Ensuite, l'information sur la profondeur peut être récupérée par triangulation (figure 2.17) avec un bon degré de précision, si les paramètres intrinsèques de la camera sont bien déterminés.

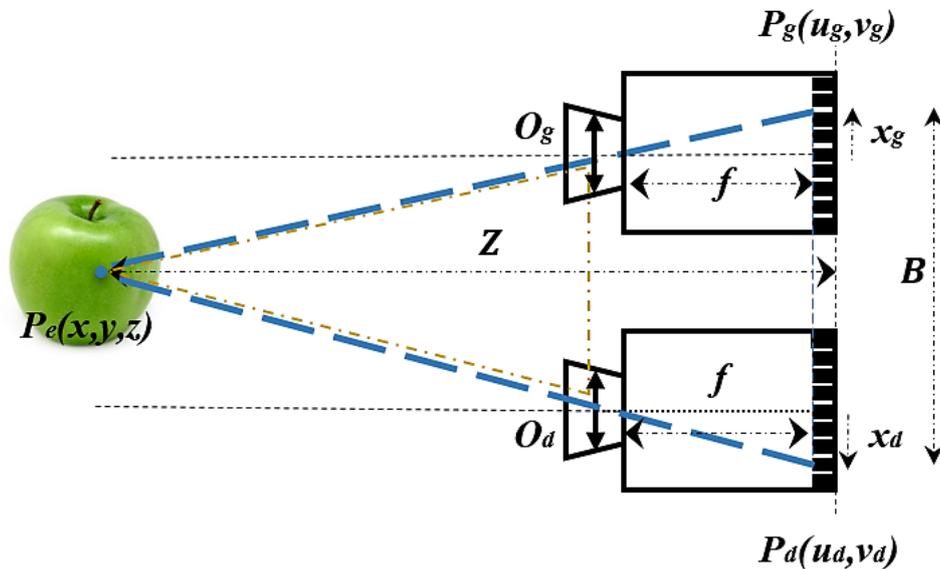


Figure 2.17 : Principe de la stéréovision.

Si on considère P_g et P_d comme les projections du point P_e de l'espace 3D sur les deux plans des images droite et gauche, et on se basant sur la relation entre les deux triangles semblables (P_e, O_d, O_g) et (P_e, P_d, P) l'expression de la profondeur Z est donnée par :

$$Z = f \frac{B}{d} \quad 2.10$$

Où la disparité d qui mesure la différence de position pixellique entre les deux points P_g et P_d , est définie par : $d = x_g - x_d$.

L'aspect le plus fastidieux des algorithmes d'estimation de la profondeur basés sur la disparité binoculaire est le problème de mise en correspondance stéréoscopique, l'un des domaines de recherche les plus actifs de la vision par ordinateur. Étant donné un point image sur l'image de gauche, comment peut-on trouver le point image correspondant dans l'image droite ? En raison des ambiguïtés inhérentes des paires d'images stéréoscopiques, telles que l'occlusion, le problème générale de mise en correspondance stéréo est difficile à résoudre. Plusieurs contraintes ont été introduites pour simplifier la résolution de ce problème. La géométrie épipolaire et l'étalonnage de la camera sont les deux contraintes les plus fréquemment utilisées, avec lesquelles, les paires d'images stéréoscopiques peuvent être rectifiées.

La géométrie épipolaire [38] [39] est utilisée en stéréoscopie afin de contraindre l'espace de recherche des points homologues (figure 2.18), qui sont les projections d'un même point P de l'espace 3D sur les différentes images. Le plan épipolaire est défini par le point P observé et les deux centres de projection O_d et O_g .

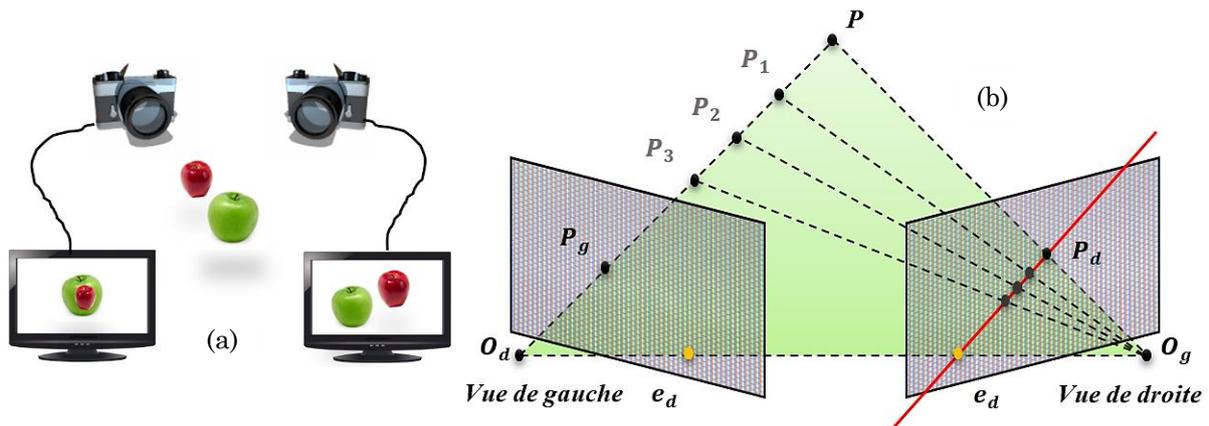


Figure 2.18 : La géométrie épipolaire. (a) : Une scène observée sous différents points de vue, la géométrie épipolaire doit décrire la relation entre les deux images résultantes. (b) Schématisation de la géométrie épipolaire.

Dans le cas des caméras directionnelles, les images sont d'abord rectifiées pour éliminer toute distorsion optique, ainsi une droite de l'espace est projetée en une droite dans l'image acquise. La recherche du correspondant d'un point de l'image de droite s'effectue donc le long d'une droite, dite épipolaire, dans l'image de gauche. Ces droites se croisent en un point, appelé épipole, qui est l'image du centre de projection.

L'étalonnage de la camera consiste à estimer les paramètres du modèle de la camera à partir d'observations ce qui revient à établir la relation entre les

coordonnées spatiales d'un point de l'espace et le point associé dans l'image acquise par la caméra.

Une autre hypothèse largement utilisée, est la contrainte photométrique, qui stipule que les intensités des pixels correspondants sont semblables les unes aux autres. La contrainte de l'ordre précise que l'ordre des points dans la paire d'images stéréoscopiques est généralement le même. La contrainte d'unicité affirme que chaque point ne peut avoir qu'un unique correspondant, Tant qu'à la contrainte de lissage aussi connu comme la contrainte de continuité, dit que la disparité change progressivement presque partout. Certaines de ces contraintes sont difficiles à implémenter, comme la géométrie épipolaire par exemple, tandis que d'autres telles que les contraintes de lissage sont moins complexes. Scharstein et Szeliski [40] ont étudié la performance d'environ 40 algorithmes de mise en correspondance fonctionnant sur une paire d'images rectifiées. Certains de ces algorithmes imposent diverses séries de contraintes.

Les algorithmes actuels de mise en correspondance stéréo sont basés sur la corrélation des fenêtres locales, par appariement d'un ensemble de caractéristiques de l'image telles que les points d'intérêt, ou par optimisation globale. Lors de la comparaison de corrélation entre les fenêtres des deux images, l'élément correspondant est donné par la fenêtre lorsque la corrélation est maximale. Une mesure de similarité traditionnelle est la somme des différences au carré (SSD). Les algorithmes locaux génèrent une carte de disparité, tant dit que les méthodes basées sur les caractéristiques sont conceptuellement similaires à celles à base de corrélation, mais à leur différence, la recherche de correspondance se fait par ensemble de caractéristiques. La mesure de similarité dans ce cas doit être adaptée au type des caractéristiques utilisées. De nos jours les méthodes d'optimisation globale sont plus populaires en raison de leur bonne performance. Ils font usage explicite des contraintes de lissage et essaient de trouver une affectation de disparité qui minimise une fonction globale d'énergie. L'énergie globale est généralement une combinaison du calcul de la similarité et le terme de continuité, où celui-ci mesure habituellement les différences entre les disparités des pixels voisins. C'est l'étape de minimisation de la différence utilisée dans ces algorithmes qui les différencie les uns des autres.

2.3.2 Contribution à la stéréovision omnidirectionnelle

La stéréovision omnidirectionnelle consiste à utiliser au moins deux capteurs à large champ de vision pour acquérir l'image de l'environnement en relief. Parmi les systèmes de vision omnidirectionnelle, les capteurs catadioptriques, qui présentent des spécificités particulièrement intéressantes pour la reconstruction 3D et la navigation. Grâce à leur champ de vue omnidirectionnel la fiabilité du suivi des objets ou des points d'intérêts le long d'une séquence d'images est assurée, un privilège qui ne peut être garanti dans le cas d'une caméra classique avec un champ de vision restreint.

Les premiers travaux dans ce domaine ont été initiés par Shree K. Nayar en 1989. Dans son article [41] *Sphero: Determining Depth using Two Specular Spheres and a Single Camera*, Nayar avait présenté un système de vision omnidirectionnelle catadioptrique, constitué de deux miroirs sphériques ainsi qu'une caméra directionnel fixe. Ce système permet l'estimation de la profondeur à partir de chaque paire de points correspondants par triangulation (figure 2.19). L'utilisation d'une seule caméra fixe permet d'éviter les complexités indésirables qui caractérisent la procédure d'étalonnage stéréo. La plage de mesure du système est considérablement améliorée par l'utilisation des miroirs sphériques et n'est pas limitée par le champ de vision de la caméra.

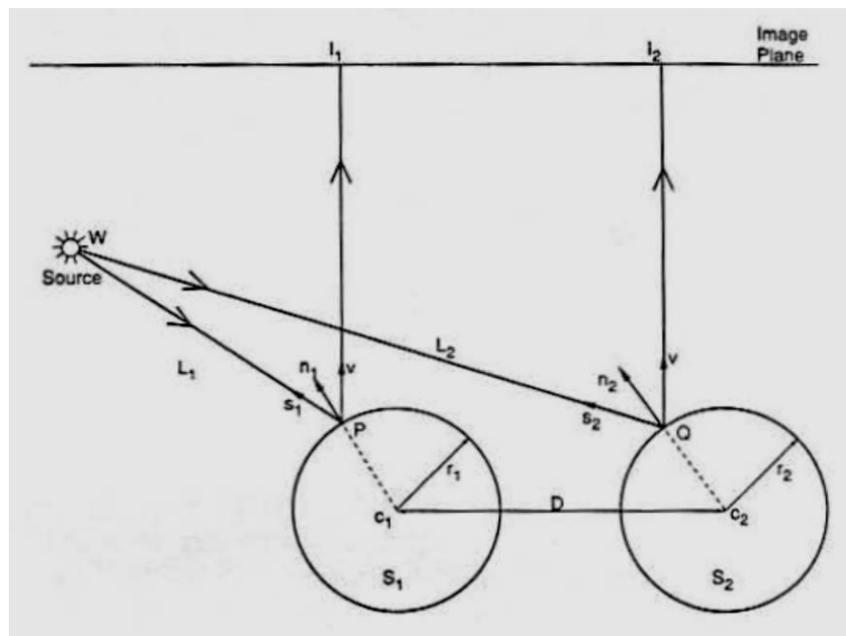


Figure 2.19 : le système de stéréovision omnidirectionnelle SPHEREO.

Depuis, la stéréovision omnidirectionnelle s'est vite répandue comme un axe de recherche en vision artificielle, riche et prometteur. Beaucoup de chercheurs ont trouvé de l'intérêt à l'exploiter dans diverses applications.

En vidéosurveillance, le coût, la complexité et la fiabilité sont trois critères à considérer lors de la conception d'un système. L'utilisation des capteurs omnidirectionnels permet à la fois de réduire le coût, vu qu'un seul capteur peut couvrir un large champ de vision qui nécessite normalement l'utilisation de plusieurs cameras classiques, et permet aussi de simplifier la complexité en évitant l'utilisation de dispositifs secondaires tels que les systèmes de motorisation des caméras. Dans [42] un système de vidéosurveillance N-Oculaire stéréo a été proposé. Le système est constitué d'un réseau de capteurs de vision omnidirectionnelle (Figure 2.20.a). L'une de ses fonctionnalités est le suivi en temps réel des personnes mobiles (figure 2.20.c). La segmentation des cibles dynamiques se fait par extraction du fond. La configuration N-Oculaire permet de

déterminer la position de la cible tout en minimisant l'erreur d'observation ; L'emplacement de la cible est mesuré par triangulation à partir des angles d'azimut qui sont détectés par chaque paire combinatoire possible des capteurs de vision arbitraires (figure 2.20.b). L'algorithme élimine par la suite toutes les positions calculées, où la cible n'est pas visible par au moins trois capteurs.

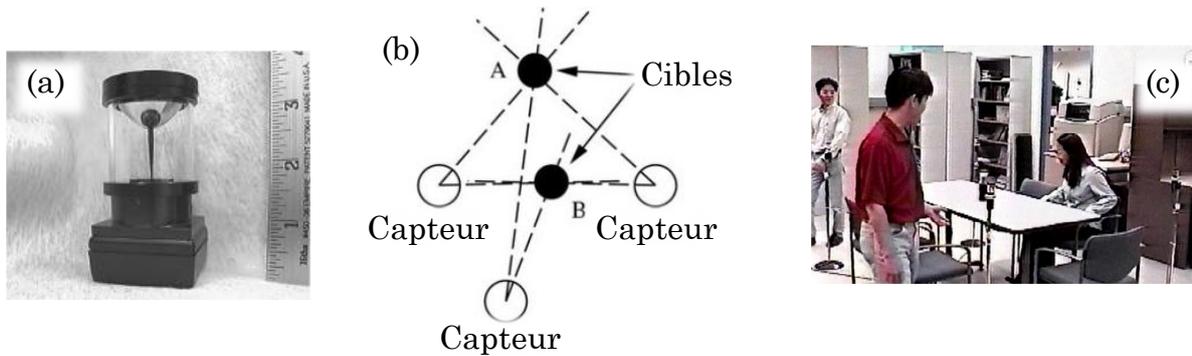


Figure 2.20 : Système de vidéosurveillance à base de capteurs omnidirectionnels. (a) capteur catadioptrique utilisé. (b) Principe de la stéréovision omnidirectionnelle N-Oculaire. (c) Aperçu du système.

Les systèmes *omnistéréo* se sont avérés utiles dans des applications d'aide à la conduite, et de la sécurité routière. Dans [43] un système d'assistance du conducteur basé sur la vision artificielle a été présenté. Il est constitué de deux cameras omnidirectionnelles catadioptriques montées sur un véhicule. Plusieurs configurations (figure 2.21) sont proposées pour la détection efficace des véhicules et des piétons. Dans ce travail une carte dense de profondeur utilisée généralement pour la détection d'obstacles est générée par mesure de corrélation, après rectification des images acquises. L'algorithme passe par cinq étapes principales, d'abord les images acquises sont rectifiées pour éliminer les distorsions, cette opération n'est possible que lorsque les caméras sont déjà calibrées. Ensuite une étape de prétraitement est exécutée où un filtre passe-bande est appliqué sur les images rectifiées par convolution avec un filtre laplacien de gaussienne. La mise en correspondance des pixels de l'image se fait par comparaison des fenêtres rectangulaires de chaque paire d'image stéréo, en utilisant la somme des différences absolues (SAD) comme mesure de corrélation. Chaque minimum des valeurs SAD correspond à un pic de corrélation. Le décalage spatial entre chaque paire de pixels mis en correspondance est connu entant que la disparité de ce pixel. La quatrième étape consiste à générer une carte de disparité par attribution de valeurs proportionnelles à chaque écart pixellique. Finalement les erreurs dues aux fausses mises en correspondance et les discontinuités de profondeur sont éliminées.

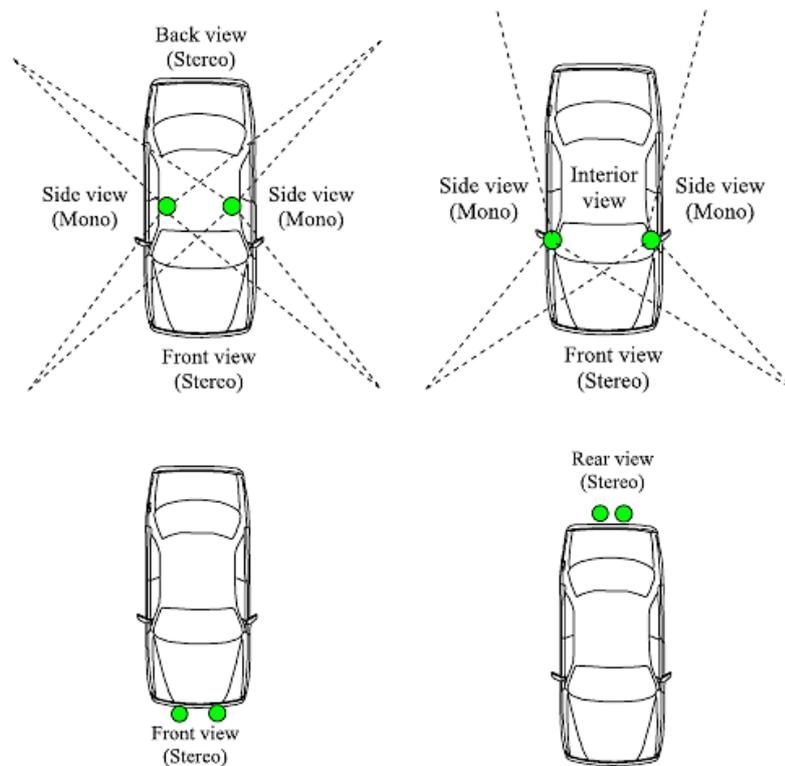


Figure 2.21 : Les différentes configurations d'un system omnistéréo monté sur un véhicule.

La réalité virtuelle qui est une simulation informatique interactive, immersive, et sensorielle d'environnements réels ou imaginaires, est l'un des domaines où les capteurs de vision omnistéréo ont été utilisés. Un système d'acquisition 3D omnidirectionnelle a été publié dans [44], Le capteur d'acquisition est constitué de deux miroirs pyramidaux hexagonaux et douze caméras CCD (figure 2.22.a).

L'approche présentée est basée sur l'acquisition des scènes dynamiques à l'aide ce capteur en temps réel. Un environnement mixte est créé à partir d'un modèle 3D

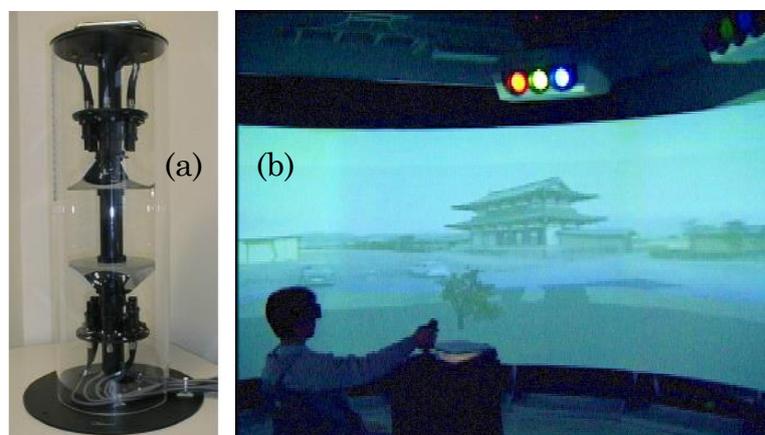


Figure 2.22 : (a) Capteur d'image stéréo omnidirectionnel. (b) Système de projection 3D immersive CYLINDRA.

cylindrique contenant les textures de la scène réelle, et d'un modèle virtuel qui comprend des images de synthèses ou des graphiques. Le modèle 3-D cylindrique est généré à partir des images panoramiques stéréo.

La profondeur d'une scène réelle existante est le principal facteur qui peut représenter correctement la relation entre la profondeur des objets virtuels et les objets réels. L'estimation de la profondeur se fait dans ce cas par mise en correspondance des images panoramiques stéréo. Cependant, il y a une forte possibilité de fausses correspondances causées par les bruits lors de l'exercice de l'appariement sur l'ensemble des images stéréo. Afin de réduire les correspondances incorrectes, la solution proposée consiste à effectuer la mise en correspondance des contours des objets et non pas l'ensemble des pixels de l'image, par la suite, les données intermédiaires sont estimées par interpolation linéaire. L'algorithme détecte d'abord les contours par application du filtre de Sobel sur l'image du haut et celle du bas. La correspondance stéréo est effectuée par mesure de la similarité corrélation croisée normalisée, avec une fenêtre de 9x9 pixels. Les bruits et les valeurs manquantes sont ensuite éliminés par application d'un filtre médian de 5x3. Le résultat est une carte de profondeur (figure 2.23) dans lequel des valeurs de profondeur sont codés en intensité ; un pixel brillant est plus proche et un pixel sombre est plus loin. Un pixel noir est un pixel dont la profondeur n'a pas pu être calculée à partir d'images stéréo. L'environnement mixte 3D généré est exposé à l'aide d'un système de projection immersive composé de six projecteurs RGB.

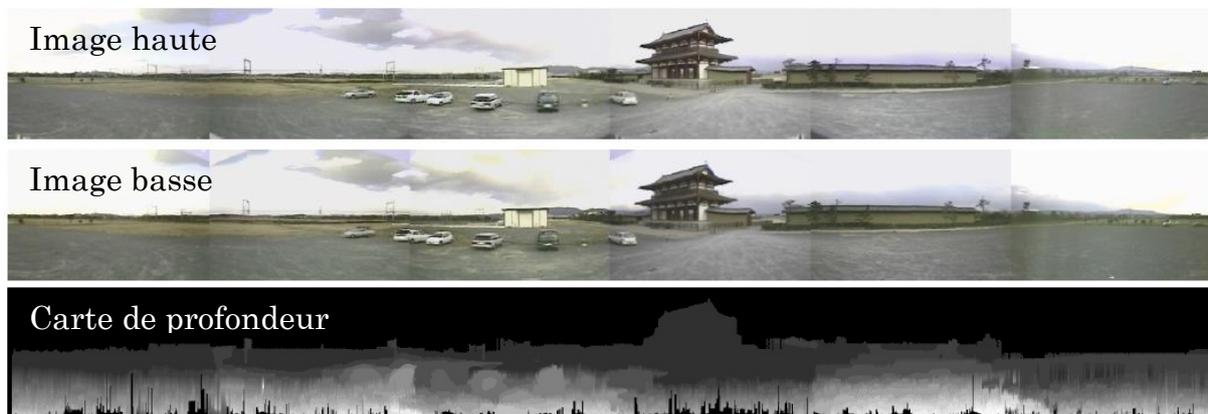


Figure 2.23 : Carte de profondeur panoramique générée à partir d'images stéréo.

La stéréovision omnidirectionnelle a été largement utilisée en robotique à diverses fins, telles que la détection d'obstacles [45], la cartographie [46], la localisation [47], et la navigation [48]. Les capteurs omnidirectionnels catadioptriques sont les capteurs à large champ de vision les plus utilisés en robotique. Rappelons que les systèmes omnidirectionnels catadioptriques procèdent de la combinaison des miroirs convexes et des caméras. L'alignement de l'axe optique de la caméra et l'axe de révolution du miroir permet de concevoir un système d'acquisition avec un champ de vision de 360°. Cette approche conserve les propriétés d'acquisition en temps-réel et permet d'obtenir des images de bonne résolution. Néanmoins une

adaptation des algorithmes des traitements classiques au modèle de projection non-linéaire et à la résolution non-uniforme des images catadioptriques est nécessaire. Notre étude sera focalisée par la suite sur ce type de capteur.

Les systèmes catadioptriques peuvent être classifiés en deux catégories : centraux et non centraux [49]. Un capteur est dit central, si le miroir convexe utilisé satisfait la contrainte du point de vue unique, sinon il est considéré non central.

La contrainte du point de vue unique SVP (*Single View Point*) est définie comme suit : à chaque point de l'image catadioptrique, est associée une seule et unique direction projective de l'espace. Les systèmes satisfaisants cette contrainte présentent certains avantages par rapport aux autres, parmi ces avantages : la possibilité d'éliminer les distorsions lors de la transformation de l'image omnidirectionnelle en image panoramique (déballage), ainsi que la simplicité relative à la modélisation et du calibrage de ce genre de cameras. Baker et Nayar [50] ont démontré qu'en se basant sur le modèle sténopé de la camera, seuls les quadriques de révolution de type ellipsoïdes, paraboloides, hyperboloïdes et planaires, satisfassent la contrainte SVP.

Chaque forme des miroirs convexes (figure 2.24) ; conique, sphérique, hyperbolique, ou parabolique, présente des spécificités propres, des avantages et des inconvénients concernant la satisfaction du point de vue unique, l'importance des aberrations, et la complexité de réalisation (cout). La conception d'un système stéréo débute d'abord par le choix du miroir de révolution et ensuite la configuration du système. Dans la partie qui suit nous allons décrire les différents miroirs existants ainsi que les diverses configurations possibles.

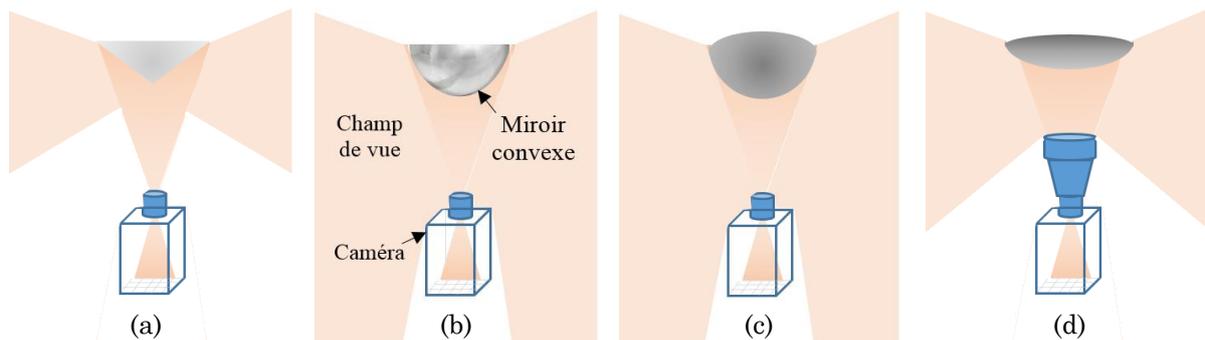


Figure 2.24 : Miroirs convexes pour caméras catadioptriques. (a) Miroir conique (b) Miroir sphérique (c) Miroir hyperbolique (d) Miroir parabolique.

2.3.2.1 Choix du miroir

Selon Baker et Nayar. [51], **les miroirs coniques** ont un point de vue unique au sommet du cône. Cela signifie que dans le but d'accomplir la contrainte SVP, le sténopé de la caméra doit être placé à l'extrémité du cône. Dans les cas particuliers où un modèle de scène 3D est disponible, l'image panoramique peut être mise en correspondance avec le modèle pour obtenir un environnement virtuel en 3D. Des chercheurs de l'Université de l'Alberta [52] ont utilisé une méthode similaire pour l'inspection de canalisation à l'aide d'une caméra catadioptrique avec un miroir conique. L'image omnidirectionnelle fournie par la caméra catadioptrique est mappée sur une surface 3D définie par un modèle précédemment disponible de la surface réelle ou sur une surface développable complexe formée par des blocs de construction de base. Même si le miroir conique a été traditionnellement mis au rebut de l'ensemble des miroirs qui sont conformes à la contrainte SVP, placer les configurations de miroirs coniques dans la catégorie des capteurs catadioptriques non-SVP est devenu contestable. En 2001, Lin et Bajcsy [53] ont montré que le miroir conique peut être utilisé dans un capteur catadioptrique SVP si le modèle de la caméra n'est pas le sténopé, mais un étendu qui prend en compte des phénomènes optiques plus complexes. Lin et Bajcsy [54] ont décrit un capteur

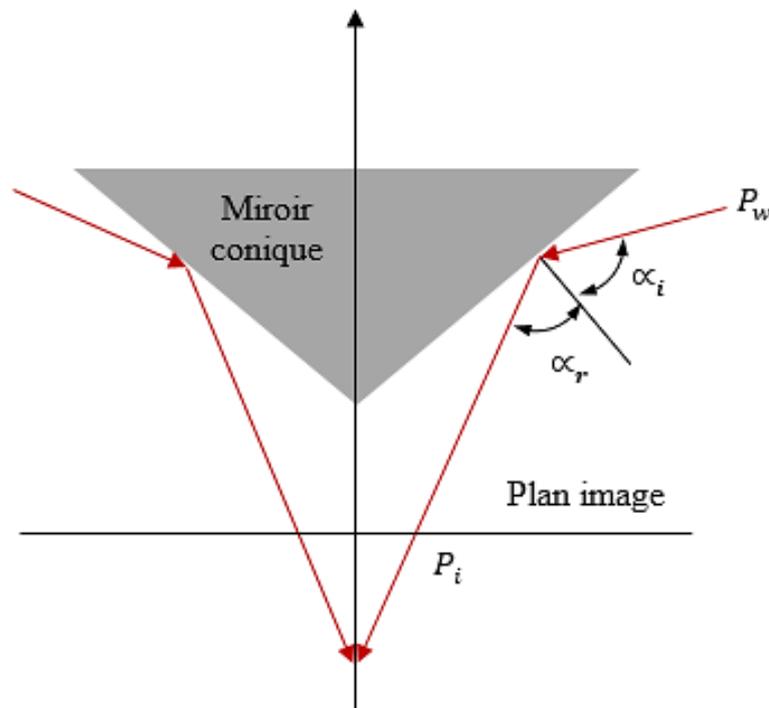


Figure 2.25 : Géométrie de formation de l'image acquise par caméra catadioptrique à miroir conique.

omnidirectionnel qui fournit des informations de profondeur à l'aide de deux caméras, un diviseur de faisceaux et un miroir conique, pour des applications en robotique. Un capteur similaire a été proposé par Spacek [55] qui a utilisé une

configuration stéréo omnidirectionnelle intégrée avec deux miroirs coniques coaxiaux.

☑ **Attributs :**

- La fabrication des miroirs coniques est relativement facile, comme ils sont peu coûteux.
- La courbure nulle de ce type de miroir, annule l'effet *fish-eye* qui se manifeste par un grossissement des objets auprès du centre.
- La direction de la normale de réflexion est constante, ce qui simplifie l'association du miroir conique à d'autres quadriques de révolution [56].

☒ **Incommodités :**

- L'astigmatisme est important.
- Les cameras catadioptriques à miroir conique ne sont pas centrales.

Les miroirs sphériques respectent la contrainte SVP seulement dans le cas où une caméra de modèle sténopé est placée au centre de la sphère. La camera dans ce cas n'acquerra que sa propre image. Donc, ce genre de miroirs n'est pas utile pour les systèmes qui exigent le respect de la contrainte du point de vue unique. Néanmoins, parfois, les miroirs de forme sphérique représentent un choix attrayant, en raison de leurs caractéristiques telles que le faible coût de fabrication, les caractéristiques spécifiques de la densité de résolution de l'image qui est en fonction de l'angle d'élévation, ainsi que leurs tailles.

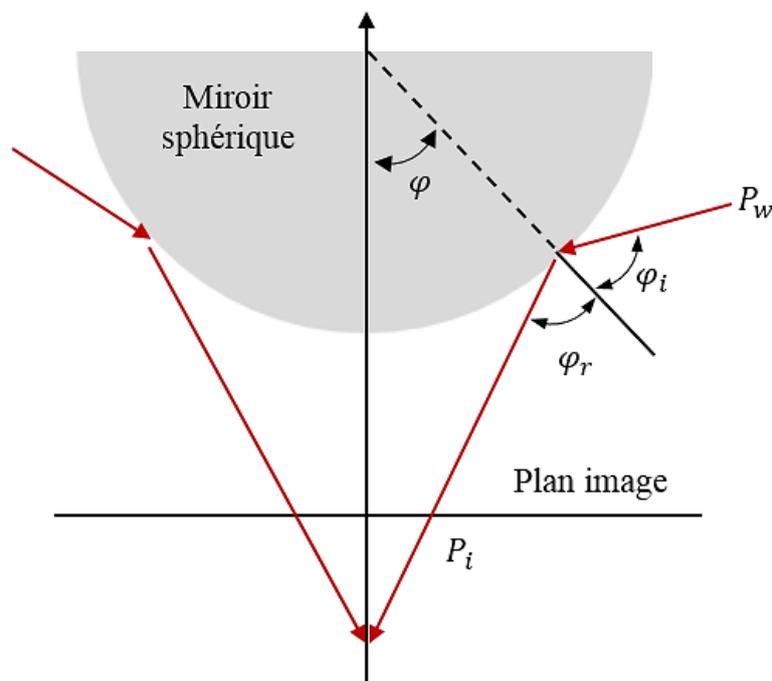


Figure 2.26 : Géométrie de formation de l'image acquise par caméra catadioptrique à miroir sphérique.

En 1996 Southwell et al., [57] ont utilisé un capteur catadioptrique non central muni d'un miroir sphérique à double lobe pour obtenir deux images séparées, ce qui permet une perception 3D omnidirectionnelle. Plus tard, Fiala et Basu [58] ont utilisé un capteur similaire à obtenir une vue panoramique stéréo dans laquelle les lignes horizontales et verticales ont été détectées à l'aide de la transformée de Hough panoramique, une extension de la transformée de Hough adaptée aux images catadioptriques. La profondeur est récupérée en faisant correspondre les lignes imagées par les deux miroirs sphériques à travers la symétrie radiale. Ils ont utilisé le miroir sphérique à deux lobes, indiqué sur la figure 2.27.

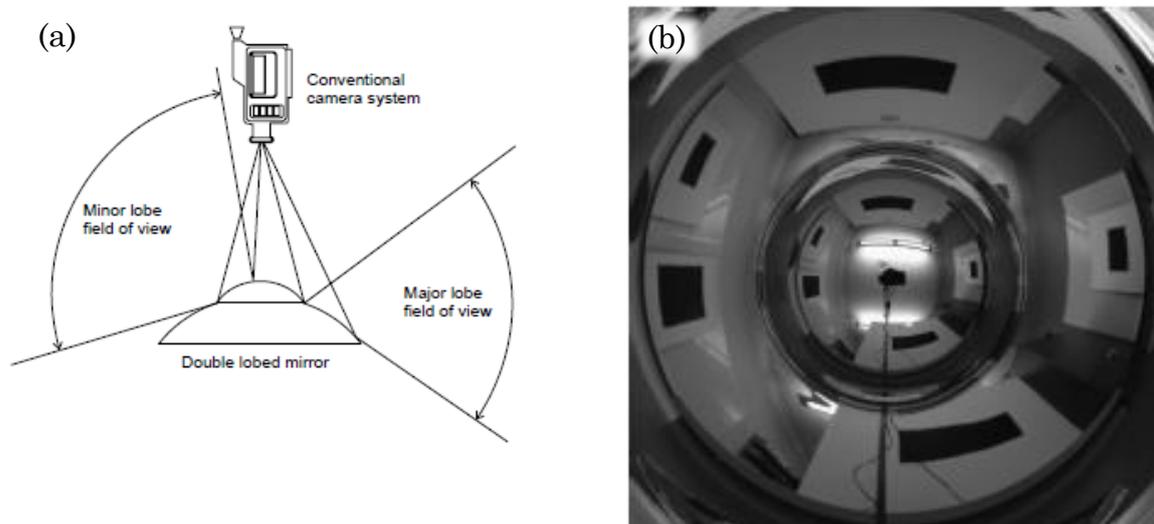


Figure 2.27 : Système catadioptrique avec miroir sphérique à double lobe. (a) Configuration du système. (b) Exemple d'image.

☑ Attributs :

- La distinction de la sphère réside dans sa symétrie par rapport à son centre, ce qui facilite l'étalonnage matériel du système catadioptrique.
- Les miroirs sphériques sont faciles à manufacturer.
- Ils présentent un faible astigmatisme notamment au centre, mais il est plus important sur la périphérie.
- La distance focale nécessaire pour faire converger les rayons est relativement faible.

☒ Incommodités :

- Les cameras catadioptriques à miroir sphérique ne sont pas centrales.
- Les distorsions sont fortes sur la périphérie de l'image catadioptrique.

Dans le cas des **miroirs hyperboliques**, la contrainte de SVP est satisfaite lorsque le sténopé et le point de vue du miroir sont placés sur l'un des deux foyers de l'hyperboloïde. En pratique, la réalisation d'une telle configuration est une tâche alambiquée. Ollis [59] avait simulé diverses configurations en utilisant deux miroirs hyperboliques avec une ou deux caméras. Les mesures de profondeur ont

été obtenues par triangulation. Le problème de correspondance a été résolu par utilisation d'une fenêtre de correspondance entre les deux images stéréoscopiques, au moyen d'un algorithme de corrélation qui tient en compte de la courbure des miroirs. En 2001, une équipe de l'Université d'Osaka a publié les résultats obtenus avec un système stéréo omnidirectionnel à deux miroirs hyperboliques (figure 2.29) pouvant opérer en temps réel, et de ses applications à la détection d'obstacles et au suivi, pour un robot mobile.

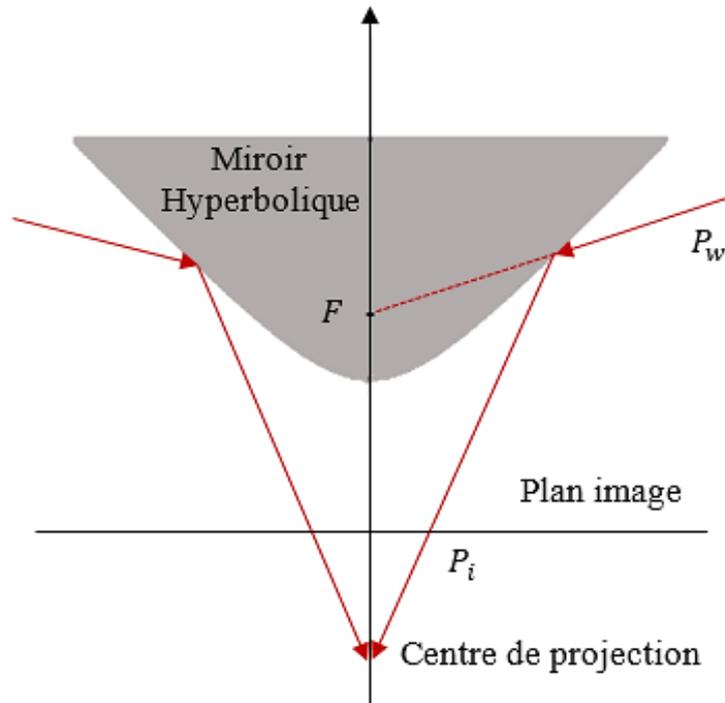


Figure 2.28 : Géométrie de formation de l'image acquise par caméra catadioptrique à miroir hyperbolique.

☑ **Attributs :**

- A centre de projection unique.
- L'astigmatisme dépend de la courbure du miroir.

☒ **Incommodités :**

- Le procédé de fabrication des miroirs hyperboliques est difficile.
- Pratiquement la contrainte du point de vue unique est difficile à satisfaire.

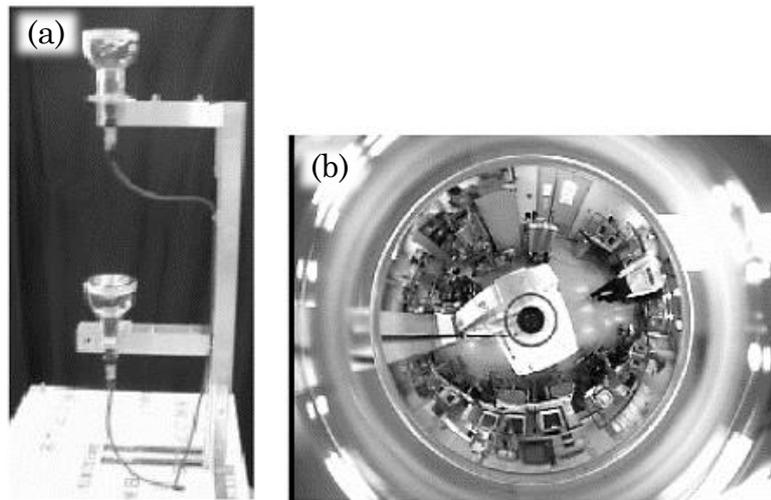


Figure 2.29 : système stéréo omnidirectionnel à miroirs hyperboliques. (a) configuration du système. (b) exemple d'image.

Une **parabole** est l'ensemble de tous les points dans le plan situés à égale distance de la directrice L , et du foyer F . Le paramètre est la distance entre la directrice et le foyer, donnée par $f = 2a$, où a est la distance du sommet à la directrice ou le foyer. La surface de révolution obtenue par rotation d'une parabole autour de son axe de symétrie est appelé un paraboloid. Toutefois, les miroirs avec cette surface sont communément appelés **miroirs paraboliques**. La forme parabolique est une solution de la contrainte SVP dans le cas particulier d'une projection orthographique. Le miroir parabolique fonctionne de la même manière que l'antenne parabolique : les rayons entrants passent par le point focal et sont réfléchis parallèlement à l'axe de rotation de la parabole. Par conséquent, un miroir parabolique doit être utilisé en conjonction avec une caméra orthographique. Une caméra en perspective peut également être utilisée si elle est placée très loin du miroir de sorte que les rayons réfléchis peuvent être estimés comme parallèle. De toute évidence, cette solution fournira une faible résolution, inacceptable, et n'a aucune valeur pratique. Gluckman et Nayar [60] ont développé un capteur panoramique de vision stéréo, avec deux miroirs paraboliques alignés sur l'axe vertical avec leurs caméras correspondantes. L'ensemble du système a été étalonné au moyen d'un procédé itératif, l'estimation de la profondeur a été basée sur la géométrie épipolaire. En 2000, Strum [61] a mis en place une méthode de reconstruction 3D à partir d'une seule image panoramique acquise avec un capteur omnidirectionnel à miroir parabolique. Le procédé obtient en entrée un ensemble de contraintes sur du structure 3D, fourni par l'utilisateur. Ces contraintes font référence à la coplanarité, parallélisme et la perpendicularité des objets dans l'image. L'auteur a décrit un procédé d'étalonnage simple en ajustant un cercle à la frontière de l'image de scène réfléchi par le miroir. L'algorithme de

reconstruction utilise la propriété SVP de cette configuration catadioptrique qui permet la reprojection inverse des points d'image 2D.

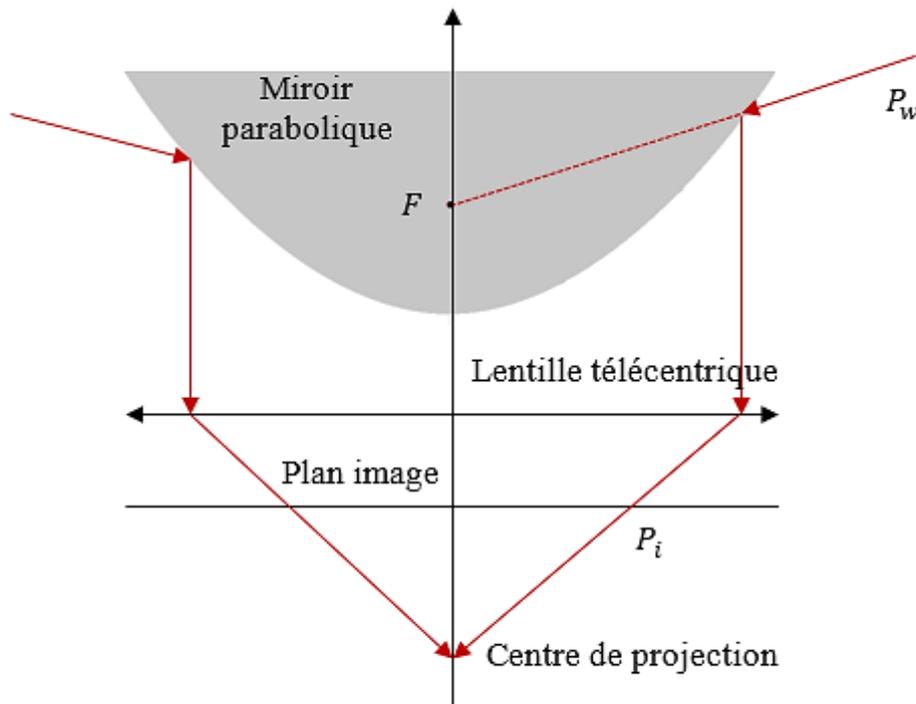


Figure 2.30: Géométrie de formation de l'image acquise par caméra catadioptrique à miroir parabolique.

☑ Attributs :

- La projection orthographique est réalisable par utilisation d'une lentille télécentrique, qui permet de sélectionner les rayons parallèles à l'axe optique de caméra.
- La contrainte SVP est satisfaite.
- Le choix de la distance lentille miroir est flexible.

☒ Incommodités :

- La taille de la lentille télécentrique est considérable.
- Le prix de la lentille télécentrique est considérable.

Le tableau 2-2 résume les propriétés des différents miroirs de révolution utilisés pour concevoir une caméra catadioptrique. Néanmoins pour concevoir un système omnistéréo un second choix doit être fait, celui de la configuration stéréo. Une multitude de dispositions sont possibles parmi elles, la configuration horizontale, verticale, N-Oculaire, et la configuration dynamique. Elles diffèrent principalement par la géométrie selon laquelle les capteurs omnidirectionnels sont

disposés. Chaque configuration a ses propres caractéristiques concernant le champ de vue, les occlusions, la géométrie épipolaire, et la précision sur l'estimation de la profondeur.

Tableau 2-2 : Tableau récapitulatif des miroirs principaux.

| Forme du miroir | Equation de la surface | Champ de vue (vertical) | SVP | Lentille |
|-----------------|---|-------------------------|-----|---------------|
| Conique | $x^2 + y^2 = z^2(\tan \alpha)$ | [-45 45] | Non | Normale |
| Sphérique | $x^2 + y^2 + z^2 = r^2$ | [-90 10] | Non | Normale |
| Hyperbolique | $\frac{(z+\frac{d}{2})^2}{a^2} - \frac{x^2}{b^2} + \frac{y^2}{b^2} = 1$ $a = \frac{\sqrt{d^2 + 4p^2} - 2p}{2}$ $b = \sqrt{p(\sqrt{d^2 + 4p^2} - 2p)}$ | [-90 45] | Oui | Normale |
| Parabolique | $\sqrt{x^2 + y^2 + z^2} = z + 2p$ | [-90 45] | Oui | Télécentrique |

2.3.2.2 Choix de la configuration

Une image omnidirectionnelle à 360° degrés, peut être considérée comme une projection cylindrique à partir d'un seul point de vue où la projection dans la direction verticale est perspective [62]. La projection (θ, z) d'un point 3D, $P(X, Y, Z)$ dans une telle image peut être représentée par :

$$(\theta, z) = \left(\tan^{-1} \left(\frac{Y}{X} \right), F \frac{Z}{D} \right) \quad 2.11$$

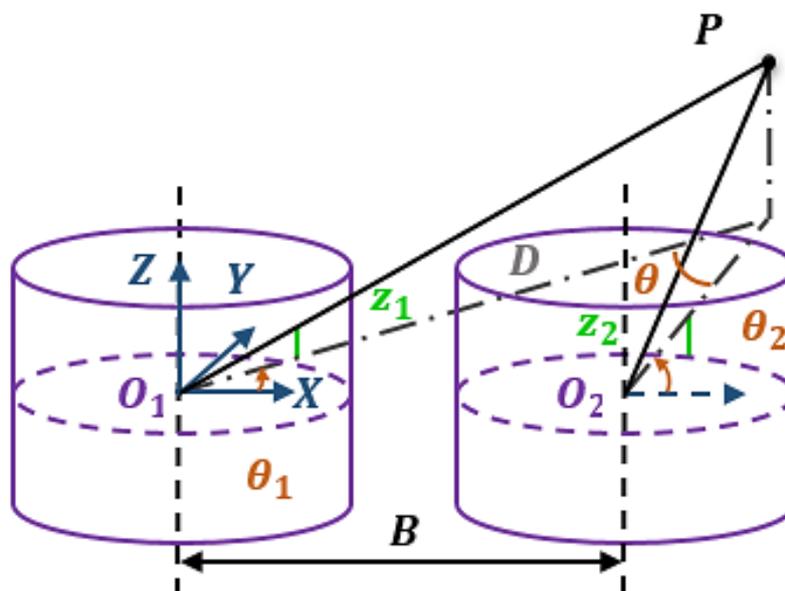


Figure 2.31 : Configuration omnistéréo horizontale.

Où $0 \leq \theta < 360^\circ$ est la direction des rayons dans le champ de vision horizontal, $D = \sqrt{X^2 + Y^2}$ est la distance perpendiculaire entre le point P et l'axe vertical OZ du cylindre tout en passant par le point focal. F est la distance focale du cylindre virtuel de la camera omnidirectionnelle (figure 2.31).

La configuration omnistéréo **binoculaire-H** [63] [64] [65] [66] est réalisable en alignant deux cameras omnidirectionnelles **horizontalement**, compte tenu des points correspondants de la paire stéréo, (θ_1, z_1) et (θ_2, z_2) d'un point P du repère 3D, la profondeur de ce point par rapport à la première camera, peut être calculée par triangulation (Figure 2.31).

$$D = B \frac{\sin\theta_2}{\sin\theta_2 - \sin\theta_1} = B \frac{\sin\theta_2}{\sin\theta} \quad 2.12$$

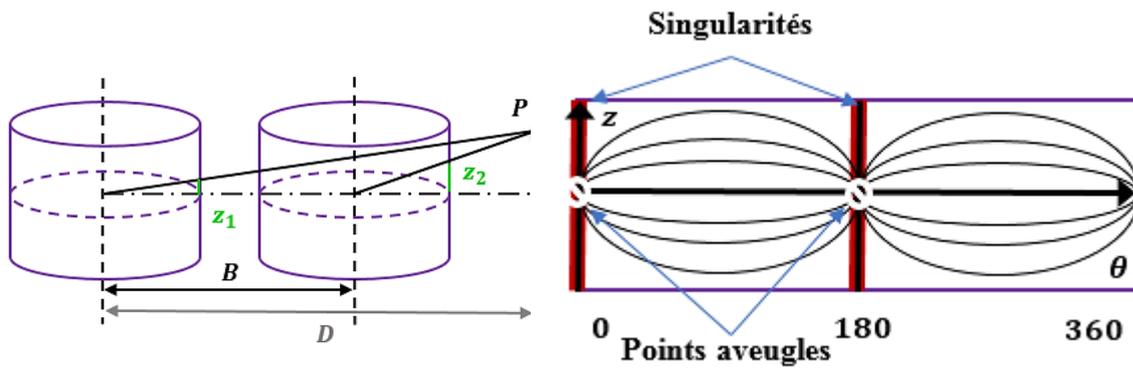


Figure 2.32 : (a) Triangulation omnistéréo horizontale. (b) Les lignes épipolaires.

$\theta = \theta_2 - \theta_1$ est la disparité horizontale, et également l'angle de vergence entre les deux rayons incidents. À la différence de la configuration stéréo perspective (planaire), cette équation de triangulation est invalide dans les cas singuliers où le point est aligné avec la ligne de base, lorsque $\theta = 0^\circ$ ou $\theta = 180^\circ$ (Figure 2.32.b). Généralement l'erreur sur le calcul de la profondeur est généralement estimée par :

$$\partial D = \frac{D\sqrt{D^2 - B^2 \sin^2\theta_2}}{B \sin\theta_2} \partial\theta \quad 2.13$$

Dans le cas où $D \gg B$ l'équation peut être simplifiée sous la forme :

$$\partial D = \frac{D^2}{B \sin\theta_2} \partial\theta \quad 2.14$$

Où $\partial\theta$ est l'erreur sur l'estimation de la disparité.

La précision sur la profondeur de l'omnistéréo binoculaire-H est anisotropique. Pour une ligne de base (*Baseline*) et une distance donnée, la meilleure estimation de la distance est obtenue lorsque l'angle de vergence entre les deux rayons incidents est au maximum ; $\theta_2 = 90^\circ$ ou $\theta_2 = 270^\circ$. Cette précision est proportionnelle au carré de la profondeur, et inversement proportionnelle à la longueur de la ligne de base. L'erreur est maximale (infinie) dans les cas de

singularité ($\theta = 0^\circ$ ou $\theta = 180^\circ$), il est toutefois possible de calculer la profondeur dans ces deux cas autrement [67] [68]:

$$D = B \frac{z_2}{z_2 - z_1} = B \frac{z_2}{z} \quad 2.15$$

Cette équation est valable que dans le cas $\theta = 0^\circ$ (figure 2.32.a), où z est la disparité verticale.

Dans ce sens, théoriquement, ce système ne présente que deux points aveugles. Cependant, en pratique, les deux caméras occluent l'une l'autre dans la direction de la ligne de base, de sorte qu'aucune information n'est disponible dans les régions d'occlusion mutuelles. Un autre inconvénient de l'omnistéréo binoculaire-H est sa géométrie épipolaire. Etant donné un point (θ_2, z_2) dans la deuxième image son point correspondant (θ_1, z_1) dans la première image se situe sur une courbe sinusoïdale (figure 2.32.b):

$$z_1 = \frac{\sin\theta_1}{\sin\theta_2} z_2 \quad 2.16$$

Cette équation n'est valable que lorsque $\sin\theta_2 \neq 0$, Dans les cas de singularité, les lignes épipolaires sont le long de z .

En littérature, cette configuration n'est pas privilégiée vu les inconvénients qu'elle présente. Dans [69] on propose un système de stéréovision omnidirectionnelle, composé de deux caméras catadioptriques alignées horizontalement (figure 2.33). Ce support d'acquisition avait été utilisé pour implémenter des algorithmes de cartographie, de localisation simultanées (SLAM) et SfM (*structure from motion*), basés sur l'extraction des primitives de type contours, qui sont suivi par la suite en utilisant l'algorithme *Lucas-Kanade*. Le système est destiné pour des applications en robotique mobile.

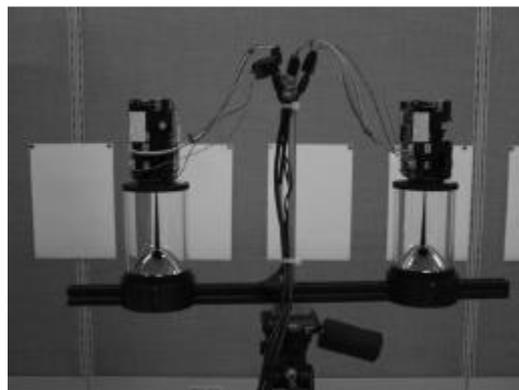


Figure 2.33 : système de vision omnistéréo binoculaire-H

Un système similaire est présenté dans [70], mais destiné à des applications dans le domaine des véhicules autonomes, et les systèmes d'assistance au conducteur. Les auteurs proposent deux solutions pour améliorer la perception de l'environnement, la première consiste en système qui regroupe une caméra

catadioptrique avec une caméra perspective (figure 2.34.a). La deuxième solution, est un système omnistéréo composé de deux cameras catadioptriques à miroirs hyperboliques, installées sur chaque côté du toit du véhicule (figure 2.34.b).

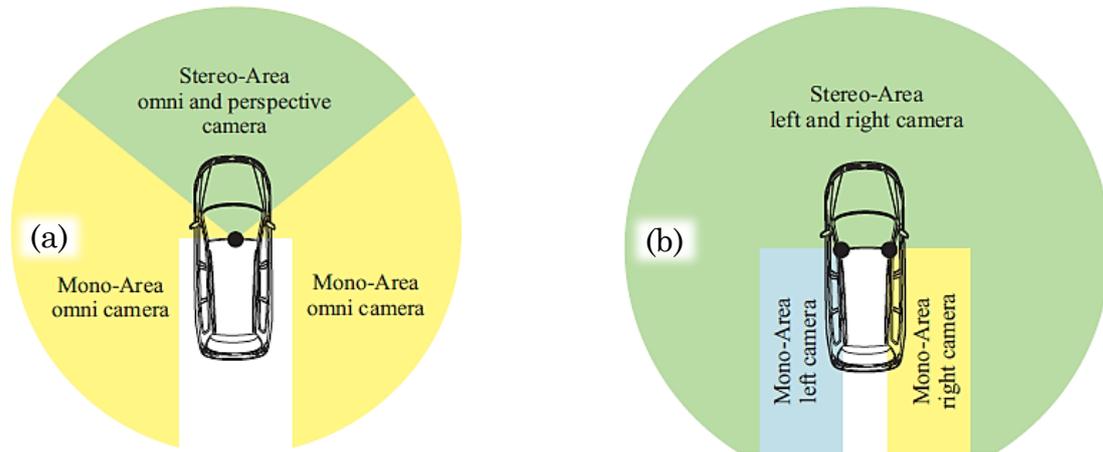


Figure 2.34 : système omnistéréo pour véhicule intelligent. (a) système hybride. (b) système catadioptrique.

L'omnistéréo **dynamique** est un cas spécial des systèmes omnistéréo à disposition horizontale. Dans ce cas la longueur de la ligne de base séparant les capteurs, n'est pas une constante, mais une variable qui change en fonction du mouvement relatif à ces capteurs. Cette technique est exploitée principalement en robotique [71].

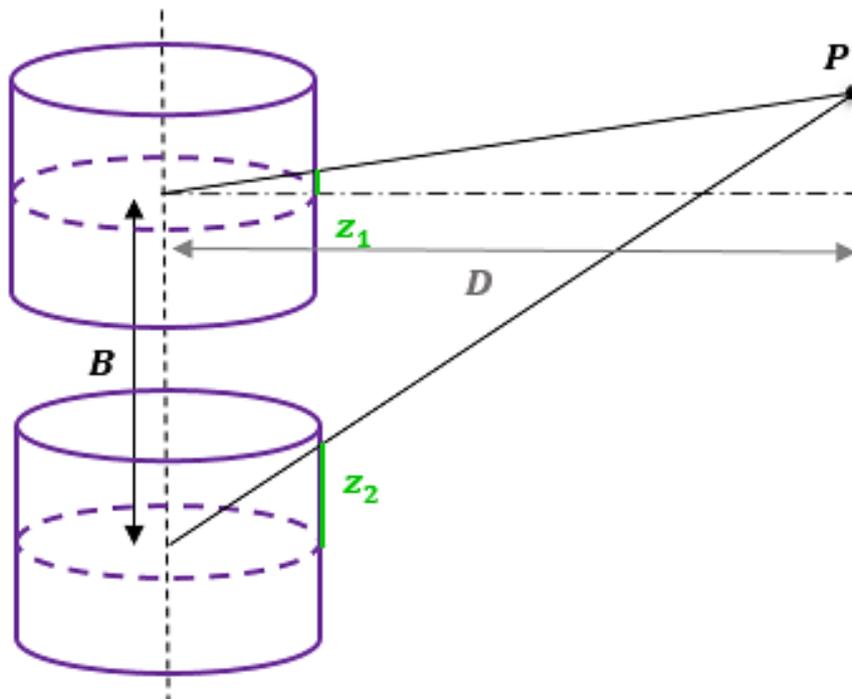


Figure 2.35 : Configuration omnistéréo binoculaire verticale.

Certains des inconvénients de l'omnistéréo binoculaire-H peuvent être surmontés en adoptant une **configuration verticale** [72] [73] [74] (figure 2.35), cela permet à la fois de simplifier la géométrie épipolaire et résout le problème de l'occlusion mutuelle.

L'équation de profondeur d'un système omnistéréo **binoculaire-V** est similaire à celle d'un système traditionnel à deux cameras perspectives :

$$D = F \frac{B}{z_2 - z_1} = F \frac{B}{z} \quad 2.17$$

B est la longueur de la ligne de base (*Baseline*), et z est la disparité verticale.

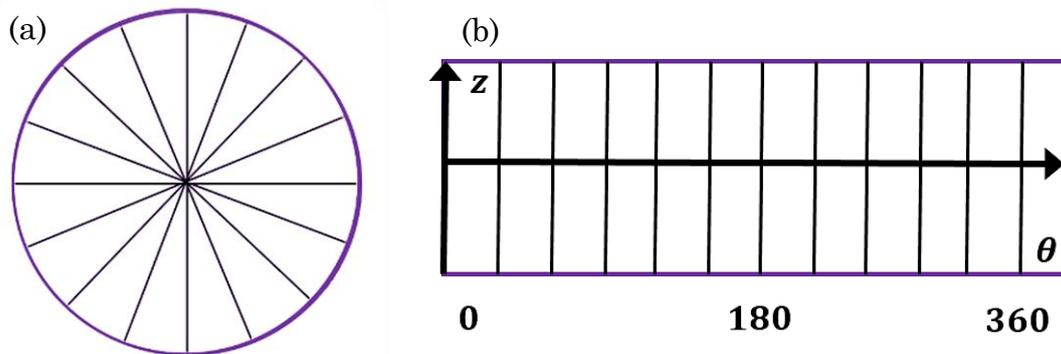


Figure 2.36 : La géométrie épipolaire d'un système omnistéréo binoculaire-V. (a) Les lignes épipolaires de l'image catadioptrique. (b) les lignes épipolaires de l'image panoramique.

Les lignes épipolaires dans ce cas sont radiales dans l'image catadioptrique, et simplement verticales dans l'image panoramique (figure 2.36). La configuration omnistéréo binoculaire-V est la plus privilégiée, et a été adoptée dans beaucoup de travaux de recherche. En 1998, Gluckman, Nayar et Thoresz [75] ont publié la conception d'un capteur stéréo panoramique compact (figure 2.37), capable de produire des cartes de profondeur panoramique en temps réel. Le système est composé de deux caméras omnidirectionnelles, coaxiales, alignées verticalement. Chaque caméra omnidirectionnelle est constituée d'une caméra classique, d'optique télécentrique, et d'un miroir parabolique. La profondeur est calculable par triangulation (figure 2.37) une fois le système calibré, et s'exprime sous la forme :

$$d = \frac{\cos(\beta)}{\sin(\alpha - \beta)} b \quad 2.18$$

Ce concept reste l'un des plus aboutis de la stéréovision omnidirectionnelle, et avait été réutilisé pour un large éventail d'applications.

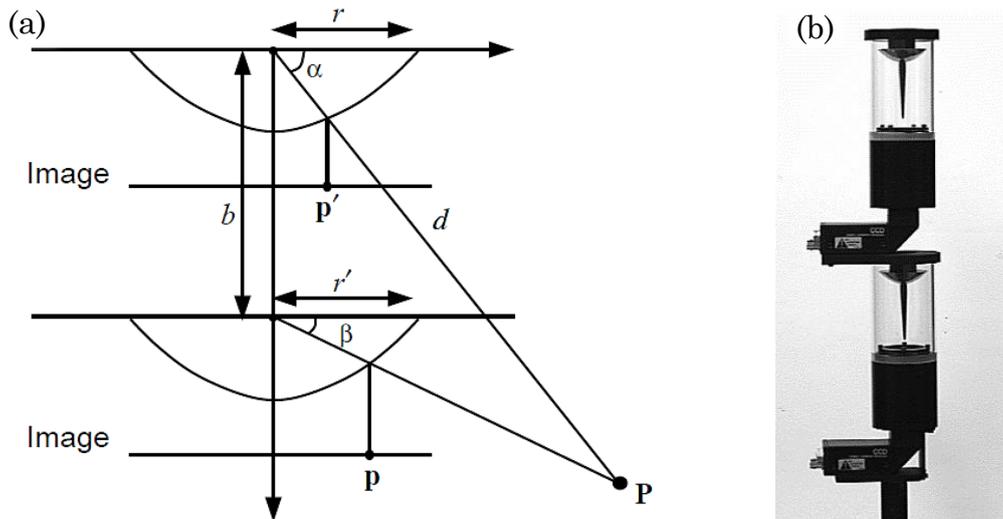


Figure 2.37 : (a) Triangulation et calcul de la profondeur (b) Système omnistéréo binoculaire-V à deux cameras catadioptriques.

Un système similaire [76] avait été utilisé pour une application en robotique mobile, constitué de deux miroirs coniques coaxiaux, ce système omnistéréo permet au robot (figure 2.38) de se situer par rapport à son environnement, et de se guider en n'exploitant que les informations visuelles acquises par le capteur.



Figure 2.38: DaX, le robot mobile Scitos G5.

Un cas particulier de cette configuration omnistéréo verticale, sont des systèmes qui n'utilisent qu'une seule camera, combinée avec deux miroirs ou un seul miroir à double lobe, ou un seul miroir avec des dispositifs optiques spéciaux. Ces systèmes ne peuvent pas être classés en tant que monoculaires, même s'ils n'impliquent qu'une seule camera, car ils acquièrent l'image de la scène à partir de points de vues espacement distincts. L'un des avantages qu'ils présentent, c'est qu'ils permettent de contourner l'étalonnage d'un système à deux cameras catadioptriques, qui est souvent une tâche compliquée.

En 2006 Liancheng Su et al., [77] ont étudié un système omnistéréo, de la calibration, à l'extraction de l'information 3D. Le système dédié à la détection d'obstacles se compose d'une unique camera orientée vers deux miroirs hyperboliques disposés verticalement (figure 2.39.a). Un trou dans le miroir du bas permet d'imager celui au-dessus ; ainsi chaque image acquise, contient deux images catadioptriques concentriques, représentant la réflexion de la scène par chacun des miroirs (figure 2.39.b).

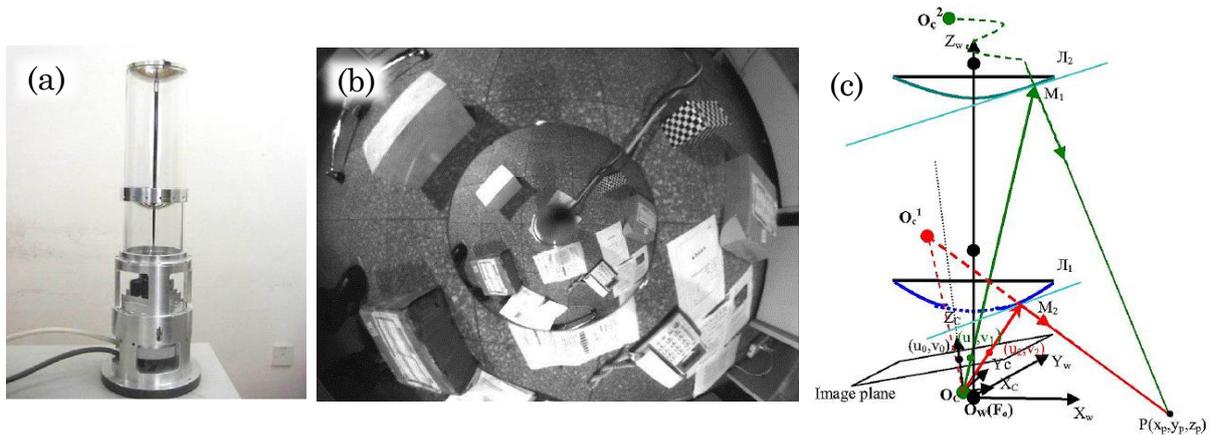


Figure 2.39 : Système omnistéréo à caméra unique et deux miroirs hyperboliques. (a) Apparence du système de vision. (b) Une image acquise avec ce système. (c) Triangulation.

Le système est conçu d'une façon que les deux foyers des miroirs coïncident avec le centre de la camera perspective. La triangulation dans ce cas (figure 2.39.c) est similaire à celle d'un système omnistéréo binoculaire-V, dans les deux cas un seul plan image est considéré.

Une autre configuration alternative n'impliquant qu'une seule camera, consiste à utiliser un miroir à double lobe. Un miroir à double lobe est une paire coaxiale de miroir, les centres de ces deux miroirs sont colinéaires avec l'axe de la caméra, et ont un profil radialement symétrique autour de cet axe (figure 2.27). Ce principe a été développé dans [78] [79] [80] [81].

D'après Sooyeong Yi et al. [82], il est possible de remplacer le miroir à double lobe par un simple miroir de révolution associé à une lentille concave, pour aboutir à un système omnidirectionnel stéréoscopique. La profondeur est calculable conformément aux lois optiques de réflexion et réfraction.

Naturellement, la précision de la profondeur des systèmes omnistéréo binoculaires peut être améliorée par l'utilisation de plus de deux caméras omnidirectionnelles. Ceci est le cas général d'omnistéréo à multiple points de vue [83] [84] [85], ou simplement l'omnistéréo **N-Oculaire** (figure 2.20.c). Les systèmes à caméra unique et à multiples points de vues, sont une classe particulière de cette catégorie [86] [87] [88]. Généralement cette dernière configuration est réalisable en pointant une caméra perspective ou orthographique, vers un groupement de miroirs réfléchissants ²⁹(figure 2.40).

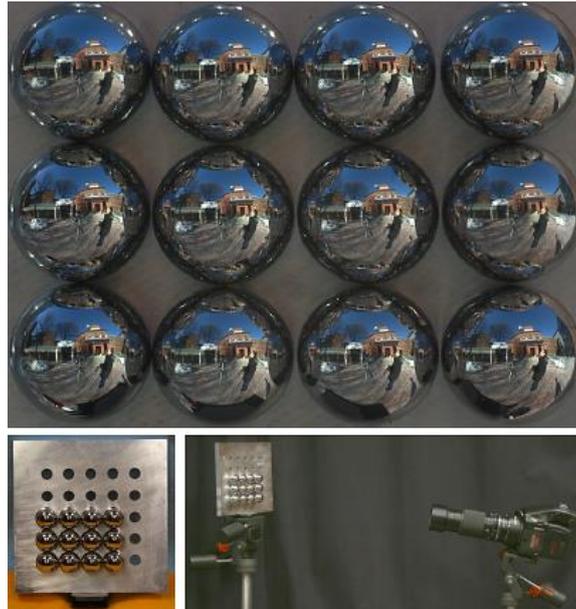


Figure 2.40 : Système omnistéréo à caméra unique et multiple miroirs (sphériques).

2.3.2.3 Notes sur la stéréovision omnidirectionnelle

La stéréovision est l'une des techniques les plus courantes pour obtenir l'information 3D sur la scène. Les systèmes omnistéréo peuvent être arrangés selon différentes configurations. Dans le cas où les deux caméras sont alignées horizontalement, quelques problèmes sont rencontrés: Le premier, est que les lignes épipolaires sont courbées sinusoidalement, ce qui mènera à l'augmentation du coût de calcul, le deuxième inconvénient est celui de l'occlusion mutuelle. L'angle de vergence n'est pas uniforme, et en plus, l'orientation des repères images et caméras influe significativement sur la géométrie épipolaire. Dans le cas de l'omnistéréo dynamique, l'erreur sur l'estimation de la longueur de la ligne de base s'ajoute aux incertitudes d'estimation de la profondeur, ce qui influe négativement sur les résultats.

²⁹ Source de figure : [86]

L'omnistéréo N-oculaire apporte de la précision, mais la géométrie épipolaire, les occlusions, et la complexité des calculs, sont les mêmes rencontrés dans le cas de l'omnistéréo horizontale.

La configuration verticale échappe aux carences révélées précédemment ; les lignes épipolaires sont simplement verticales, et l'occlusion n'affecte pas l'image de la scène environnante.

L'utilisation d'une caméra unique présente plusieurs avantages par rapport à l'utilisation de plusieurs caméras. Puisque la réponse spectrale, le gain et l'offset sont identiques pour les deux images stéréoscopiques. En outre, un seul ensemble de paramètres d'étalonnage interne est à déterminer. Sans doute le plus important avantage, est que l'utilisation d'une seule caméra simplifie l'acquisition de données, et aucun matériel ou logiciel n'est nécessaire pour la synchronisation d'acquisition. Les systèmes omnistéréo avec miroir à double lobe, peuvent produire des vues panoramiques de la scène en une seule image. Cet arrangement est compact et assure l'alignement des deux images, mais puisque les deux miroirs sont tellement rapprochés, la ligne de base devient assez restreinte, ce qui limite la disparité. Dans ce cas de figure la résolution de l'image interne est assez faible, et dans certains cas, elle est inexploitable pratiquement. Le tableau 2-3 en dessous synthétise la comparaison entre les différentes configurations omnistéréo.

Tableau 2-3 : Tableau récapitulatif des principales configurations omnistéréo.

| Configuration omnistéréo | Occlusion mutuelle | Erreur sur la profondeur | Géométrie épipolaire | Ligne de base | Points de vue |
|--------------------------|--------------------|--------------------------|----------------------|---------------|---------------|
| Binoculaire-H | Oui | $\propto D^2$ | sinusoïdal | fixe | 2 |
| Dynamique | Probable | $\propto D^{1.5}$ | sinusoïdal | variable | ≥ 2 |
| Binoculaire-V | Non | $\propto D^2$ | verticale | fixe | 2 |
| N-oculaire | Oui | $\propto D^2$ | sinusoïdal | fixe | N |

2.4 Conception d'un système de stéréovision omnidirectionnel

Cette section décrit la conception et l'étude expérimentale d'un nouveau système de vision omnistéréo à camera unique, conçu et réalisé au sein du Laboratoire d'Électronique et de Traitement de Signal/Géomatique. Vu que ce travail de recherche vise l'intégration de la vision omnidirectionnelle à des applications en robotique, il était nécessaire de développer un système de vision artificielle capable de fournir suffisamment d'informations à une machine pour qu'elle puisse interagir efficacement avec son environnement. L'étude menée, présentée dans la section précédente, et les conclusions mentionnées, nous ont convaincu d'opter pour une solution matérielle qui consiste en un système de vision catadioptrique à disposition verticale [89]. Le capteur catadioptrique est constitué d'une caméra perspective orientée vers un miroir sphérique. Le système est conçu mécaniquement de façon que le capteur soit mobile sur l'axe vertical. Ce qui permet

de pouvoir l'utiliser comme une simple camera omnidirectionnelle, ou comme un système de stéréovision omnidirectionnelle. Dans la section qui suit nous allons détailler l'architecture matérielle du capteur proposé, sa configuration, et aussi les résultats d'estimation de la profondeur obtenus.

2.4.1 Architecture matérielle du capteur proposé

Le capteur catadioptrique est composé d'une caméra CCD à haute résolution ³⁰(Logitech C310), d'interface USB 2.0, avec une fréquence d'acquisition de 30 fps, et une résolution de 1280 x 720 pixels. La caméra est orientée vers un miroir sphérique d'un rayon de 60mm, avec un champ de vision de 360° horizontal et 200° vertical. Pour faciliter l'étalonnage du système, chacune des composantes du capteur catadioptrique détient deux degrés de liberté sur le plan horizontal, et un degré de liberté sur l'axe verticale. Cette architecture est illustrée sur la figure 2.41.a :

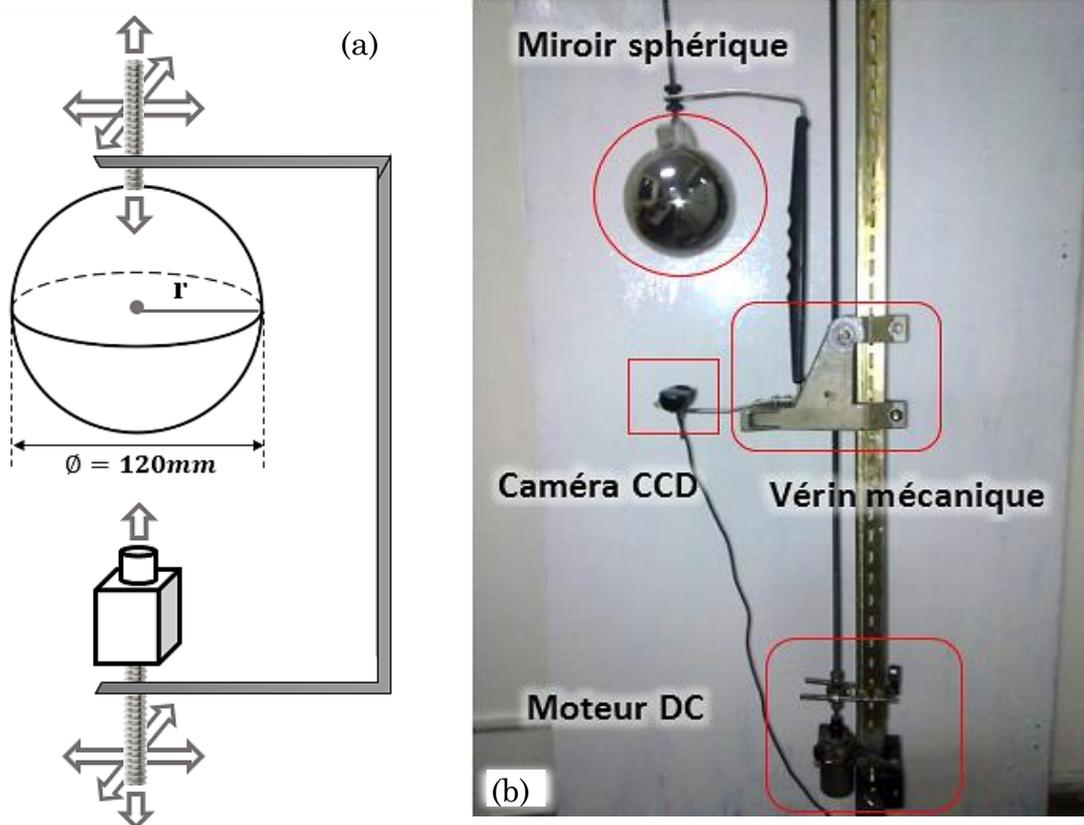


Figure 2.41 : (a) Architecture du capteur catadioptrique proposé. (b) Configuration du système omnistéréo.

Pour créer de la disparité spatiale lors de l'acquisition, nous avons muni le capteur catadioptrique d'un système de motorisation à vérin mécanique, lui permettant de mouvoir sur l'axe verticale (figure 2.41.b). Ainsi l'acquisition des images peut se

³⁰ Lien : [159]

faire à partir de positions distantes. Cette configuration est équivalente à un système omnistéréo binoculaire-V à *Baseline* variable.

2.4.2 Étalonnage du capteur

Pour pouvoir utiliser une caméra catadioptrique convenablement, quelques contraintes géométriques concernant sa configuration, doivent être satisfaites. Dans le cas d'un capteur catadioptrique à miroir sphérique ; une forme géométrique qui se distingue des autres quadriques par sa symétrie centrale, le réglage du système se réalise essentiellement par alignement de l'axe optique de la caméra avec le centre du miroir sphérique. Ce processus est à la fois mécanique et algorithmique. Nous avons développé une interface logicielle qui permet de faciliter cette tâche, en délimitant la projection théorique idéale du miroir sur l'image catadioptrique. Nous agissons ensuite mécaniquement sur l'agencement de la caméra et du miroir, jusqu'à ce que la projection réelle du miroir coïncide avec sa position optimale, ce processus est illustré sur la figure 2.42. Nous ajustant la distance focale de la camera par la suite, pour avoir une netteté maximale de l'image. Cette procédure n'est faite qu'une seule fois, puisque la configuration du système demeure intacte mécaniquement.

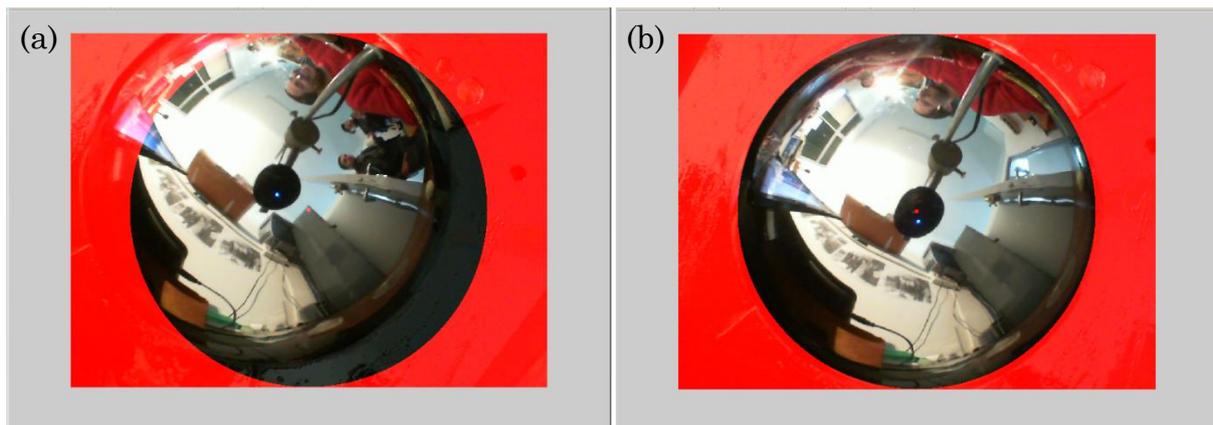


Figure 2.42 : Illustration du processus d'ajustement du capteur catadioptrique. (a) Capteur non ajusté. (b) Capteur ajusté

L'étalonnage du système est basé sur le modèle de réflexion [90] [91]; un modèle qui décrit la formation de l'image en se basant sur les lois de réflexion de la lumière, sur la surface d'un miroir et suppose que le miroir est parfaitement spéculaire ; c'est-à-dire que l'angle d'incidence est égal à l'angle de réflexion. Par conséquent, pour un point 3D donné dans l'espace, il est possible de trouver son correspondant correct par réflexion sur le miroir et figurant sur l'image catadioptrique. La projection du point miroir sur le plan image de la caméra est alors directe [92]. Notons qu'en utilisant cette méthode, la propriété du point de vue unique, dans le cas des caméras centrales, devient une conséquence et non pas une contrainte, ce qui signifie que le modèle peut être appliqué à tout type de miroir, peu importe s'il possède un point de vue unique ou pas.

2.4.3 Estimation de la profondeur par triangulation passive

La triangulation passive est une technique qui permet de déterminer la profondeur d'un point de l'espace 3D, à partir des coordonnées des pixels qui lui correspondent dans les deux images stéréoscopiques. Dans sa forme la plus simple, on associe un point 3D à deux uniques pixels appariés, chacun est situé sur une image de la paire des images stéréo.

Dans le cas du système omnistéréo proposé l'acquisition de la scène se fait à partir de deux points de vue espacés verticalement. On note que la longueur de la ligne de base est ajustable, et que les deux images étant acquises avec le même capteur, sont superposables.

Considérant un point P_w de l'espace, avec P_1 et P_2 ses deux points correspondants sur le plan image. La profondeur D est la distance qui sépare P_w de l'axe optique Δ du système omnistéréo. La figure 2.43 représente le schéma de la triangulation permettant de calculer cette profondeur.

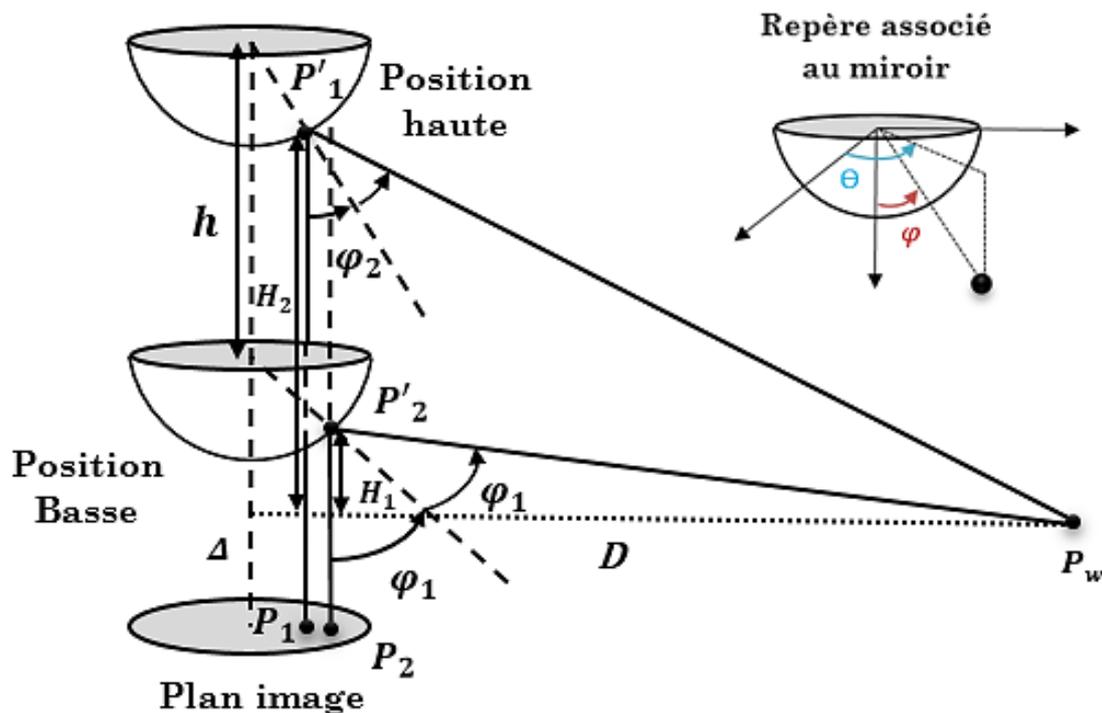


Figure 2.43 : Triangulation dans le cas du système omnistéréo proposé.

φ_1 et φ_2 désignent successivement les demi-angles formés par les rayons incidents, à position basse et à position haute, et par ceux réfléchis. H_1 et H_2 sont les distances entre les deux points de réflexion P'_1 et P'_2 et leurs projection sur le plan horizontal passant par le point P_w . Par inspection de la configuration géométrique ainsi obtenue nous pouvons poser les équations suivantes :

$$\tan(2\varphi_1) = \frac{D_1}{H_1} \text{ et } \tan(2\varphi_2) = \frac{D_2}{H_2} \quad 2.19$$

R étant le rayon du miroir sphérique nous avons :

$$D_i = D - R \sin(\varphi_i), i=1, 2. \quad 2.20$$

Et :

$$H_2 = H_1 + h - \lambda \quad 2.21$$

Avec λ définit comme par :

$$\lambda = R(\cos(\varphi_2) - \cos(\varphi_1)) \quad 2.22$$

h est la longueur de la ligne de base, qui représente le pas du mouvement du capteur le long de l'axe vertical. En exploitant les équations 1 et 2 nous obtenons l'équation suivante :

$$\tan(2\varphi_2) = \frac{(D - R \sin(\varphi_2))}{\frac{(D - R \sin(\varphi_1))}{\tan(2\varphi_1)} + h - \lambda} \quad 2.23$$

À partir de laquelle nous déduisons l'expression de la profondeur :

$$D = \varepsilon + \frac{(h - \lambda) \times \tan(2\varphi_1) \times \tan(2\varphi_2)}{\tan(2\varphi_1) - \tan(2\varphi_2)} \quad 2.24$$

Avec ε définit comme suit :

$$\varepsilon = \frac{R(\tan(2\varphi_2) \sin(\varphi_1) - \tan(2\varphi_1) \sin(\varphi_2))}{\tan(2\varphi_2) - \tan(2\varphi_1)} \quad 2.25$$

Il s'avère que la profondeur est une fonction du rayon du miroir, qui est une constante, de la disparité, et de la longueur de la ligne de base h . Dans le but d'évaluer la précision du système, nous proposons d'étudier l'expression de l'erreur sur l'estimation de la profondeur ; considérons les deux points P_1 et P_2 du plan image est leurs angles correspondants φ_1 et φ_2 reliés au point P_w de l'espace 3D :

Puisque $\lambda \ll h$ et $\varepsilon \ll \frac{h \times \tan 2\varphi_1 \times \tan 2\varphi_2}{\tan(2\varphi_1) - \tan(2\varphi_2)}$

L'expression de la profondeur peut être simplifiée sous la forme :

$$D \approx \frac{h \times \tan 2\varphi_1 \times \tan 2\varphi_2}{\tan(2\varphi_1) - \tan(2\varphi_2)} = - \frac{h \sin(2\varphi_1) \sin(2\varphi_2)}{\sin(\varphi)} \quad 2.26$$

$\varphi = 2\varphi_1 - 2\varphi_2$ étant la disparité verticale, l'erreur sur la disparité s'exprime comme suit :

$$\frac{\delta D}{\delta \varphi} = \frac{D}{h \sin(2\varphi_1) \sin(2\varphi_2)} \sqrt{D^2 - (h \sin(2\varphi_1) \sin(2\varphi_2))^2} \quad 2.27$$

Etant donné que $h \ll D$ l'expression générale de l'erreur sur la profondeur estimée est la suivante :

$$\delta D \approx \frac{D^2}{h \sin(2\varphi_1) \sin(2\varphi_2)} \delta\varphi \quad 2.28$$

Avec $\delta\varphi$ l'erreur sur la disparité verticale.

Nous pouvons conclure que l'erreur générée par le système est proportionnelle au carré de la distance D est inversement proportionnelle à la longueur de la ligne de base h .

2.4.4 Résultats expérimentaux

Afin d'évaluer les performances du système prototype réalisé, notamment sa précision d'estimation de la profondeur, nous avons établi une procédure d'évaluation expérimentale qui consiste à utiliser ce système de stéréovision omnidirectionnel pour estimer la profondeur d'objets disposés à des distances préalablement connues.

L'objet est représenté par une tâche laser ponctuelle, issue d'un projecteur laser pointé sur un tableau. Ce choix est justifié par la volonté de minimiser l'erreur engendrée lors de la détermination du centre de l'objet, qui représentera le point P_w de l'espace 3D, lors de la triangulation.

Les expérimentations initiales menées visent l'étude de la variation de l'erreur sur la profondeur, en fonction de la variation de la longueur de la ligne de base ; la tâche laser représentant le point P_w est positionnée à des distances prédéterminées de l'axe optique Δ du système omnistéréo, à chaque position la profondeur est déterminée à différentes valeurs de h (*Baseline*).

La profondeur est calculée par deux prises de vues successives du même objet, le capteur étant initialement à position basse, une première image est acquise. Le

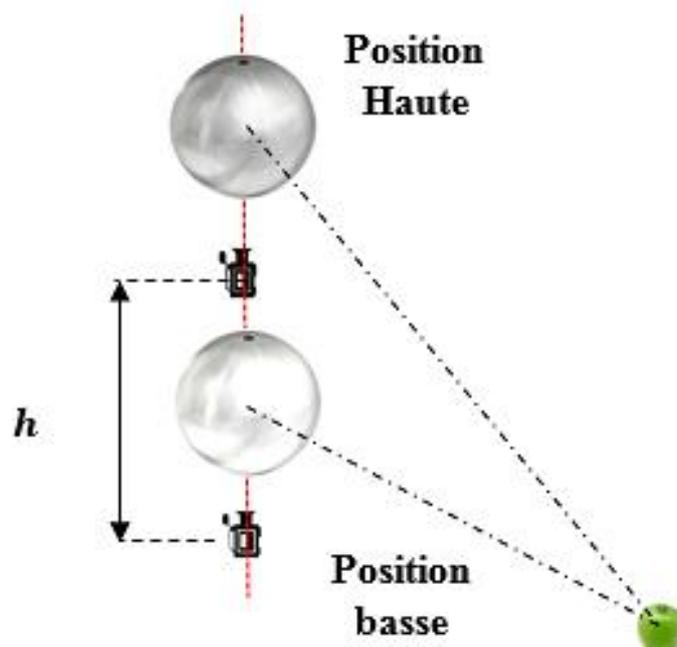


Figure 2.44 : Procédure d'estimation de la profondeur.

capteur est déplacé vers sa position haute, ensuite une seconde image est capturée, cette procédure est illustrée sur la figure 2.44.

Les coordonnées du point P_w dans les deux images catadioptriques sont extraites manuellement, Cela permet de déterminer les valeurs de φ_1 et φ_2 , la valeur de h étant antérieurement choisie, la profondeur peut alors être calculée en exploitant l'équation 2.24. La figure 2.45 représente un échantillon d'une paire d'images stéréo, sur lesquelles le spot laser est visible.

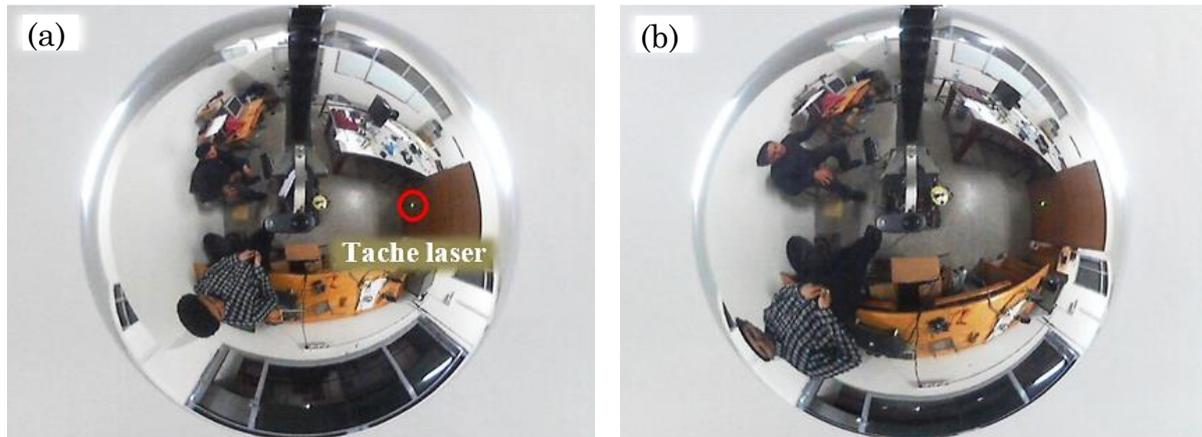


Figure 2.45 : Une paire d'images omnistéréo. (a) Image acquise à position haute. (b) Image acquise à position basse.

Les résultats de cette première phase d'expérimentation sont reportés sur le tableau 2-4, tableau 2-5, et tableau 2-6.

Tableau 2-4 : Variation de l'erreur sur la profondeur en fonction de h pour $D=2.18m$

| Profondeur réelle (m) | 2.18 | | | | | | |
|------------------------|------|------|------|------|------|-------------|-------|
| h (m) | 0.80 | 0.90 | 1.00 | 1.10 | 1.20 | 1.30 | 1.40 |
| Profondeur estimée (m) | 2.38 | 2.37 | 2.32 | 2.28 | 2.23 | 2.19 | 2.13 |
| Erreur (m) | 0.2 | 0.19 | 0.14 | 0.1 | 0.05 | 0.01 | -0.05 |
| Erreur (%) | 9.17 | 8.72 | 6.42 | 4.59 | 2.29 | 0.46 | 2.29 |

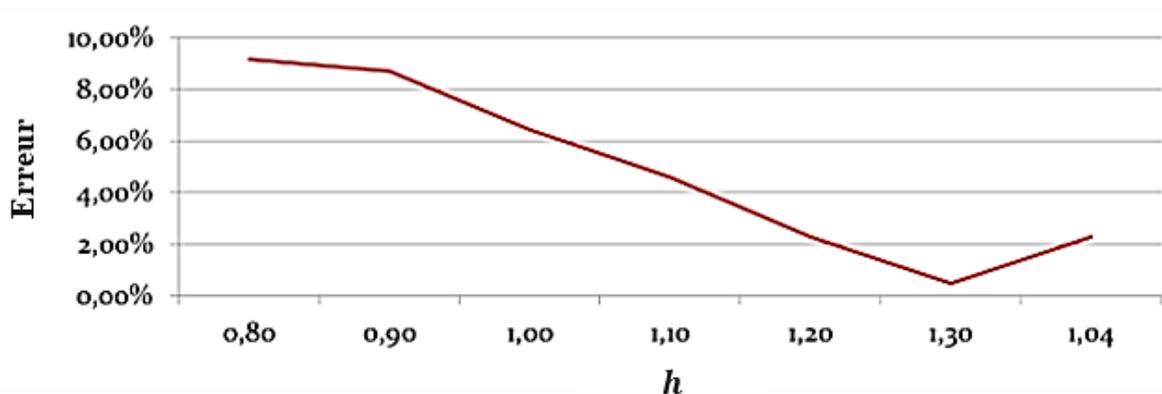
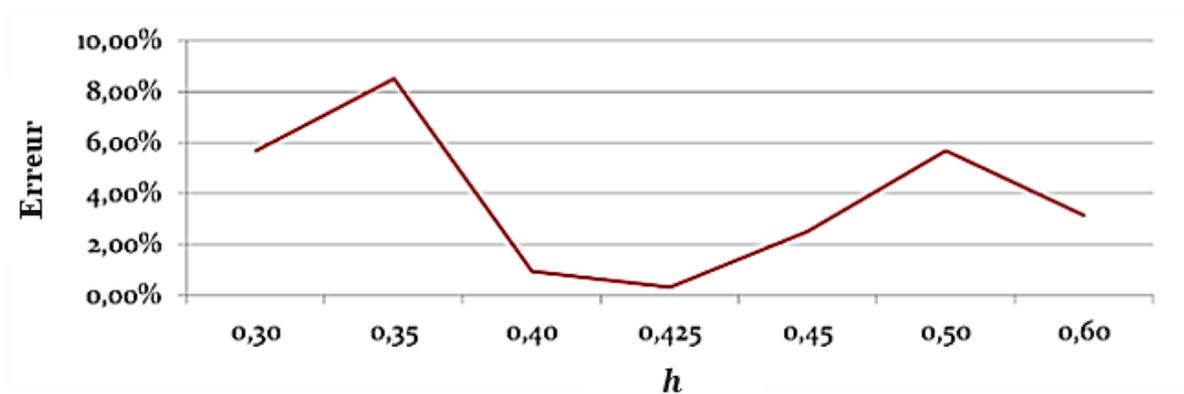


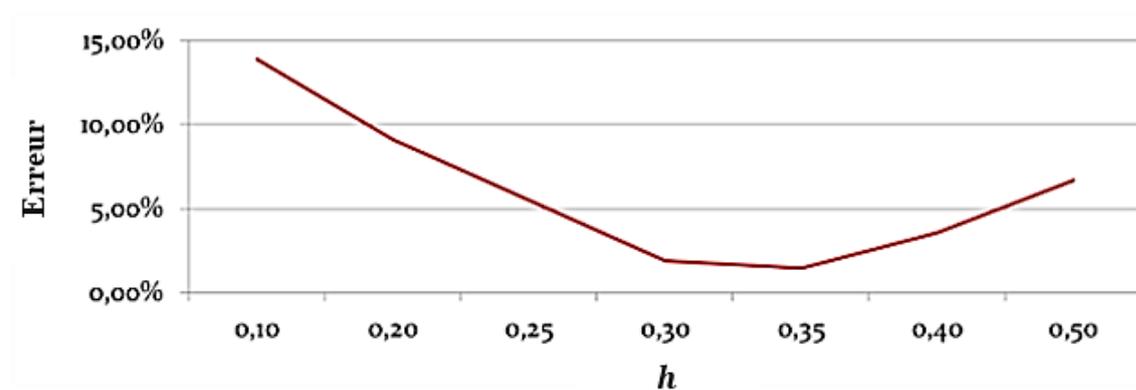
Figure 2.46 : Courbe de variation de l'erreur, $D=2.18$.

Tableau 2-5 : Variation de l'erreur sur la profondeur en fonction de h pour $D=3.18m$

| Profondeur réelle (m) | 3.18 | | | | | | |
|------------------------|-------|-------|-------|-------------|-------|-------|------|
| h (m) | 0.30 | 0.35 | 0.40 | 0.425 | 0.45 | 0.50 | 0.60 |
| Profondeur estimée (m) | 3 | 2.91 | 3.15 | 3.17 | 3.1 | 3.00 | 3.08 |
| Erreur (m) | -0.18 | -0.27 | -0.03 | -0.01 | -0.08 | -0.18 | -0.1 |
| Erreur (%) | 5.66 | 8.49 | 0.94 | 0.31 | 2.52 | 5.66 | 3.14 |

Figure 2.47 : Courbe de variation de l'erreur, $D=3.18$.Tableau 2-6 : Variation de l'erreur sur la profondeur en fonction de h pour $D=4.18m$

| Profondeur réelle (m) | 4.18 | | | | | | |
|------------------------|-------|------|-------|------|-------------|-------|-------|
| h (m) | 0.10 | 0.20 | 0.25 | 0.30 | 0.35 | 0.40 | 0.50 |
| Profondeur estimée (m) | 4.76 | 4.56 | 3.95 | 4.26 | 4.12 | 4.03 | 3.90 |
| Erreur (m) | 0.58 | 0.38 | -0.23 | 0.08 | -0.06 | -0.15 | -0.28 |
| Erreur (%) | 13.88 | 9.09 | 5.50 | 1.91 | 1.44 | 3.59 | 6.70 |

Figure 2.48 : Courbe de variation de l'erreur, $D=4.18$.

2.4.5 Discussion et synthèse

En essayant d'interpréter l'expression générale de l'erreur (Eq. 2.28), nous pouvons conclure précocement, que la valeur de cette erreur est inversement proportionnelle à la longueur de la ligne de base h ; plus on augmente h moins d'erreur est générée. Sauf que les résultats empiriques illustrés par les courbes 2.46, 2.47 et 2.48 contredisent cette conclusion ; la précision s'accroît avec h jusqu'à une certaine limite, et commence à décroître. Cela se justifie par le fait que l'équation 2.28 s'établit par supposition que $h \ll D$, dès que cette supposition n'est plus valable, l'erreur n'obéit plus à cette équation, mais à celle qui la précède : Eq. 2.27.

A chaque distance D donnée, correspond une unique valeur optimale de h qui minimise l'erreur générée.

Le système de stéréovision omnidirectionnelle développé, se distingue par sa longueur de ligne de base ajustable, ce qui s'avère utile, vu qu'à chaque gamme de profondeurs manipulées, correspond une valeur de h qui optimise la précision du système. Néanmoins l'acquisition d'une paire d'images stéréo nécessite le déplacement du capteur le long de l'axe vertical, une opération qui ne dure qu'environ sept secondes pour la plus grande des valeurs possibles de la longueur de la ligne de base (1.4m), ce qui réduit le champ d'application d'un tel système aux scènes statiques.

L'adoption d'un concept à camera unique simplifie la procédure d'étalonnage du système, comme elle élimine le problème d'occlusion mutuelle, la configuration verticale simplifie elle à son tour la géométrie épipolaire du système ; la recherche des pixels homologues se fait le long d'une droite radiale dans l'image catadioptrique, et le long d'une droite verticale dans l'image panoramique.

2.5 Conclusion

Ce chapitre présentait une classification des technologies des capteurs d'acquisition de données 3D. Parmi ces technologies : les systèmes actifs qui utilisent l'émission et la réception d'ondes pour mesurer la distance entre le capteur et la cible, et les systèmes passifs, il s'agit de techniques d'acquisition par vision.

Les systèmes actifs tels que les radars et les lidars sont principalement destinés à des applications exigeant des portées importantes, vu que les caractéristiques des ondes utilisées ne permettent pas une mesure précise de distances courtes. Les télémètres à ultrason sont plutôt appropriés aux environnements intérieurs. Les capteurs actifs souffrent également d'une résolution azimutale faible. Cet inconvénient peut être pallié en utilisant des systèmes de balayage, une solution qui requiert des déplacements consécutifs de la source d'émission. Cette solution est donc peu envisageable dans le cas d'applications temps-réel.

Les technologies d'acquisition 3D passive se distinguent par leur capacité à acquérir une quantité d'information importante de façon instantanée. La stéréovision omnidirectionnelle est l'une de ces techniques, qui s'avère adaptée pour la numérisation en temps-réel des scènes en trois dimensions et à large champ de vue. Nous avons souligné les systèmes omnistéréo catadioptriques comme solution particulière qui répond davantage aux exigences des applications dans le domaine de la robotique mobile. Ces systèmes peuvent être classés selon le nombre de point de vue sous lesquels la scène est observée, la forme des miroirs utilisés, ainsi que l'arrangement spatial des capteurs catadioptriques constituant le système. La configuration verticale est particulièrement intéressante, car elle présente plusieurs avantages, parmi lesquels la simplicité de la géométrie épipolaire, et la préservation du large champ de vue.

Il n'y a pas de système d'acquisition 3D à large champs de vue qui convient partout, mais à chaque application correspond une configuration qui respecte le mieux son cahier de charges. Le choix doit prendre en considération les attributs et les inconvénients de chaque système.

Nous avons présenté une solution matérielle qui s'est orienté vers un système non-intrusif, didactique, et configurable. Il consiste en un capteur omnistéréo, à camera unique et à disposition verticale. La capacité à la fois basique et capitale du prototype réalisé à estimer la profondeur avait été évaluée expérimentalement. Les résultats sont conformes aux fondements théoriques du système. La conclusion majeure à retenir est que le choix du dimensionnement, notamment la longueur de la ligne de base, d'un système omnistéréo doit tenir compte de la gamme de profondeurs dans laquelle le système opère. Un concept à baseline ajustable tel que celui proposé, offre l'avantage d'agir sur la précision du capteur en fonction des conditions de son utilisation.

En résumé, l'utilisation des systèmes de stéréovision catadioptrique permet de réduire les coûts de fabrication des capteurs 3D omnidirectionnels. Ce qui rend ce type de capteurs plus concurrentiel sur le marché des solutions de vision par ordinateur, qui est un facteur déterminant pour la croissance de cette branche de vision artificielle. Vu leur potentiel, les caméras catadioptriques ont été profondément étudiées dans les vingt dernières années. Aujourd'hui elles bénéficient d'un fondement mathématique important.

3. Contribution au traitement des images omnidirectionnelles

3.1 Introduction

Le traitement d'image est une discipline qui peut être considérée comme un maillon de la chaîne de vision artificielle, puisqu'elle englobe l'ensemble des transformations permettant à une machine d'interpréter les images traitées avec plus de fiabilité. Comme elle peut être considérée plus généralement comme l'ensemble des techniques permettant d'agir sur les caractéristiques d'une image, ainsi que sa représentation. Il n'y a pas de limites claires dans le continuum du traitement d'image d'un côté, et de la vision par ordinateur à l'autre côté. Cependant, un paradigme utile est de considérer trois types de traitements informatisés dans ce continuum : bas niveau, moyen niveau, et les traitements de haut niveau. Les traitements de bas niveau impliquent des opérations primitives telles que le prétraitement de l'image pour réduire le bruit, amélioration du contraste, ou accentuation de l'image. Le traitement de bas niveau est caractérisé par le fait que l'entrée et la sortie sont des images. Les traitements de moyen niveau sur les images impliquent des tâches telles que la segmentation, et la classification. Ils sont caractérisés par le fait que leurs entrées sont généralement des images, mais leurs résultats sont des attributs extraits de ces images tels que les contours ou les points d'intérêts. Le traitement de haut niveau comprend l'interprétation de l'ensemble des attributs de l'image, et leur analyse tout en exploitant les fonctions cognitives normalement associés à la vision.

Les caméras catadioptriques offrent un large champ de vision, cependant les images acquises avec ce genre de capteurs présentent des spécificités géométriques différentes de celles des images perspectives acquises avec des caméras classiques. Le processus de formation des images catadioptriques par réflexion sur un miroir convexe introduit de fortes distorsions radiales. Il est bien connu aussi que ces images ne sont pas uniformément échantillonnées ; la quantité d'informations contenues dans chaque pixel n'est pas la même ; plus proche est le pixel du centre de l'image, plus grande est la quantité d'informations qu'il contient. Une autre particularité qui caractérise les images omnidirectionnelles est que la disposition spatiale du voisinage d'un pixel n'est pas la même que celle des images perspectives.

Dans la littérature les approches de traitement des images omnidirectionnelles peuvent être classées en deux catégories : l'utilisation des outils de traitements classiques développés pour les images perspectives et les appliquer directement aux images omnidirectionnelles, ou bien l'adaptation des opérateurs de traitement à la géométrie des images omnidirectionnelles.

La performance des algorithmes adaptés est remarquable par rapport à l'utilisation des méthodes classiques, pour cela nous nous sommes orientés vers le

développement d'outils de traitement spécialement adaptés aux images catadioptriques.

Dans ce chapitre nous présentons les différentes approches adoptées dans le traitement des images omnidirectionnelles. Un premier algorithme permettant le déballage de ces images est proposé. Par la suite, le modèle de projection unifié, et la projection stéréographiques sont introduits. Ces deux outils ont été utilisés pour développer deux algorithmes d'extractions de primitives, adaptés à la géométrie de formation des images sphériques. Il s'agit d'un opérateur de détection de contours, basé sur un modèle sphérique de charges électriques virtuelles. Ainsi qu'une adaptation de l'algorithme de détection des coins de Harris aux images catadioptriques.

3.2 Application des traitements classiques aux images omnidirectionnelles

De nombreux chercheurs dans le domaine de la vision à large champ, ont vu que les images omnidirectionnelles peuvent être traitées simplement en utilisant les outils classiques de traitements des images perspectives. Cela permet à la fois d'optimiser le temps d'exécution des algorithmes, et de simplifier leur complexité. Mais il est à noter que les caméras catadioptriques interceptent la scène sous une forme polaire (disque). Il est généralement nécessaire de transformer ces images en une forme plus compréhensible pour le système de perception visuelle humain, afin qu'elles puissent être traitées ou visualisées. L'opération de transformation de l'image catadioptrique en une image panoramique est connue sous le nom : déballage (*unwrapping*) [93] [94] [95] [96], illustrée sur la figure 3.1.

Grzegorz Cielniak, et al., ont proposé dans [97] une méthode de suivi des personnes avec camera omnidirectionnelle unique. Cette méthode est basée sur la modélisation automatique de l'aspect d'une personne et le filtre de Kalman pour estimer la position de la personne. Des indices de profondeur sont d'abord extraits des images panoramiques, puis un réseau neuronal artificiel est formé pour estimer la distance entre la personne et la caméra. L'opération du déballage était le premier prétraitement à appliquer sur l'image omnidirectionnelle, pour qu'elle soit ensuite traitée comme une image perspective.

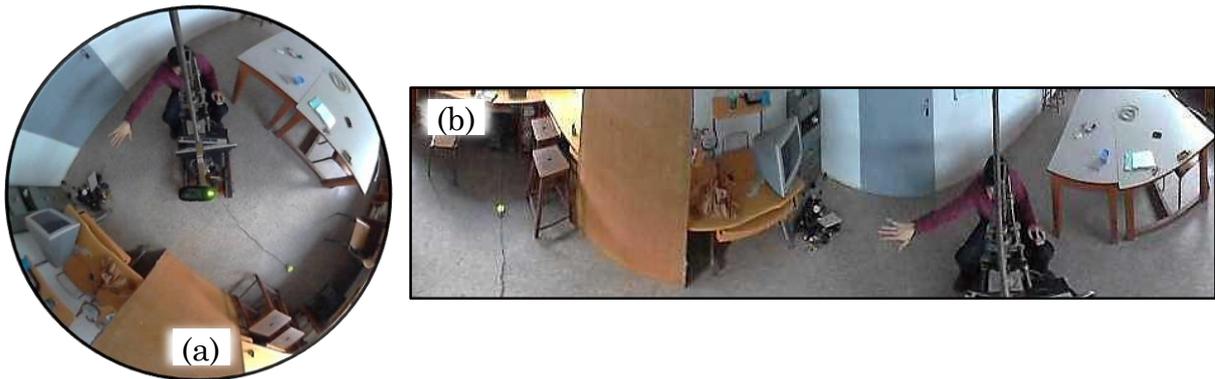


Figure 3.1 : Déballage d'une image catadioptrique (a) image catadioptrique. (b) l'image panoramique correspondante.

Dans [98] Nikos Vlassis et al., ont utilisé le filtre de Sobel pour extraire les contours à partir des séquences d'images omnidirectionnelles, le filtre a été appliqué directement sur l'image, sans déballage. Ces contours sont utilisés comme des attributs, dans un algorithme de localisation d'un robot mobile.

De notre côté, nous nous sommes intéressé initialement par l'évaluation des performances des algorithmes de traitements classiques appliqués sur les images catadioptriques sans adaptation. Pour cela nous avons développé un algorithme de transformation directe d'images catadioptriques en images panoramiques.

3.3 Déballage des images omnidirectionnelles

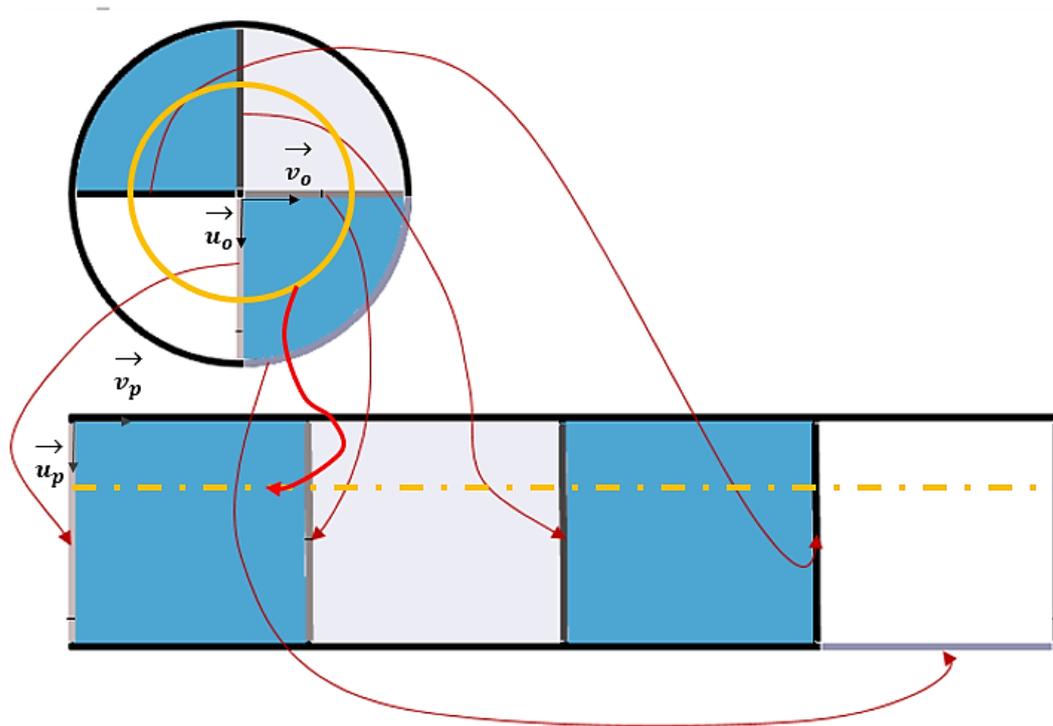


Figure 3.2 : Principe du déballage des images omnidirectionnelles.

L'image catadioptrique peut être convertie en une image panoramique pour qu'elle puisse être traitée. Certaines applications telles que la vidéosurveillance ou la visio-conférence l'exige, vu que ce type d'image sous sa forme brute, ne peut pas être interprété facilement par l'œil. Le déballage des images omnidirectionnelles, consiste en une opération de remappage des pixels de l'image originale en représentation panoramique (figure 3.2).

La procédure de déballage que nous proposons [99] se déroule en quatre étapes :

- La première étape consiste en une translation du repère associé à l'image catadioptrique vers son centre (figure 3.3). Chaque pixel était initialement repéré par ses coordonnées (x, y) , après ce changement de repère, il sera désormais repéré par ses nouvelles coordonnées (x', y') , avec :

$$(x', y') = (x - x_c, y - y_c) \quad 3.1$$

x_c et y_c étant les coordonnées du centre de l'image.

- Les coordonnées polaires (ρ, θ) sont calculées pour chaque pixel :

$$(\rho, \theta) = (\sqrt{x'^2 + y'^2}, \text{atan2}(y'/x')) \quad 3.2$$

La fonction *atan2* (*Four-quadrant inverse tangent*) est une variante de la fonction arc tangente, elle est définie de \mathbb{R} vers $]-\pi, \pi[$. Les valeurs de retour d'*atan2* peuvent être translatées vers l'intervalle $]0, 2\pi[$ en ajoutant 2π aux valeurs négatives.

- Les pixels de l'image omnidirectionnelle sont ensuite projetés sur l'image panoramique, leurs coordonnées (x_p, y_p) dans l'image panoramique sont définies comme suit :

$$(x_p, y_p) = (\alpha(R_{max} - \rho), \beta(\theta \cdot R_{max})) \quad 3.3$$

Avec R_{max} le rayon extérieur de l'image catadioptrique. α et β sont deux coefficients qui permettent d'agir sur la taille et l'aspect ratio de l'image panoramique. Leur valeur par défaut est égale à 1, dans ce cas de figure la taille de l'image de sortie est maximale : $(R_{max}, 2\pi R_{max})$.

Cette projection introduit de fortes discontinuités (des trous) dans l'image panoramique (figure 3.4), cela est dû à deux raisons : l'équation 3.3 affecte des valeurs réelles aux coordonnées (x_p, y_p) , qui subissent par la suite une conversion explicite (cast) vers le type entier. Cela contribue d'une part à la création des discontinuités visibles sur l'image de sortie. D'une autre part si on assimile l'image catadioptrique par un empilement de cercles concentriques avec des rayons qui varient de 0 à R_{max} , l'opération de déballage étire les cercles avec des rayons inférieurs à R_{max} pour qu'ils remplissent la longueur de l'image panoramique, exemple du cercle jaune sur la figure 3.2. Cet étirement est une autre cause spontanée de ces discontinuités. Il est à noter que la réduction des valeurs de α et β allège la gravité du problème des discontinuités introduite dans l'image résultante, mais agit aussi sur sa résolution. Ce problème peut être résolu par

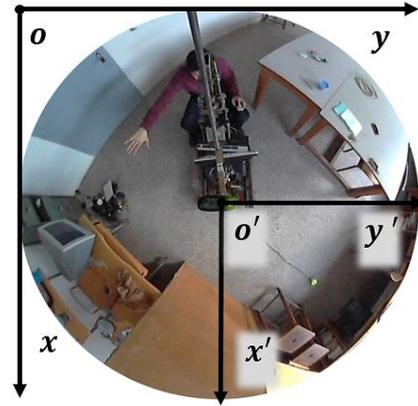


Figure 3.3: changement du repère de l'image catadioptrique.

application d'un filtre médian avec réduction des coefficients α et β mais cette solution est peu efficace.



Figure 3.4: Image panoramique non interpolée.

- Finalement le processus de déballage doit être suivi par une étape d'interpolation, cela permet de résoudre le problème des discontinuités plus efficacement. L'interpolation permet de remplir les pixels qui n'ont subi aucune affectation de valeur (les trous) par les valeurs des pixels les entourant. Plusieurs types d'interpolation sont envisageables, mais nous avons opté pour une interpolation par plus proche voisin, qui est plus favorable vis-à-vis de la complexité et du temps d'exécution.

L'approche proposée avait été évalué expérimentalement notamment en regard du temps d'exécution qui est un facteur déterminant, si l'extension de l'algorithme à des applications temps réel est visée, le déballage des séquences vidéo à titre d'exemple. Nous avons testé cet algorithme sur différentes images catadioptriques, l'expérimentation a été réalisée sur un ordinateur avec les caractéristiques suivantes :

Tableau 3-1 : Caractéristiques de la plateforme d'expérimentation.

| | |
|---|----------------------------------|
| Environnement de développement | Matlab 7.10.0.499 (R2010a 32bit) |
| Processor | Intel Celeron CPU 2.26 GHz |
| RAM | 2Gb DDR2 |
| Système d'exploitation | Ubuntu 10.10 |
| Temps de remise à zéro (Redémarrage de l'ordinateur) | ~35s |

Les durées de temps nécessaires pour l'exécution du déballage d'une image omnidirectionnelle (figure 3.5) à différentes résolutions sont rapportées sur le tableau 3-2.

Tableau 3-2 : Temps d'exécution de l'algorithme du déballage

| La résolution de l'image d'entrée | Les facteurs d'échelle (α, β) | La résolution de l'image de sortie | Temps d'exécution |
|--|--|---|--------------------------|
| 64 X 64 | (1, 0.5) | 17 X 100 | 16,021 ms |
| 128 X 128 | (1, 0.5) | 35 X 204 | 62,509 ms |
| 256 X 256 | (1, 0.5) | 67 X 409 | 240,115 ms |
| 473 X 482 | (1, 0.5) | 130 X 769 | 944,725 ms |
| 1683 X 1689 | (1, 0.5) | 669 X 2642 | 19.781 s |

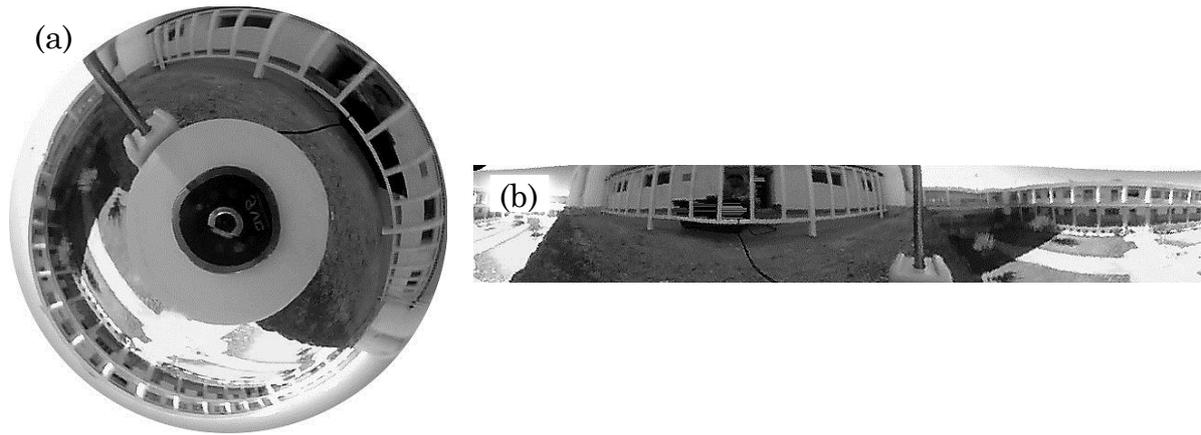


Figure 3.5: (a) Image omnidirectionnelle utilisée dans l'expérimentation du déballage.
 (b) Image panoramique correspondante, résolution : 122 X 749, $\alpha=1$, $\beta=0.5$

D'après ces résultats on peut déduire que le temps nécessaire pour le traitement de chaque pixel de l'image catadioptrique, lors de l'opération du déballage, varie légèrement entre 0.0037ms et 0.007ms en fonction de la taille de l'image d'entrée. Pour la transformation d'une image ou d'un ensemble d'images omnidirectionnelles, la performance de l'algorithme développé est satisfaisante. Néanmoins son applicabilité directe sur un flux vidéo demeure limitée par certaines contraintes. Parmi lesquelles, la résolution des images de la séquence vidéo, ainsi que le nombre de frames (images) par seconde. L'obligation à respecter est la suivante : le temps nécessaire pour le déballage de chaque image ne doit pas dépasser le temps de rafraîchissement du flux vidéo.

L'efficacité de cet algorithme dépend évidemment des caractéristiques de la plateforme matérielle d'implémentation (tableau 3-1), elle peut grandement être améliorée une fois transplantée vers une plateforme dédiée au traitement en temps réel. Des ordinateurs avec des architectures supportant le traitement parallèle de l'information, ou des circuits logiques programmables (*Field-Programmable Gate Array*) peuvent être envisagés en perspectives.

3.4 Adaptation des traitements

D'autres chercheurs ont vu que les particularités géométriques des images catadioptriques, ne peuvent pas être négligées lors du traitement, et que l'adaptation de ces traitements est inévitable. En effet, lorsque les opérateurs de traitements classiques conçus pour des images perspectives, sont appliqués directement sur les images omnidirectionnelles, le problème majeur rencontré est celui du voisinage pixellique. Ce problème peut être bien illustré, si on considère le cas simple de la convolution d'une image omnidirectionnelle avec un masque quelconque. Sur la figure 3.6.a on remarque que lorsqu'on convolue l'image avec un filtre selon les deux axes OX et OY , le voisinage spatiale du pixel central P_0 représenté par les pixels $[P_1 \dots P_8]$ est mal positionné. Pour corriger cette situation, le balayage de l'image peut se faire selon les vecteurs unitaires d'un repère polaire

$(O, \vec{e}_\rho, \vec{e}_\theta)$ situé au centre de l'image omnidirectionnelle. Ou bien, considérer que cette image est équivalente à une représentation sphérique de la scène, et de la traiter dans un espace qui lui convient davantage. Comme illustré sur la figure 3.6.b, le balayage de l'image selon un repère sphérique $(O, \vec{e}_\theta, \vec{e}_\varphi)$ résout naturellement le problème du voisinage pixellique. Nous allons voir ultérieurement dans ce chapitre le fondement mathématique de l'équivalence entre image catadioptrique et image sphérique.

De nombreux auteurs ont souligné l'incompatibilité des traitements classiques avec les images omnidirectionnelles, des solutions ont été proposées dans la littérature pour remédier à ce problème.

Dans [100] on propose de redéfinir le voisinage pixellique on le projetant sur un cylindre, et d'utiliser des opérateurs de traitement à noyaux flous. Cette solution avait été adoptée dans [101] pour détecter les contours dans les images omnidirectionnelles.

D'autres auteurs ont opté pour l'équivalence sphérique comme solution plus adéquate à l'aspect géométrique de ce type d'image. Dans [102] Javier Cruz-Mota et al., ont présenté un algorithme de détection d'attributs de type SIFT (*Scale Invariant Feature Transform*) en coordonnées sphériques. La méthode proposée permet de générer deux types de descripteurs ; des descripteurs locaux sphériques et des descripteurs locaux planaires, qui sont utilisés pour faire la mise en correspondance entre les images sphériques, ou entre une image sphérique et une autre planaire (panoramique). Les résultats expérimentaux de cette étude ont

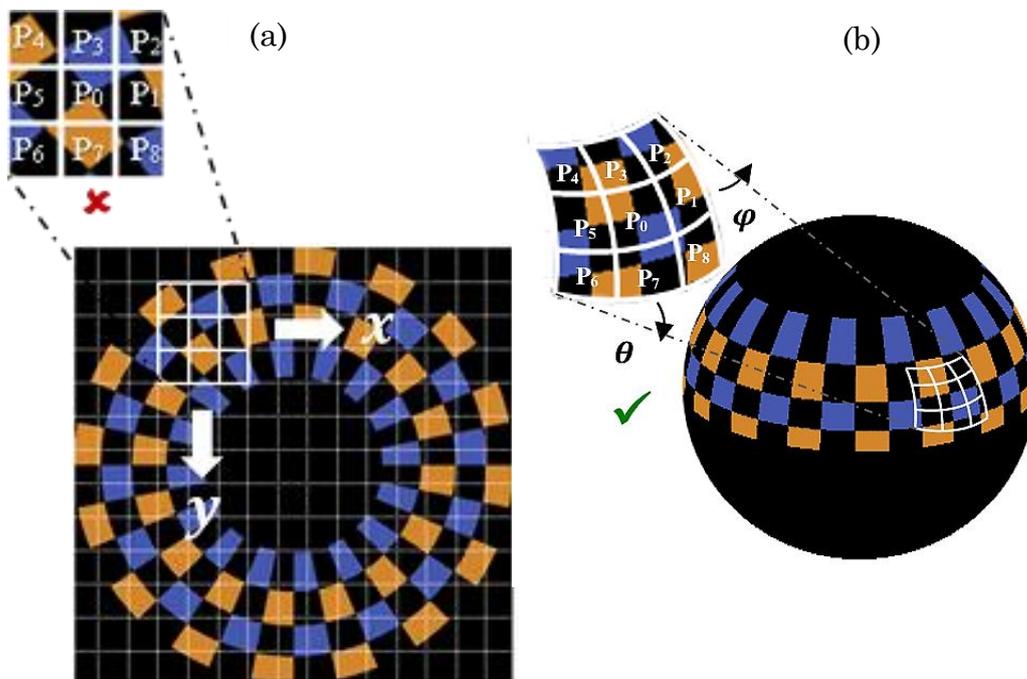


Figure 3.6 : Problème du voisinage pixellique. (a) Voisinage pixellique incorrecte dans le cas d'un filtre planaire. (b) Voisinage correcte dans le cas d'un filtre sphérique.

prouvé la performance de l'algorithme adapté, en comparaison avec un algorithme SIFT standard.

3.4.1 Un modèle de projection unifié

L'équivalence sphérique a été introduite par Geyer et Daniilidis dans [103]. Dans cet article les auteurs ont présenté un modèle de projection unifié valable pour tout type de caméra catadioptrique à centre de projection unique. Selon ce modèle la projection sur la surface d'une quadrique est équivalente à une projection sur la surface d'une sphère S de rayon unitaire (sphère d'équivalence). Ainsi la projection d'un point P de l'espace 3D sur le plan image se décompose en deux projections centrales (figure 3.7) :

- La projection centrale du point $P(X, Y, Z)$ sur la sphère S depuis son centre O en $P'(X', Y', Z')$.
- La projection centrale de P' sur le plan image depuis le point principal F en $p(u, v)$.

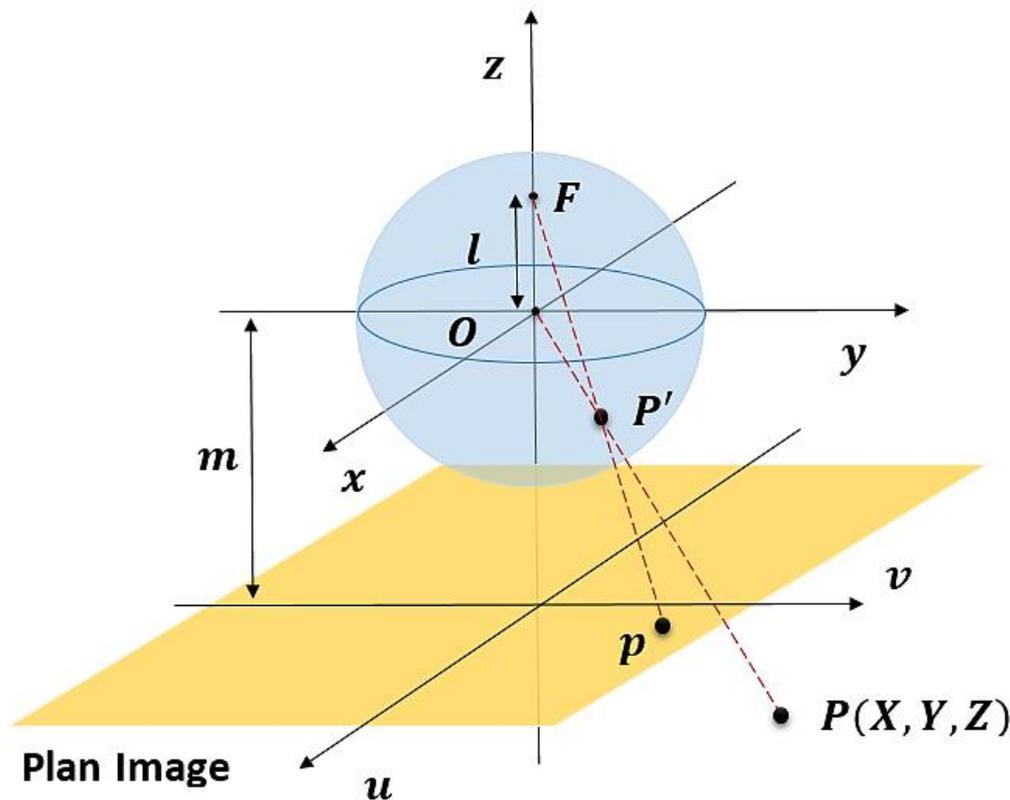


Figure 3.7 : Projection d'un point 3D par le modèle unifié.

Deux paramètres (l et m) sont nécessaires pour établir la correspondance entre un point P de l'espace 3D et sa projection $p(u, v)$ sur l'image catadioptrique. Dans le cas d'un miroir parabolique à titre d'exemple le point F est représenté par le pôle nord de la sphère unitaire.

Le modèle unifié peut être étendu à d'autres capteurs catadioptriques non centraux, par approximation du point de vue unique. Cette approximation a été accomplie pour les caméras à lentille *fish-eye* [104], les caméras catadioptriques à miroirs coniques [105], à miroirs sphériques [106] [107], et autres formes de quadriques spéciaux [108].

Ce modèle a permis surtout de coaliser les travaux de recherche sur l'adaptation des traitements aux images omnidirectionnelles, autour d'un espace de représentation unifié qui est la sphère d'équivalence. En se basant sur ce modèle l'image catadioptrique peut être mappée sur cette sphère via projection centrale inverse qui est équivalente à une projection perspective. Dans le cas singulier d'un système orthographique cette projection est équivalente à une projection stéréographique. Par conséquent, toute une classe d'images omnidirectionnelles peut être traitée par des algorithmes de traitement d'images sphériques.

3.4.2 La projection stéréographique

La 2-sphère ($S^2 \in R^3$) est une variété compacte à courbure constante positive. En coordonnées sphériques, chaque point de la sphère est un vecteur à trois dimensions :

$$\omega = (x_0, x_1, x_2) = (r \cos \theta, r \sin \theta \sin \varphi, r \sin \theta \cos \varphi) \quad 3.4$$

Avec $r \in [0, \infty[$, $\theta \in [0, \pi]$, $\varphi \in [0, 2\pi]$.

La projection stéréographique est une transformation conforme (figure 3.8), qui permet de projeter une sphère sur un plan à partir d'un point de référence fixe de la sphère. Si nous considérons la sphère S^2 de Riemann ($r = 1$) et le plan tangent comme plan complexe C^2 , la projection stéréographique à partir du pôle sud est une bijection définie par :

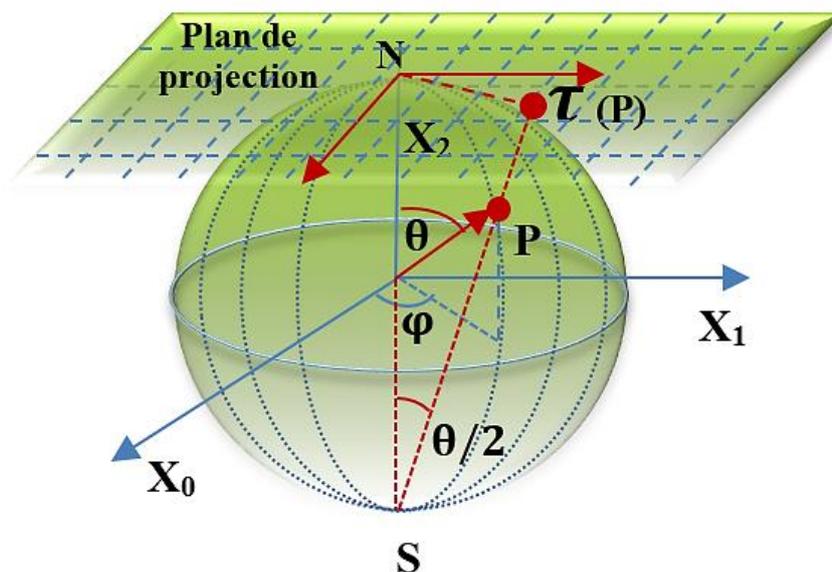


Figure 3.8 : La projection stéréographique.

$$\tau(p) = 2 \tan \frac{\theta}{2} (\cos \varphi; \sin \varphi) \quad 3.5$$

Avec $p = (\theta, \varphi)$, $\theta \in [0, \pi]$, $\varphi \in [0, 2\pi]$.

Les traitements d'image adaptés aux images omnidirectionnelles, que nous allons présenter dans le reste de ce chapitre, sont basés sur l'utilisation de la projection stéréographique inverse pour transformer l'image catadioptrique en image sphérique. Cela permet de traiter ces images tout en respectant les propriétés géométriques de leur formation.

3.4.3 Les opérateurs classiques de détection des contours

L'une des plus importantes habilités du système visuel humain est sa capacité à utiliser les contours ou des bords d'objets pour la reconnaissance et la perception de la distance et de l'orientation. Cela a conduit à une théorie pour le système visuel humain qui est basée sur l'idée que le cortex visuel contient un réseau complexe de détecteurs de caractéristiques qui sont raccordés comme des segments de différentes largeurs et orientations. Pour cette raison, la détection de contours dans une image peut jouer un rôle important dans l'analyse d'une image.

La détection de contours revient à repérer les points d'une image numérique qui indiquent un changement brutal de l'intensité lumineuse. Ces changements de caractéristiques de l'image déchiffrent en général des événements importants ou des changements dans les propriétés de la scène. Ils impliquent les discontinuités de la profondeur, de l'orientation d'une surface, des propriétés d'un matériau ou de l'éclairage de cette scène. La détection de contour est un champ de la recherche qui appartient au traitement d'image et à la vision artificielle, et plus particulièrement au domaine d'extraction des caractéristiques.

La détection des contours d'une image réduit significativement la quantité de données à analyser et élimine les informations les moins pertinentes, tout en préservant les caractéristiques structurelles importantes de l'image.

La détection des contours utilise des opérateurs différentiels pour détecter des changements du gradient des niveaux de gris. Elle est divisée en deux catégories principales : (i) détection de contours du premier ordre, (ii) et la détection de contours du second ordre. Comme son nom l'indique, la détection de contours du premier ordre est basée sur l'utilisation d'une dérivée de premier ordre tandis que la détection de contours du second ordre est basée sur l'utilisation d'une dérivée du

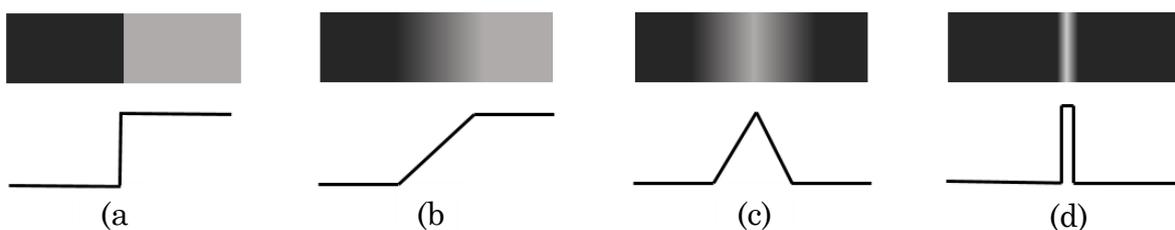


Figure 3.9 : Modèles de contours. (a) Marche d'escalier. (b) Rampe. (c) Toit. (d) Ligne.

second ordre, en particulier, le Laplacien ∇^2 . La première dérivée $\partial/\partial x$ est extrême sur un contour alors que la seconde dérivée $\partial^2/\partial x^2$ s'annule là où le contour est à sa plus grande pente. Dans une image à niveaux de gris différents, quatre modèles de contours peuvent être distingués (figure 3.9). Une « marche d'escalier » si le contour est distinct, une « rampe » si le contour est plus flou, un « toit » s'il s'agit d'une ligne floue sur un fond uniforme, et une « ligne » dans le cas d'une ligne bien distinguée.

3.4.3.1 Détecteurs de contours du premier ordre

La détection des contours du premier ordre est basée sur le calcul du gradient de l'image en x et y , et en observant les emplacements dans l'image où il change brusquement. Si l'image est représentée par la fonction $f(x, y)$, alors l'idée de base est de calculer ce gradient :

$$\nabla f(x, y) = \hat{x} \frac{\partial}{\partial x} f(x, y) + \hat{y} \frac{\partial}{\partial y} f(x, y) \quad 3.6$$

Et puis afficher l'amplitude du gradient défini par :

$$|\nabla f| = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad 3.7$$

Un seuillage peut être ensuite appliqué sur l'image résultante. Seules les valeurs de l'amplitude du gradient dépassants un certain seuil, sont retenues. Une image binaire représentant les contours, est alors obtenue.

Les différentes méthodes basées sur le gradient pour la détection de contours, sont résultantes d'approximations discrètes de ∇f . Les approximations disponibles sont agglomérées dans une classe d'opérateurs appelés **gradients numériques**. Une fonction continue $f(x)$ peut être étendue à point $x = \Delta x$ comme une série de Taylor :

$$f(x + \Delta x) = f(x) + \Delta x \frac{\partial}{\partial x} f(x) + \frac{(\Delta x)^2}{2!} \frac{\partial^2}{\partial x^2} f(x) + \dots \quad 3.8$$

Si nous négligeons tous les termes à partir du deuxième, alors on obtient une approximation de la dérivée de $f(x)$ à Δx , donnée par l'équation de différence :

$$\frac{\partial}{\partial x} f(x) = \frac{f(x+\Delta x) - f(x)}{\Delta x} \quad 3.9$$

Aux dérivées partielles que nous avons :

$$\frac{\partial}{\partial x} f(x, y) = \frac{f(x+\Delta x, y) - f(x, y)}{\Delta x} \quad 3.10$$

Et

$$\frac{\partial}{\partial y} f(x, y) = \frac{f(x, y+\Delta y) - f(x, y)}{\Delta y} \quad 3.11$$

Ainsi avec une image numérique f_{ij} , nous pouvons remplacer les dérivées partielles par les différences :

$$D_x f_{ij} = f_{(i+1)j} - f_{ij} \quad 3.12$$

Et

$$D_y f_{ij} = f_{i(j+1)} - f_{ij} \quad 3.13$$

Ces opérations sont équivalentes à une convolution de f_{ij} avec le noyau $(-1,1)$ dans la direction de x , et avec le même noyau $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ dans la direction y . Les noyaux ou les filtres qui sont utilisés pour la convolution d'une image numérique de cette manière, sont appelés masques. Ils sont des opérateurs invariants par translation, qui nous permettent d'écrire les anciennes équations de différences sous la forme :

$$D_x f_{ij} = D_x \otimes f_{ij} \quad 3.14$$

Et

$$D_y f_{ij} = D_y \otimes f_{ij} \quad 3.15$$

Avec

$$D_x = (-1,1) \text{ et } D_y = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad 3.16$$

Le filtre de Roberts : Les masques examinés jusqu'ici fournissent une première approximation pour le calcul du gradient d'une image dans les directions x et y qui se traduit par une détection des contours horizontaux et verticaux au détriment des contours diagonaux. Le filtre de Roberts est basé sur l'approximation du gradient de premier ordre, où l'amplitude du gradient est donnée par :

$$G_{ij} = \sqrt{(f_{ij} - f_{(i+1)(j+1)})^2 + (f_{(i+1)j} - f_{i(j+1)})^2} \quad 3.17$$

Ce qui revient à appliquer les masques :

$$D_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad 3.18$$

Et :

$$D_y = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad 3.19$$

Toutefois, ces masques opèrent sur des zones de pixels relativement petites, et sont par conséquent plus sensible au bruit. Les applications pratiques de gradients numériques tels que l'opérateur Roberts, nécessitent habituellement du prétraitement pour réduire le bruit inhérent à une image. Une autre approche consiste à considérer les gradients numériques qui opèrent sur une grande matrice de pixels.

Le filtre de Sobel : Le détecteur de contours de Sobel est une extension qui comprend un degré de lissage pour réduire automatiquement certains effets

parasites dus au bruit. Plus grand est le filtre, meilleur est la réduction du bruit, mais quand le filtre devient trop large, des contours utiles peuvent ne pas être détectés. Le filtre de Sobel est basé sur les gradients numériques suivants :

$$D_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, D_y = \begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix} \quad 3.20$$

Le filtre de Prewitt : Le filtre de Prewitt est basé sur les gradients numériques suivants :

$$D_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}, D_y = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix} \quad 3.21$$

Le détecteur de contours de Canny : En appliquant les filtres décrits ci-dessus, deux problèmes se posent. Comme les éléments de données sont discrets, il n'est pas toujours évident de calculer la valeur du gradient, et, en présence de bruit, des contours parasites peuvent être détectés. Le détecteur de contours de Canny est conçu pour lutter contre certains de ces problèmes et se compose de trois phases principales :

- Lissage de l'image avec un filtre gaussien pour réduire la quantité de bruit notamment de type *speckle*. Il est important d'éliminer les composantes à très haute fréquence qui dépassent celles qui sont associées au filtre de gradient utilisé, autrement, ceux-ci peuvent provoquer de fausses détections de contours.
- Calcul du gradient de l'image, cela conduit à la création de deux images, l'une contenant les amplitudes de gradient G , et un autre contenant de l'orientation $\theta(G)$. Les implémentations les plus courantes, utilisent la simple dérivée discrète du premier ordre.
- Seuillage des amplitudes de gradient à partir d'une valeur minimale, de sorte que seuls les contours signifiants sont détectés. Un seuillage avec une valeur maximale est également spécifié. Sur toute une ligne continue, au moins l'un des points du contour doit dépasser cette valeur maximale. Cela supprime les petits ou les insignifiants segments de ligne.

En contrôlant l'écart type de la Gaussienne du filtre de lissage, ainsi que les deux valeurs des seuils haut et bas, la plupart des bords généraux peuvent être détectés. Mais en présence de bruit, l'ajustement adéquat de ces paramètres, n'est pas évident.

3.4.3.2 Détecteurs de contours du second ordre

Les détecteurs de contours du second ordre sont basés sur le calcul de la dérivée seconde en x et y et en observant les emplacements dans l'image où sa valeur passe par zéro. Le Laplacien de l'image est donnée par :

$$\nabla^2 f(x, y) = \frac{\partial^2}{\partial x^2} f(x, y) + \frac{\partial^2}{\partial y^2} f(x, y) \quad 3.22$$

Par extension de la fonction $f(x + \Delta x)$ comme une série de Taylor, et en suivant un développement similaire à celui fait pour la dérivée du premier ordre, l'expression du Laplacien discret devient :

$$\nabla^2_{ij} f_{ij} = f_{(i+1)j} + f_{(i-1)j} + f_{i(j+1)} + f_{i(j-1)} - 4f_{ij} \quad 3.23$$

Cette opération peut s'écrire en termes d'une convolution de f_{ij} par un masque approprié :

$$\nabla^2_{ij} f_{ij} = \text{masque}_{ij} \otimes \otimes f_{ij} \quad 3.24$$

Avec :

$$\text{masque}_{ij} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad 3.25$$

Le principal avantage de la dérivée seconde est qu'elle est indépendante par rapport à l'orientation. Mais plus sensible au bruit.

3.4.4 Détecteur de contours dédié aux images omnidirectionnelles

La détection des contours est un processus local de traitement de l'image, systématiquement lié à la disposition du voisinage de chaque pixel traité. Dans une image perspective, la résolution est uniforme, et les pixels sont disposés selon une topologie cartésienne. Dans une image catadioptrique, la topologie du voisinage pixellique est de forme sphérique. L'application des opérateurs classiques de détection des contours sur une telle image, est géométriquement incorrecte. D'où l'intérêt de développer de nouveaux opérateurs, adaptés à la géométrie de formation des images catadioptriques.

Dans ce même contexte, nous avons présenté dans [109] un algorithme de détection de contours adapté aux images omnidirectionnelles. L'algorithme modélise l'image catadioptrique comme une distribution virtuelle sphérique, de charges électrostatiques en équilibre.

3.4.4.1 Modèle sphérique de charges électriques virtuelles

L'image est fragmentée en blocks de 3x3 pixels (figure 3.10), Nous supposons que les forces exercées par les charges au-delà de la fenêtre 3x3 sur la charge centrale sont négligeables. L'analogie suivante est faite : l'intensité de chaque pixel correspond à la charge électrique associée à ce pixel, et la charge centrale q_0 est soumise aux forces électrostatiques des huit charges voisines. Nous distinguons deux types de forces : attractives et répulsives. Selon la condition d'équilibre électrostatique, nous avons :

$$\sum_{i=1}^n \vec{F}_i = 0 \quad 3.26$$

Considérons deux charges électriques de q_0 et q_1 séparées par une distance r . Un champ électrostatique est présent entre ces deux charges fixes. La charge q_1 génère un champ électrostatique \vec{E}_1 :

$$\vec{E}_1 = \frac{q_1}{4\pi\epsilon_0} \frac{\vec{r}}{r^3} \quad 3.27$$

La force exercée par le champ \vec{E}_1 sur q_0 est $\vec{F} q_1/q_0$. Selon la loi de Coulomb cette force s'exprime comme suit :

$$\vec{F} q_1/q_0 = q_1 \vec{E}_1 = q_1 \frac{q_0}{4\pi\epsilon_0} \frac{\vec{r}}{r^3} \quad 3.28$$

$$\vec{F} q_1/q_0 = k q_1 q_0 \frac{\vec{r}}{r^3}, \quad k = \frac{1}{4\pi\epsilon_0} \quad 3.29$$

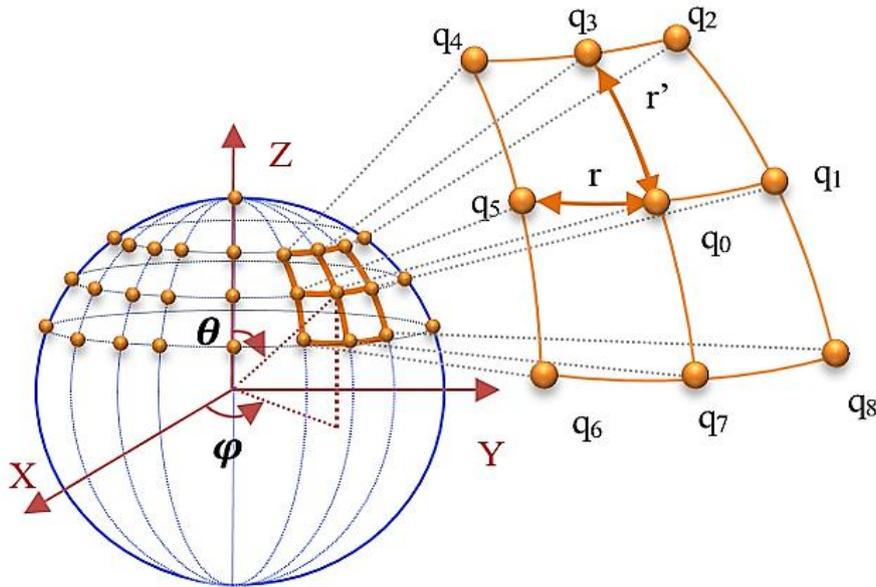


Figure 3.10 : Distribution sphérique de charges électrostatiques.

Dans le cas de forces attractives. Dans un repère sphérique $R(o, \vec{e}_r, \vec{e}_\varphi, \vec{e}_\theta)$, les charges q_i avec $i = 1, 2, \dots, 8$ exercent des forces attractives sur la charge q_0 qui possède une charge électrique de signe opposé. Soit \vec{T}_φ et \vec{L}_θ respectivement, les forces exercées sur q_0 en direction de \vec{e}_φ et \vec{e}_θ :

$$\vec{T}_\varphi = \vec{F} q_1/q_0 + \vec{F} q_5/q_0 = k \frac{q_0}{r^2} (q_5 - q_1) \vec{e}_\varphi \quad 3.30$$

$$\vec{L}_\theta = \vec{F} q_3/q_0 + \vec{F} q_7/q_0 = k \frac{q_0}{r'^2} (q_3 - q_7) \vec{e}_\theta \quad 3.31$$

Avec : $r = R \sin \theta d\varphi$ et $r' = r$ nous aurons :

$$\vec{T}_\varphi = k \frac{q_0}{(R \sin \theta d\varphi)^2} (q_5 - q_1) \vec{e}_\varphi \quad 3.32$$

$$\vec{L}_\theta = k \frac{q_0}{(R \sin \theta d\varphi)^2} (q_3 - q_7) \vec{e}_\theta \quad 3.33$$

Les forces diagonales s'expriment comme suit :

$$\vec{F}D_1 = \vec{F} q_2/q_0 + \vec{F} q_6/q_0 = k \frac{q_0}{\delta^2} (q_6 - q_2) \vec{D}_1 \quad 3.34$$

$$\vec{F}D_2 = \vec{F} q_8/q_0 + \vec{F} q_4/q_0 = k \frac{q_0}{\delta^2} (q_4 - q_8) \vec{D}_2 \quad 3.35$$

$$\vec{D}_1 = \cos(\lambda) \vec{e}_\varphi - \sin(\lambda) \vec{e}_\theta \quad 3.36$$

$$\vec{D}_2 = \cos(\lambda) \vec{e}_\varphi + \sin(\lambda) \vec{e}_\theta \quad 3.37$$

Avec $\lambda = \pi/4$ et $\delta^2 = 2r^2 = 2(R \sin \theta d\varphi)^2$

Les forces diagonales s'écrivent :

$$\vec{F}D_1 = k \frac{q_0 \sqrt{2}}{4(R \sin \theta d\varphi)^2} (q_6 - q_2) (\vec{e}_\varphi - \vec{e}_\theta) \quad 3.38$$

$$\vec{F}D_2 = k \frac{q_0 \sqrt{2}}{4(R \sin \theta d\varphi)^2} (q_4 - q_8) (\vec{e}_\varphi + \vec{e}_\theta) \quad 3.39$$

Soit $k' = k/R^2$

Et $\vec{F}_\varphi = F_\varphi \vec{e}_\varphi$, $\vec{F}_\theta = F_\theta \vec{e}_\theta$

Les équations (4), (5), (6), (7) deviennent :

$$\vec{T}_\varphi = k' \frac{q_0}{(\sin(\theta) d\varphi)^2} (q_5 - q_1) \vec{e}_\varphi \quad 3.40$$

$$\vec{L}_\theta = k' \frac{q_0}{(\sin(\theta) d\varphi)^2} (q_3 - q_7) \vec{e}_\theta \quad 3.41$$

$$\vec{F}D_1 = k' \frac{q_0 \sqrt{2}}{4(\sin(\theta) d\varphi)^2} (q_6 - q_2) (\vec{e}_\varphi - \vec{e}_\theta) \quad 3.42$$

$$\vec{F}D_2 = k' \frac{q_0 \sqrt{2}}{4(\sin(\theta) d\varphi)^2} (q_4 - q_8) (\vec{e}_\varphi + \vec{e}_\theta) \quad 3.43$$

Finalement :

$$\vec{F}_\varphi = k' q_0 \frac{1}{(\sin(\theta) d\varphi)^2} [q_5 - q_1 + \frac{\sqrt{2}}{4} (q_6 - q_2 + q_4 - q_8)] \vec{e}_\varphi \quad 3.44$$

$$\vec{F}_\theta = k' q_0 \frac{1}{(\sin(\theta) d\varphi)^2} [q_3 - q_7 + \frac{\sqrt{2}}{4} (q_2 - q_6 + q_4 - q_8)] \vec{e}_\theta \quad 3.45$$

Si q_0 correspond au pixel central de la fenêtre 3×3 étudiée, nous obtenons les deux filtres G_φ et G_θ en direction de φ et θ définis ainsi :

$$G_\varphi = 1/(\sin(\theta) d\varphi)^2 \begin{pmatrix} \sqrt{2}/4 & 0 & -\sqrt{2}/4 \\ 1 & 0 & -1 \\ \sqrt{2}/4 & 0 & -\sqrt{2}/4 \end{pmatrix} \quad 3.46$$

Et :

$$G_\theta = 1/(\sin(\theta) d\varphi)^2 \begin{pmatrix} \sqrt{2}/4 & 1 & \sqrt{2}/4 \\ 0 & 0 & 0 \\ -\sqrt{2}/4 & -1 & -\sqrt{2}/4 \end{pmatrix} \quad 3.47$$

G_φ et G_θ représentent deux masques de convolution permettant la détection des contours, l'un dans la direction de \vec{e}_φ et l'autre dans la direction de \vec{e}_θ , applicables sur une image sphérique. Ce même raisonnement fait dans le cas de forces attractives, peut être fait dans le cas de forces répulsives. Dans ce cas on suppose que toutes les charges électriques sont de même signe. Nous obtiendrons deux masques G'_φ et G'_θ avec :

$$G'_\varphi = -G_\varphi \text{ et } G'_\theta = -G_\theta$$

Pratiquement les deux masques G'_φ et G'_θ sont équivalents à G_φ et G_θ . Car le raisonnement dans le cas de forces répulsives, est équivalent à inverser les sens des deux vecteurs \vec{e}_φ et \vec{e}_θ .

3.4.4.2 Algorithme de détection des contours dans les images sphériques

Nous avons vu dans les sections précédentes, que les filtres conventionnels ne sont pas adaptés à la géométrie des images catadioptriques. L'un des moyens disponibles pour surmonter cette anomalie, est de mapper l'image catadioptrique sur la sphère d'équivalence qui est un espace conforme à la géométrie de formation de ce type d'images (figure 3.11).

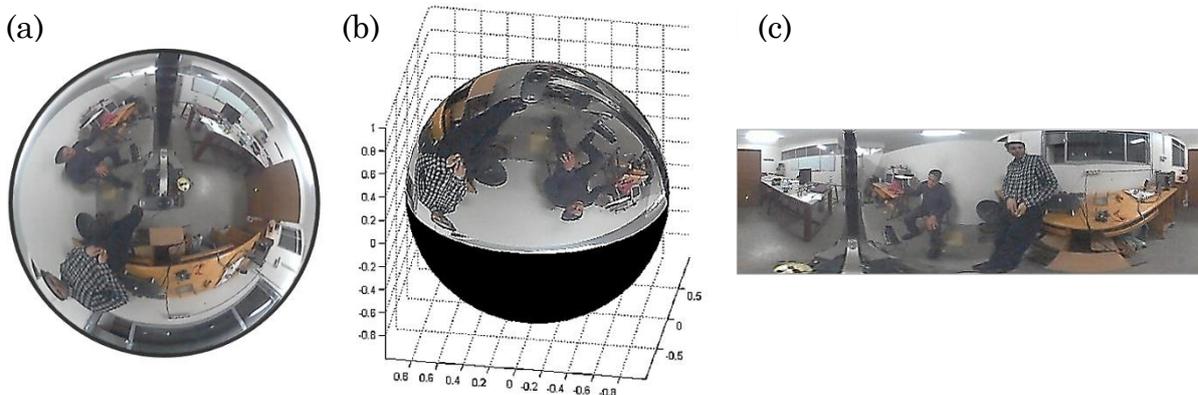


Figure 3.11 : (a) Image catadioptrique. (b) Image mappée sur la sphère. (c) L'image panoramique correspondante.

L'algorithme de détection de contours que nous proposons est basé sur ce fondement. L'image catadioptrique à traiter est initialement mappée sur la sphère via projection stéréographique inverse, une image sphérique est ainsi obtenue. Nous appliquons ensuite les masques de détection de contours G_φ et G_θ sur cette image. Comme pour tout algorithme de détection de contours, une étape de seuillage des valeurs obtenues est nécessaire. Cela permet d'éliminer une grande partie des fausses détections, et d'obtenir une image binaire qui réduit l'image originale en un ensemble de contours séparant les régions de cette image. Cet algorithme est schématisé dans la figure 3.12 ci-dessous.

Tout algorithme de détection de contours ne peut être validé, qu'après une étude expérimentale qui permettra d'évaluer ses performances, notamment en présence de bruit. Nous proposons d'estimer quantitativement la robustesse de l'algorithme présenté vis-à-vis du bruit en utilisant la méthode d'évaluation de Fram & Deutsch. Les expériences ont été menées des images synthétiques, ainsi que sur des images réelles. L'algorithme a été confronté aux opérateurs classiques de détection des contours.

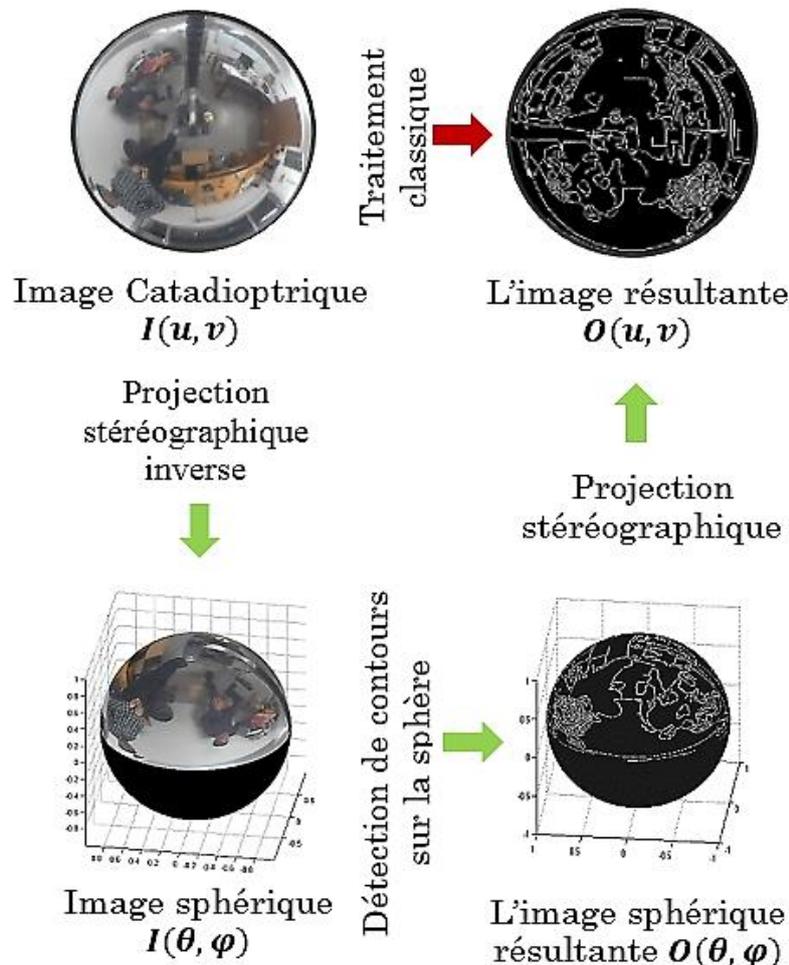


Figure 3.12 : Algorithme de détection de contours dans l'image sphérique.

3.4.4.3 La méthode d'évaluation de Fram & Deutsch

En 1975, Fram et Deutsch ont présenté dans [110] une méthode d'évaluation quantitative de la performance des détecteurs de contours. La robustesse de détection est évaluée par calcul de deux paramètres P_1 et P_2 qui prennent des valeurs dans l'intervalle $[0,1]$. Le premier paramètre reflète la distribution des vrais points détectés du contour en contraste avec les faux points qui résultent du bruit. La seconde mesure peut être considérée comme une estimation du maximum de vraisemblance du rapport entre le nombre total de vrais points du contour étudié, et le nombre total de points détectés. P_1 et P_2 sont égaux à 1 pour une détection parfaite du contour. Leurs expressions peuvent maintenant être données. Soit R la région du contour prise en considération (figure 3.13). R renferme R_{in} comme région intérieure du contour, et R_{out} comme région extérieure. Ainsi n_{tot} dénote le nombre total de points de R . n_{in} est le nombre de points de R_{in} , et n_{out} est le nombre de points de R_{out} avec $n_{tot} = n_{in} + n_{out}$. n^0 indique le nombre de points détectés au seuil déterminé dans R_{out} et n^e comme définit par Fram et Deutsch dans [111] est le nombre de points détectés dans R_{in} . Soit w_1^e le nombre de colonnes contenues dans R_{in} et w_1^{stan} le nombre de colonnes contenues dans R . Finalement, on définit n^r comme nombre de lignes de R_{in} , et w_2 le nombre de lignes de R . Les deux paramètres P_1 et P_2 sont définis comme suit :

$$P_1 = \frac{n_{sig}^e}{n_{sig}^e + (n_{noise}^e + n^0) \frac{n_{in}}{n_{tot}}} \quad 3.48$$

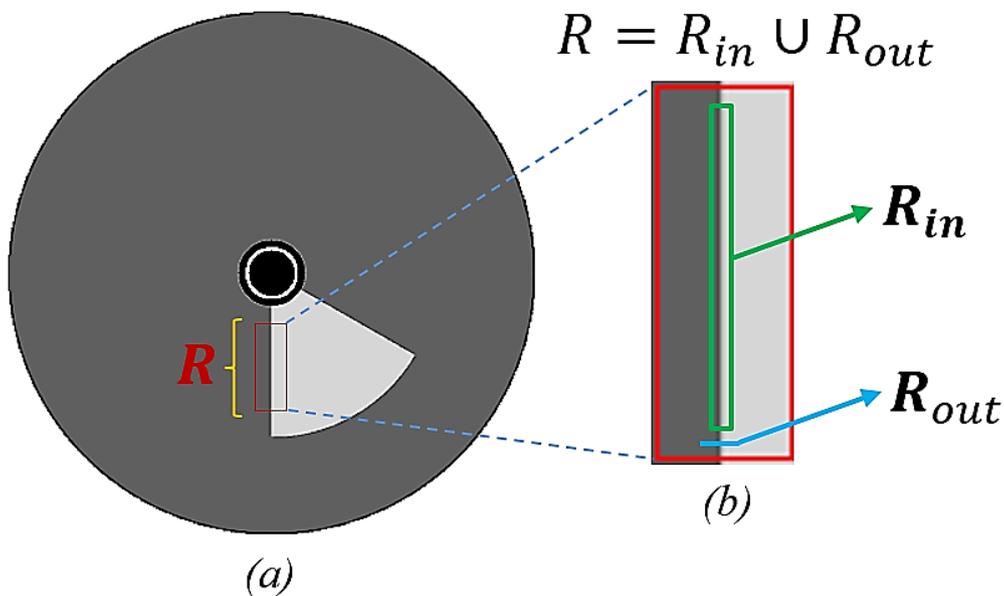


Figure 3.13 : (a) L'image synthétique générée pour l'évaluation quantitative de la détection de contours, avec un contour verticale à 0° et un contour diagonale à 60°. (b) La région de contour considérée.

$$\text{Avec : } n_{sig}^e = \frac{n^e - n_{noise}^e}{1 - \frac{n_{noise}^e}{n_{in}}}, n_{noise}^e = \frac{n^0 n_{in}}{n_{out}} \text{ et } f = \frac{w_1^e}{w_{stan}^e}$$

$$P_2 = \frac{\frac{n_r}{w_2} \left\{ 1 - \left[1 - \frac{n_{noise}^e}{n_{in}} \right]^{w_1^e} \right\}}{\left[1 - \frac{n_{noise}^e}{n_{in}} \right]^{w_1^e}} \quad 3.49$$

3.4.4.4 Résultats expérimentaux

Le jeu d'images de synthèses utilisé pour évaluer l'algorithme proposé se compose de trois images artificielles avec deux types de contours, le premier est vertical, et le second est un contour diagonal incliné d'un angle de 60 °. La première image a un pas de 50 niveaux de gris entre les deux régions considérées, le pas de la deuxième est de 100 niveaux et le troisième est de 150 niveaux. Un bruit gaussien avec une variance croissante progressivement σ est ajouté d'abord à ces images afin d'évaluer le comportement de chaque opérateur étudié, en présence du bruit. Afin de mettre en œuvre la méthode d'évaluation de Fram et Deutsch, un processus de seuillage est nécessaire après l'application des masques de détection des contours. Le paramètre P_1 est fortement corrélé à la valeur du seuil choisi, donc nous avons opté pour adopter le seuil offrant la meilleure valeur de P_1 pour chaque image traitée. Les opérateurs classiques de Sobel, Prewitt et le modèle planaire de charges électriques virtuelles (PMCEV), ont été appliqués directement sur les images catadioptriques planaires. Par contre l'algorithme de détection de contours basé sur le model sphérique de charges électriques virtuelles (SMCEV) a été appliqué sur les images catadioptriques sphériques.

Les graphiques ci-dessous illustrent l'étude comparative de la robustesse de chaque opérateur vis-à-vis d'un bruit gaussien de variance croissante. Les figures 3.14, 3.15, et 3.16 représentent la situation d'un contour vertical, avec un saut de 50, 100 et 150 niveaux de gris entre les deux régions considérées.

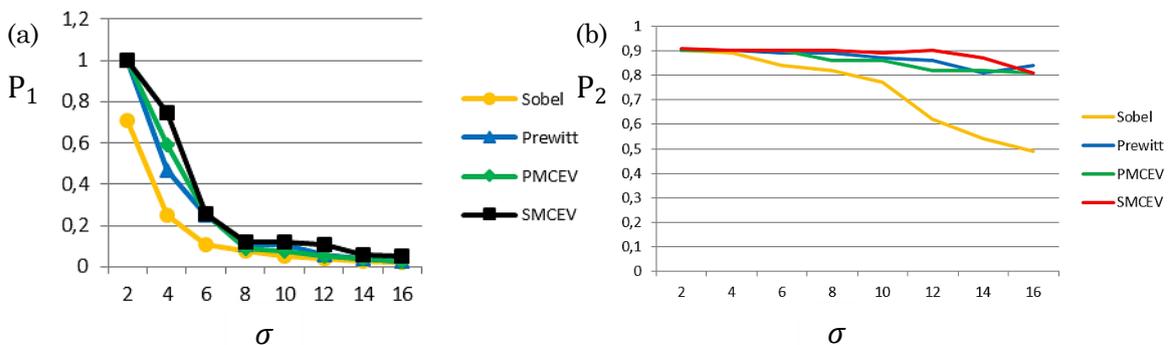


Figure 3.14 : (a) Valeurs de P_1 avec bruit gaussien de variance croissante (contour vertical, un saut de 50 niveaux de gris). (b) Valeurs de P_2 .

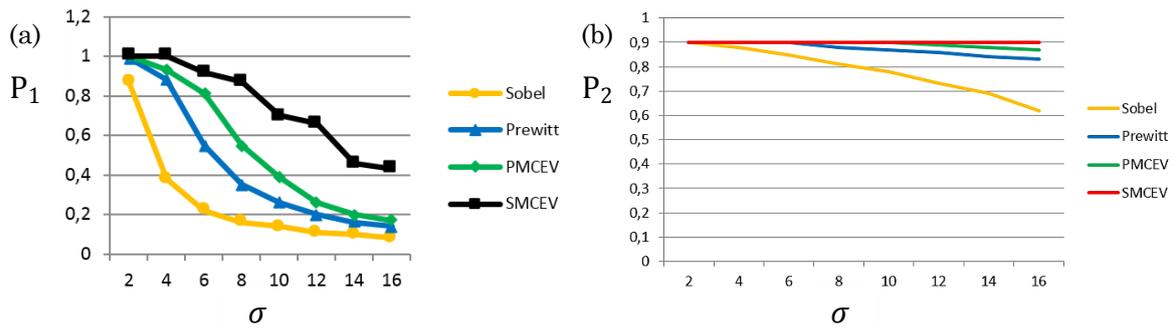


Figure 3.15 : (a) Valeurs de P_1 avec bruit gaussien de variance croissante (contour vertical, un saut de 100 niveaux de gris). (b) Valeurs de P_2 .

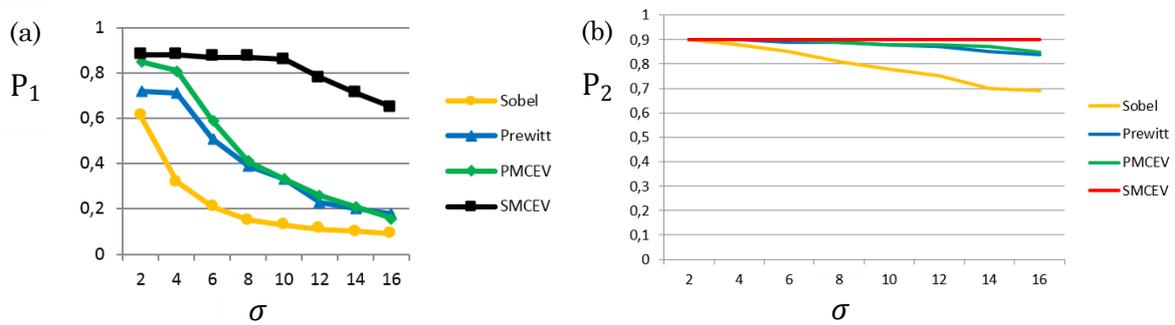


Figure 3.16 : (a) Valeurs de P_1 avec bruit gaussien de variance croissante (contour diagonal, un saut de 150 niveaux de gris). (b) Valeurs de P_2 .

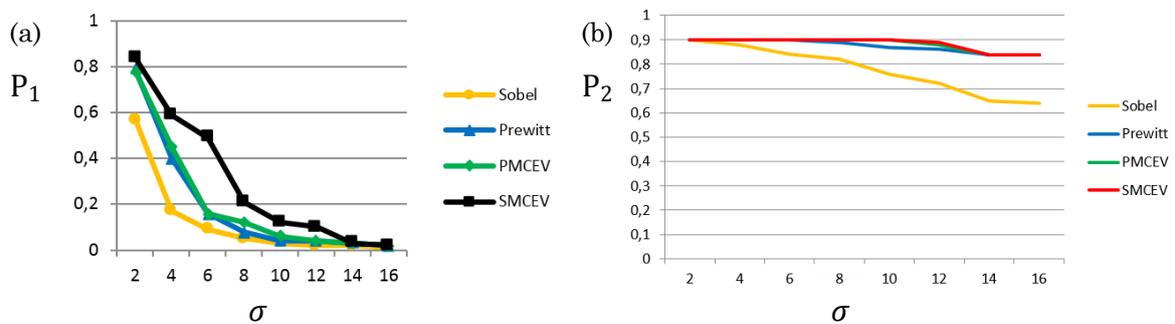


Figure 3.17 : (a) Valeurs de P_1 avec bruit gaussien de variance croissante (contour diagonal, un saut de 50 niveaux de gris). (b) Valeurs de P_2 .

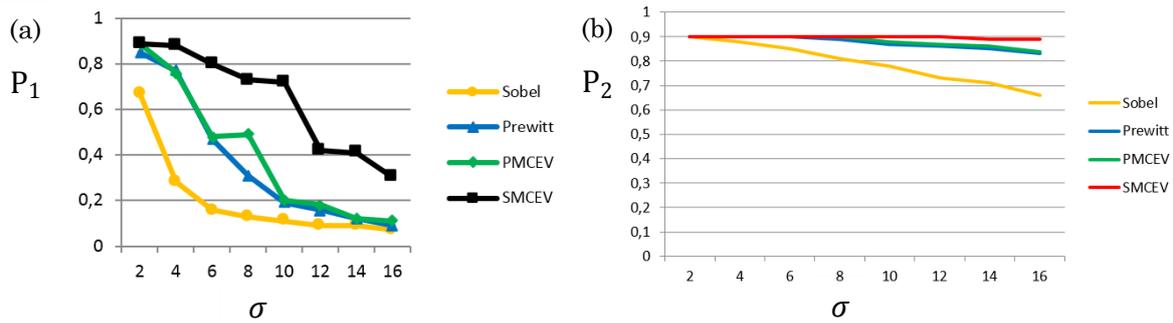


Figure 3.18 : (a) Valeurs de P_1 avec bruit gaussien de variance croissante (contour diagonal, un saut de 100 niveaux de gris). (b) Valeurs de P_2 .

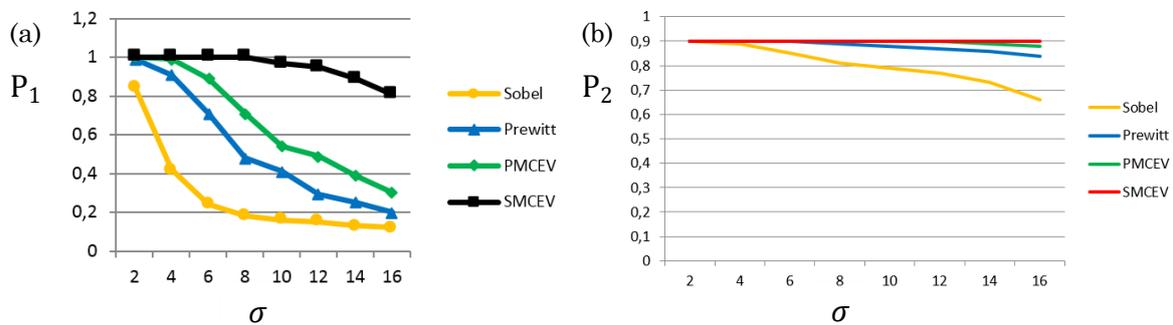


Figure 3.19 : (a) Valeurs de P_1 avec bruit gaussien de variance croissante (contour vertical, un saut de 150 niveaux de gris). (b) Valeurs de P_2 .

Comme P_1 mesure la performance de chacun des opérateurs de détection de contours en présence du bruit, les résultats ci-dessus montrent que dans les deux situations de contours verticaux et obliques, l'approche SMCEV est moins affectée par le bruit vu qu'elle détecte moins de faux points du contour considéré. On constate que l'opérateur de Sobel fournit les résultats les moins appréciables, avec une grande sensibilité au bruit ; son paramètre P_1 commence à décroître, même à faible variance du bruit gaussien. Alors que P_1 pour les opérateurs PMCEV et Prewitt décroît plus progressivement que celui de Sobel. Ces résultats sont remarquablement plus perceptibles lorsque le pas de niveau de gris entre les deux régions considérées augmente, dans ce cas, l'analyse est plus discriminante. L'efficacité de tous les opérateurs de détection de contours étudiés, se dégrade moyennement tout en opérant sur les contours obliques, contrairement aux contours verticaux. Considérant le paramètre P_2 . La répartition des points détectés sur la longueur des contours, de la méthode proposée est meilleure que celle des autres méthodes comparées. Une image omnidirectionnelle d'un damier (mire de calibration) présentée sur la figure 3.20 a été examinée pour étudier le comportement du détecteur proposé dans des conditions réelles d'utilisation. Les images ci-dessous montrent les résultats obtenus avec chaque détecteur.



Figure 3.20 : Image réelle utilisée pour l'évaluation de la détection de contours.

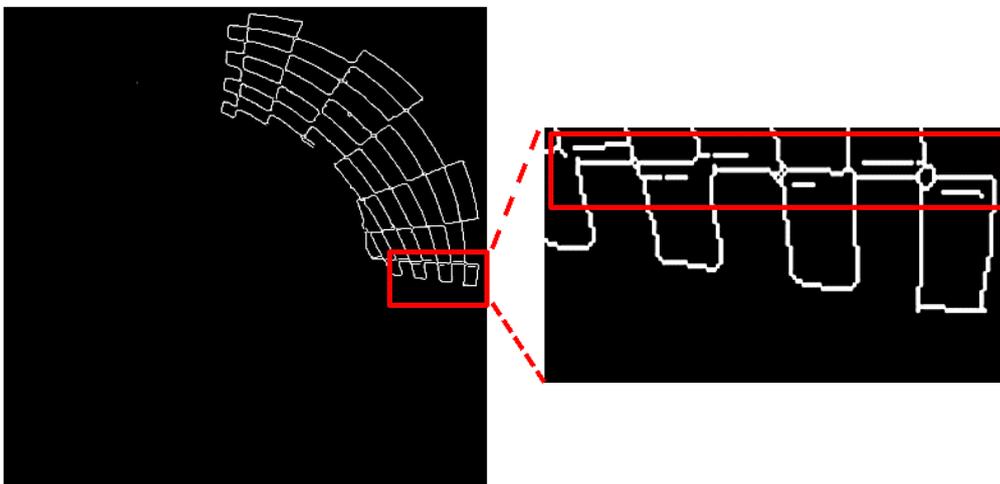


Figure 3.21 : Image réelle filtrée avec l'opérateur de Sobel.

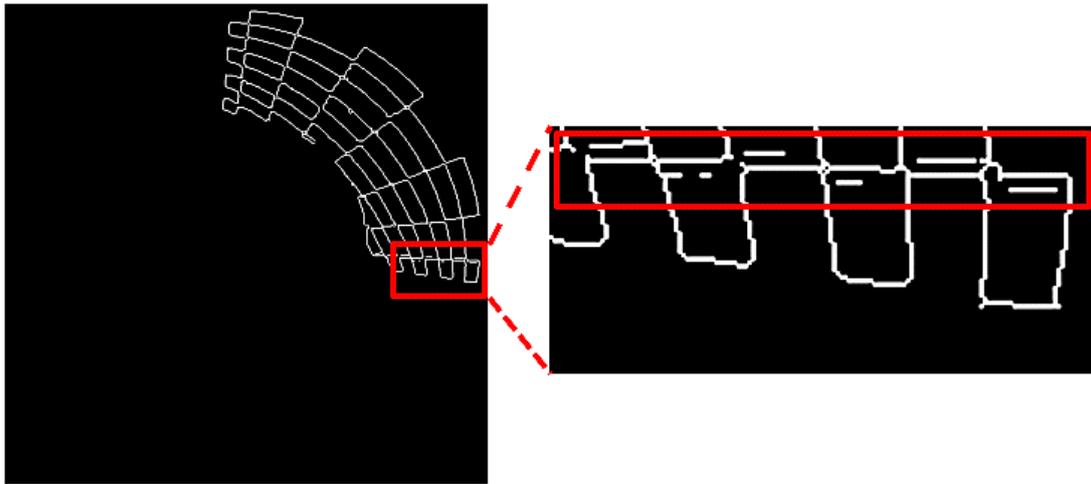


Figure 3.22 : Image réelle filtrée avec l'opérateur de Prewitt.

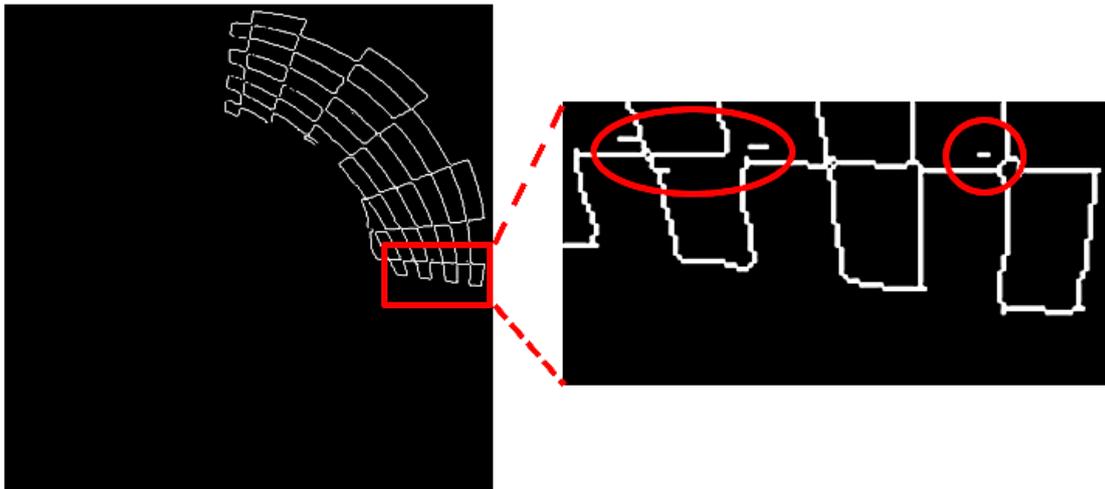


Figure 3.23 : Image réelle filtrée avec l'opérateur PMCEV.

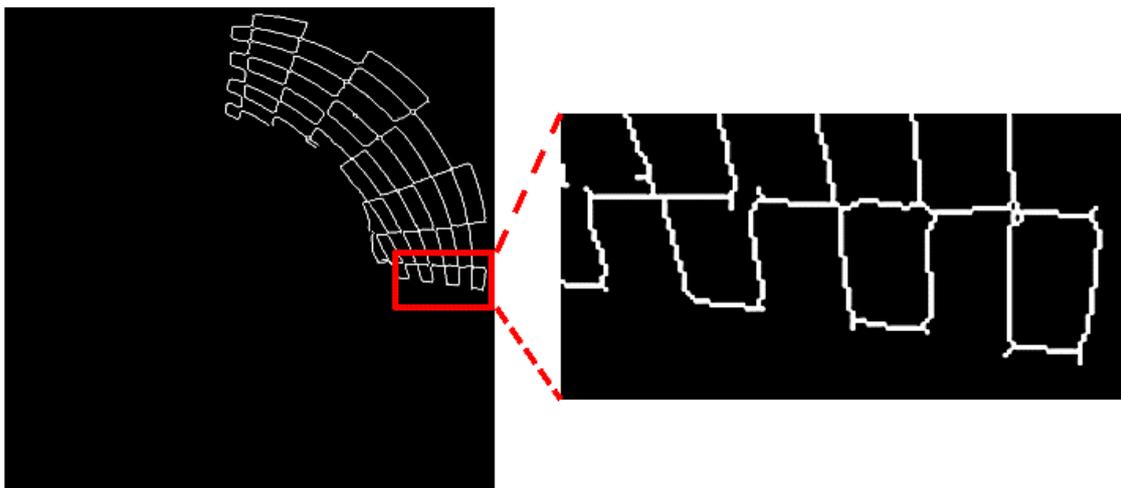


Figure 3.24 : Image réelle filtrée avec l'opérateur SMCEV.

Les images omnidirectionnelles réelles incorporent des dégradations et des bruits spécifiques issus de la chaîne d'acquisition. Les opérateurs de détection de contours basés sur le gradient (première dérivée) sont plus sensibles au bruit, alors que ceux basés sur le modèle électrostatique le sont moins. Les contours dans les images réelles sont moins distincts et plus flous, par conséquent on constate que les opérateurs classiques tendent à détecter le vrai contour ainsi que ses bordures, ce qui explique les doubles ou triples faux contours détectés par les opérateurs classiques. L'algorithme SMCEV non basé sur le gradient, est visiblement moins sensible aux dégradations de l'image, et par respect du vrai voisinage pixellique, il offre une reproduction fiable des contours présents dans l'image omnidirectionnelle.

3.4.5 Adaptation du détecteur de points d'intérêt de Harris

3.4.5.1 La détection des points d'intérêt

Parmi les différents modèles de caractéristiques bas-niveau d'une image, les primitives point, connues aussi comme points d'intérêt ou coins. Les points d'intérêt s'avèrent assez pertinents dans le sens de réduction de la quantité d'informations à analyser par un système de vision (figure 3.25). Les primitives point sont définis comme le point de courbure maximale plane dans la ligne de la plus grande pente de niveau de gris [112] [113], comme elles peuvent faire référence à une intersection d'au moins deux contours principaux et distincts. Fondamentalement, un coin représente la forme la plus simple des attributs bidimensionnels présents dans une image. L'extraction de points d'intérêt est d'une importance cruciale dans le domaine du traitement d'images et de vision artificielle. En effet c'est une tâche dont dépendent beaucoup d'applications, telle que la mise en correspondance, le suivi d'un objet, ou SLAM. Par conséquent, les performances de ces applications sont liées à la robustesse du détecteur de coins adopté.

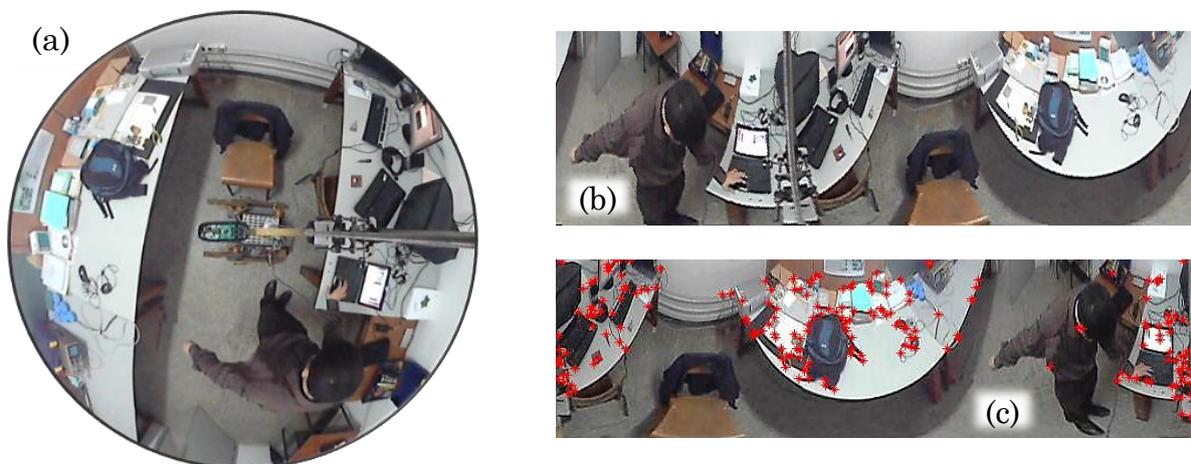


Figure 3.25 : Exemple d'extraction des points d'intérêt. (a) Image catadioptrique acquise avec un miroir parabolique. (b) Image panoramique correspondante. (c) Extraction de points d'intérêt.

Il existe principalement deux catégories de détecteurs de coins, à savoir, les détecteurs basés sur un modèle et ceux basés sur la géométrie. Un bon détecteur doit satisfaire quatre critères de robustesse principaux [114]:

- Bonne détection : Le détecteur doit détecter même les coins les plus subtiles, tout en ignorant le bruit.
- Bonne localisation : Les coins doivent être détectés le plus près possible de leurs véritables emplacements.
- La stabilité : La position détectée d'un coin ne doit pas bouger lorsque plusieurs images sont acquises de la même scène.
- Complexité : La complexité réduite de l'algorithme contribue à l'automatisation efficace du processus, et à une implémentation plus rapide.

3.4.5.2 Les détecteurs de coins basés modèle

La détection de coins à base de modèle, consiste à déterminer la similarité, ou la corrélation entre un modèle de taille $n \times n$ et toutes les sous-fenêtres de taille $n \times n$ dans une image donnée [115]. On suppose que le modèle est carré, $n = 2k + 1$. Soient $p_{i,j}$ et $q_{i,j}$ les éléments (i,j) du modèle P et de l'image Q . On définit respectivement la moyenne et la variance du modèle comme α et $\sigma(p)$. On définit aussi la moyenne β de la sous-fenêtre de l'image centrée au pixel (i,j) et sa variance $\sigma(q)_{ij}$. La formule généralisée de corrélation entre le modèle et l'image à la position (i,j) est donnée par :

$$\sigma_{ij} = \frac{\frac{1}{n} \sum_{l=-k}^k \sum_{m=-k}^k (p_{l,m} - \alpha)(q_{i+l,j+m} - \beta)}{\sqrt{\sigma(p)\sigma(q)_{ij}}} \quad 3.50$$

Malheureusement, il est impossible de concevoir des modèles qui peuvent couvrir toutes les orientations et les angles du coin. Ainsi, l'erreur est inévitable du fait de la structure complexe des coins.

3.4.5.3 Les détecteurs de coins basés sur la géométrie

La détection de coins basée sur la géométrie repose sur la mesure de la géométrie différentielle des primitives point. Les méthodes peuvent être divisées en trois catégories : la détection de coins liée aux contours, la détection de coins liée à la topologie et de détection de coins liée à l'autocorrélation.

- La détection de coins liée aux contours : on considère le point de coin comme jonction de deux ou plusieurs lignes de contours. Alors que le coin devient la caractéristique supplémentaire d'un point contour.
- La détection topologique des coins : La détection basée sur la topologie, considère les coins comme attribue géométrique interne sur la surface de l'image. Elle opère directement sur une image en niveaux de gris. Ces détecteurs sont basés sur la mesure des caractéristiques topologiques de la géométrie différentielle d'un coin sur la surface de l'image. Beudet [116] avait proposé une mesure invariante par rotation, qui dérive de la matrice

hessienne H . L'expression du second ordre de Taylor de la surface intensité est donnée par :

$$\Delta = \det(H) = \begin{vmatrix} I_{xx} & I_{xy} \\ I_{xy} & I_{yy} \end{vmatrix} \quad 3.51$$

La détection de coins est basée sur le seuillage de la valeur maximale de cette mesure de saillance (*cornerness measure*).

- La détection de coins à base d'autocorrélation : Moravec [117] a proposé de considérer la somme des différences au carré (*SSD*) autour d'un point considéré. Ce détecteur de coins fonctionne en considérant une fenêtre locale dans l'image, et en déterminant les variations moyennes d'intensité qui résultent du léger déplacement de cette fenêtre dans des différentes directions. Un coin peut être détecté lorsque la variation minimale produite par l'un des changements est supérieure à une valeur de seuil. En désignant les intensités de l'image par I , le changement E produit par un déplacement (x, y) est donnée par :

$$E_{xy} = \sum_{u,v} W_{u,v} [I(u+x, v+y) - I(u, v)]^2 \quad 3.52$$

Où W désigne la fenêtre actuelle de l'image, le détecteur de Moravec consiste tout simplement à détecter le maximum de $\min\{E_{xy}\}$ supérieur à un certain seuil.

3.4.5.4 Le détecteur de coins de Harris

Harris [118] a publié une extension de la méthode de Moravec, vu que le détecteur de Moravec est anisotrope en ce qu'il se base seulement sur les directions à 45 degrés. Celui de Harris prend en considération toutes les combinaisons possibles. La réponse du détecteur de Moravec est très sensible au bruit. Harris a ajouté un filtre Gaussien à sa méthode. À la différence de Moravec, Harris propose d'estimer la mesure d'autocorrélation du signal en utilisant les dérivées du premier ordre de l'image. Le détecteur de Harris exploite la fonction d'autocorrélation pour localiser les positions où le signal change dans les deux directions simultanément. La mesure de saillance proposée est donnée de la manière suivante :

$$c(x, y) = \det(M) - k(\text{trace}(M))^2 \quad 3.53$$

M est la matrice hessienne : $M = \begin{bmatrix} A & C \\ C & B \end{bmatrix}$, $\det(M) = \lambda_1 \lambda_2 = AB - C^2$.

Et : $\text{trace}(M) = \lambda_1 + \lambda_2 = A + B$,

A, B and C sont définis comme suit :

$$A = \left(\frac{dI}{dx}\right)^2 \otimes w \quad 3.54$$

$$B = \left(\frac{dI}{dy}\right)^2 \otimes w \quad 3.55$$

$$C = \left(\frac{dI}{dx} \frac{dI}{dy} \right) \otimes w \quad 3.56$$

$$\text{Avec } w = G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad 3.57$$

k est un paramètre d'ajustement d'amalgamer l'information contour contenue dans la trace de M avec l'information d'angularité donnée par son déterminant. Cette mesure est positive dans le cas d'un coin. Le détecteur est donc invariant par rotation. Notons que les valeurs propres de la matrice hessienne représentent les courbures capitales de la fonction d'autocorrélation, nous distinguons entre trois cas :

- Lorsque $c(x, y)$ est petite, ce qui arrive quand les valeurs propres λ_1 et λ_2 sont de petites valeurs. La région considérée est plate.
- Si $c(x, y)$ est négative, lorsque $\lambda_1 \gg \lambda_2$ ou l'inverse, la région est un contour.
- Si $c(x, y)$ est grande, ce qui arrive lorsque λ_1 et λ_2 sont de valeurs grandes mais proches. La région dans ce cas est un coin.

Le détecteur de Harris est sans doute le plus connu et adopté dans la littérature; il est robuste, fiable et standard, vu qu'il est opérationnel sur un large spectre d'images numériques.

L'inconvénient majeur du détecteur de Harris est son instabilité face au bruit, car il se base sur le calcul des dérivées du premier ordre qui représente en quelque sorte une procédure de détection de contours male conditionnée.

3.4.5.5 Détecteur de coins dédié aux images omnidirectionnelles

Dans le même contexte d'adaptation des traitements conventionnels aux images omnidirectionnelles, on s'est proposé de redéfinir le détecteur de Harris dans le cadre de la géométrie sphérique des images catadioptriques.

Comme une première étape de l'algorithme de détection de coins proposé, l'image catadioptrique $I(u, v)$ est mappée par projection stéréographique inverse sur la sphère, nous obtenons ainsi une image sphérique $I(\theta, \varphi)$. À la différence de la méthode de Harris conventionnelle, la détection de contours se fait en utilisant les masques de convolution $G_s(\theta, \varphi)$ issus du modèle sphérique de charges électriques virtuelles, développé dans la section 3.4.4.1. Ces masques ont été définis en coordonnées sphériques comme suit :

$$G_\varphi = 1/(\sin(\theta) d\varphi)^2 \begin{pmatrix} \sqrt{2}/4 & 0 & -\sqrt{2}/4 \\ 1 & 0 & -1 \\ \sqrt{2}/4 & 0 & -\sqrt{2}/4 \end{pmatrix} \quad 3.58$$

Et :

$$G_\theta = 1/(\sin(\theta) d\varphi)^2 \begin{pmatrix} \sqrt{2}/4 & 1 & \sqrt{2}/4 \\ 0 & 0 & 0 \\ -\sqrt{2}/4 & -1 & -\sqrt{2}/4 \end{pmatrix} \quad 3.59$$

L'étape suivante consiste à calculer la matrice d'autocorrélation M à chaque pixel d'image, nous définissons cette matrice en coordonnées sphériques par :

$$M = \begin{bmatrix} I_{\theta\theta} & I_{\theta\varphi} \\ I_{\varphi\theta} & I_{\varphi\varphi} \end{bmatrix} \quad 3.60$$

Le filtre de lissage gaussien est calculé à l'aide de la formule suivante :

$$G = \frac{1}{2\pi\sigma^2} e^{-d/2\sigma^2} \quad 3.61$$

Avec d la distance séparant le pixel central $p_c(\theta_c, \varphi_c)$ de son voisin $p_n(\theta_n, \varphi_n)$, donnée par : $d = \arccos(\sin \varphi_c \sin \varphi_n + \cos \varphi_c \cos \varphi_n \cos(\theta_n - \theta_c))$ 3.62

Finalement l'image de sortie $O(\theta, \varphi)$ est projetée stéréographiquement en $O(u, v)$ sur le plan. Cet algorithme est schématisé globalement de la sorte :

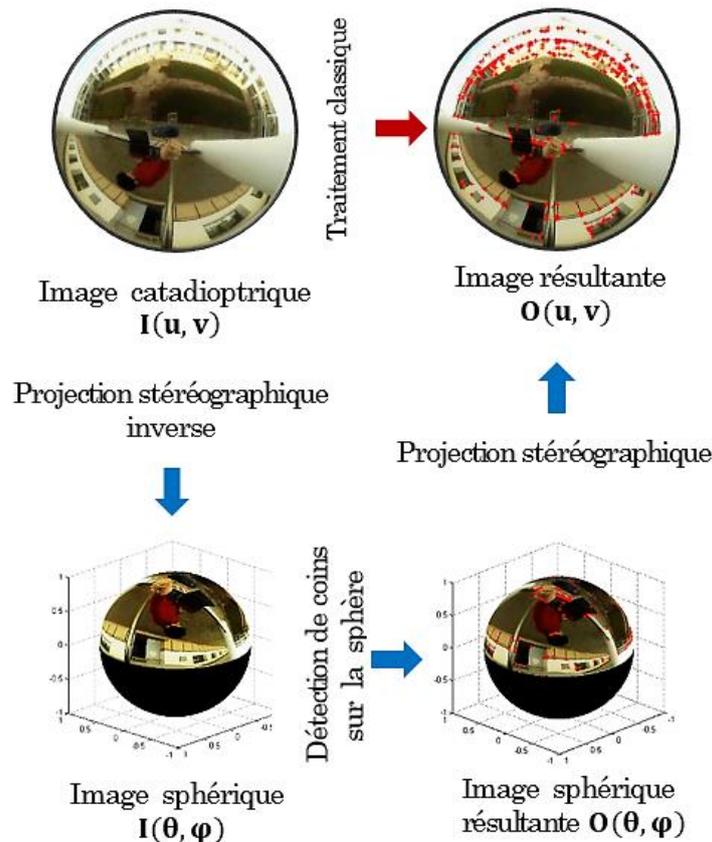


Figure 3.26 : Algorithme de détection de coins dans l'image sphérique.

La phase de détection de coins par calcul de la matrice d'autocorrélation se déroule elle-même en quelques étapes :

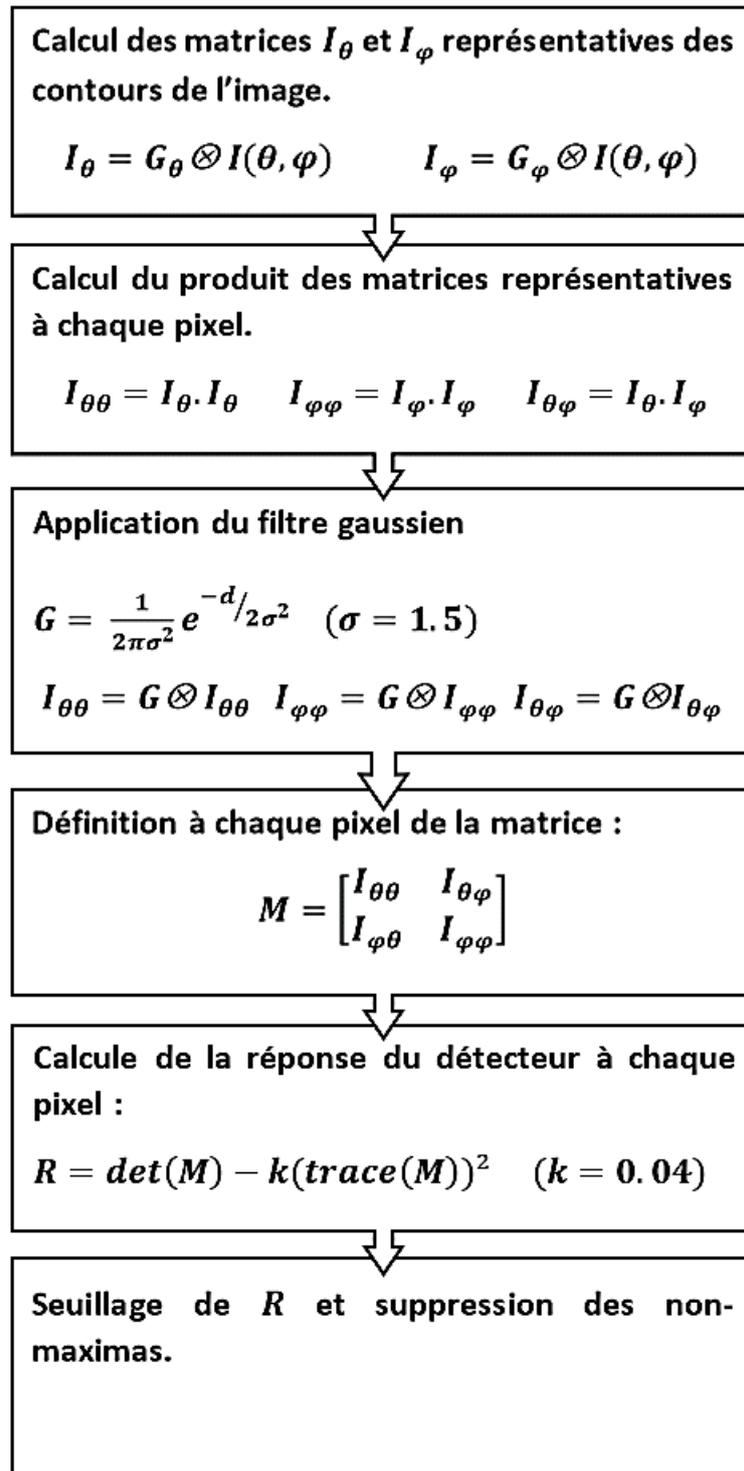


Figure 3.27: L'algorithme proposé de détection des coins.

L'algorithme proposé diffère de celui de Harris, en ce qu'il n'utilise pas le gradient pour la détection des contours de l'image. Cela, dans le but de renforcer la stabilité du détecteur en présence de bruit. L'algorithme opère sur l'image sphérique via projection stéréographique, cette transformation permet de pallier au problème du voisinage pixellique.

La redéfinition du filtre gaussien utilisée pour le lissage était nécessaire ; la distance entre les éléments de ce filtre est défini à l'origine selon une métrique euclidienne, nous nous sommes basés sur une métrique sphérique pour redéfinir cette distance. Les performances de l'algorithme proposé ont été évaluées expérimentalement en confrontation avec le détecteur standard de Harris.

3.4.5.6 Résultats illustratifs et comparaisons

Afin d'évaluer les performances du détecteur de coins proposé, nous avons mené une étude comparative avec l'algorithme standard de Harris. L'évaluation des deux détecteurs mis en confrontation est faite à base des deux critères suivants :

- Détection : Le détecteur doit repérer le maximum des vrais coins, et le minimum des faux points dus au bruit.
- Localisation : La position des coins détectés doit être la plus proche de leurs vrais emplacements.

Initialement une image synthétique a été générée, d'une taille de 2180x2180 pixels montrée sur la figure 3.28.a. Cette image est de géométrie omnidirectionnelle, incluant une distorsion radiale similaire à celle présente dans les images omnidirectionnelles réelles. Elle représente un damier constitué de deux classes de pixels, avec des intensités de 0 et 204 en niveaux de gris. Les coins sur cette image sont relativement rapprochés. Pour tester les performances de l'algorithme sous différents niveaux de résolution, et un nombre total de coins diversifié, nous avons générer une seconde image synthétique d'une taille de 2390x2390 (figure 3.28.b) avec des caractéristiques similaires à la première image, mais incluant moins de coins qui sont relativement plus espacés.

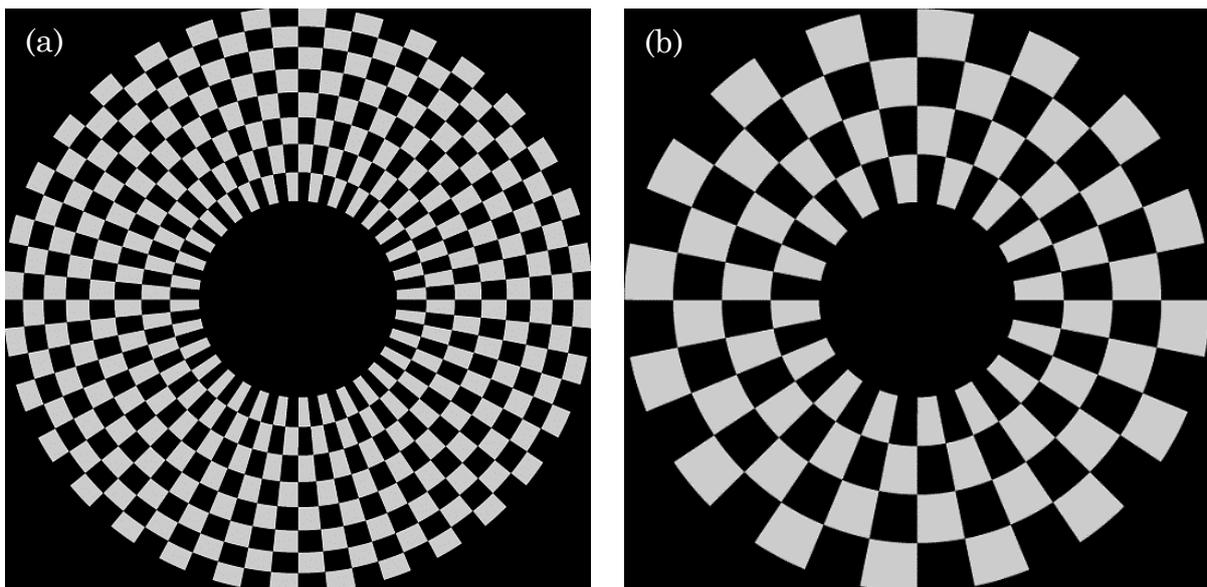


Figure 3.28 : (a) La première image synthétique utilisée pour l'évaluation de la détection des coins. (b) La seconde image synthétique

A chaque expérience menée nous comparons les résultats obtenus avec l'algorithme proposé opérant sur la sphère, avec ceux obtenus en utilisant l'algorithme de Harris opérant sur le plan. L'algorithme proposé utilise les masques SMCEV pour détecter les contours, alors que Harris standard utilise le filtre de Sobel.

Nous évaluons quantitativement chaque détecteur en calculant trois scores inspirés des méthodes d'évaluation de performance de la biométrie :

- Le taux des faux rejets (*False Rejection Rate*), que nous définissons comme :

$$FRR = \frac{u_c}{t_c} \quad 3.63$$

Avec u_c le nombre de vrais coins non détectés. Et t_c le nombre total de coins.

- Le taux des fausses acceptations (*False Acceptance Rate*) :

$$FAR = \frac{f_c}{t_p} \quad 3.64$$

Avec f_c le nombre de faux coins détectés, et t_p le nombre total de tous les points détectés.

- Le taux global de l'erreur (*Total Error Rate*) :

$$TER = \frac{u_c + f_c}{t_c + t_p} \quad 3.65$$

Notons que, afin d'assurer une évaluation objective, nous avons utilisé le même coefficient de lissage gaussien (*standard deviation*) σ égal à 1,5 pour les deux détecteurs à chaque expérience. Étant donné que le choix d'une valeur appropriée de ce coefficient reste un problème ouvert. La figure 3.29 et la figure 3.30 illustrent respectivement les résultats obtenus avec la première et la seconde image

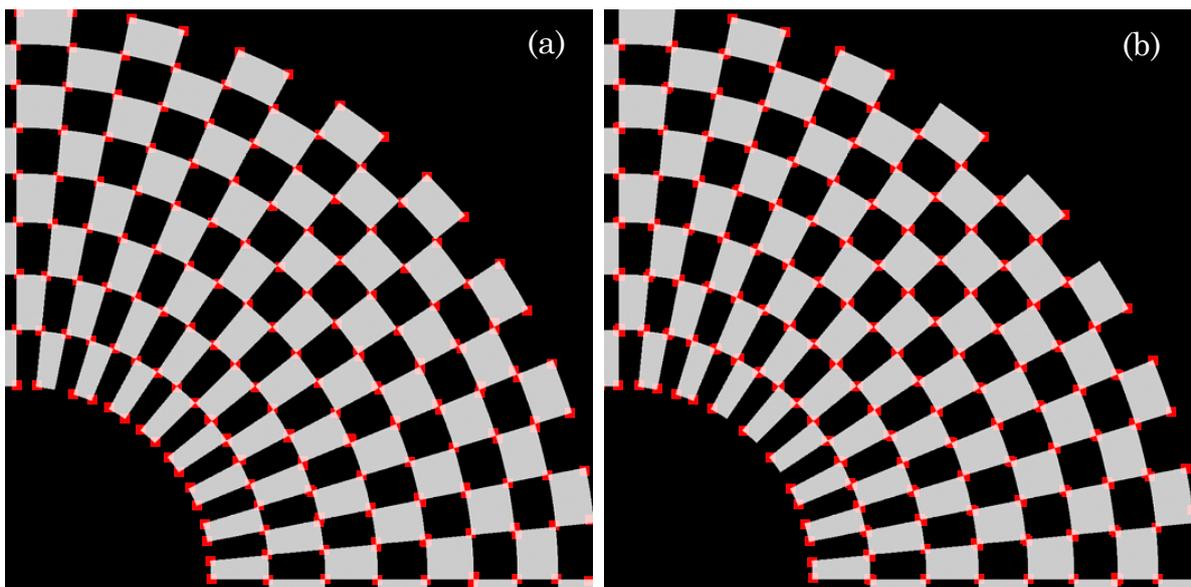


Figure 3.29 : Résultat de détection de coins dans la première image synthétique. (a) L'algorithme proposé. (b) Harris standard.

synthétique. Ensuite le tableau 3-3 et le tableau 3-4 présentent les taux de détection correspondants.

Tableau 3-3 : Taux d'évaluation de la première image synthétique.

| | Approche proposée | Harris standard |
|-----|-------------------|-----------------|
| FRR | 0.65% | 6.63% |
| FAR | 18.71% | 44.78% |
| TER | 10.58% | 30.58% |

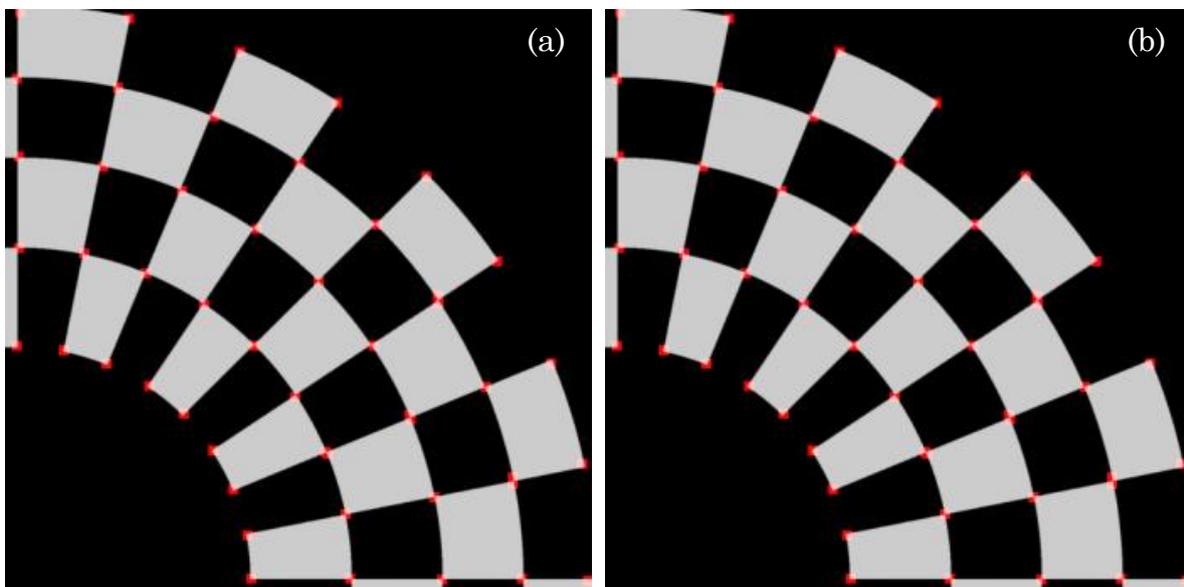


Figure 3.30 : Résultat de détection de coins dans la seconde image synthétique originale. (a) L'algorithme proposé. (b) Harris standard.

Tableau 3-4 : Taux d'évaluation de la seconde image synthétique

| | Approche proposée | Harris standard |
|-----|-------------------|-----------------|
| FRR | 0% | 11.11% |
| FAR | 8.16% | 42.02% |
| TER | 4.25% | 29.82% |

Afin d'étudier le comportement de l'approche proposée devant certaines dégradations de l'image. Tout d'abord un bruit gaussien de moyenne nulle et variance 0,01 est ajouté à la seconde image synthétique. Nous avons obtenu les résultats suivants :

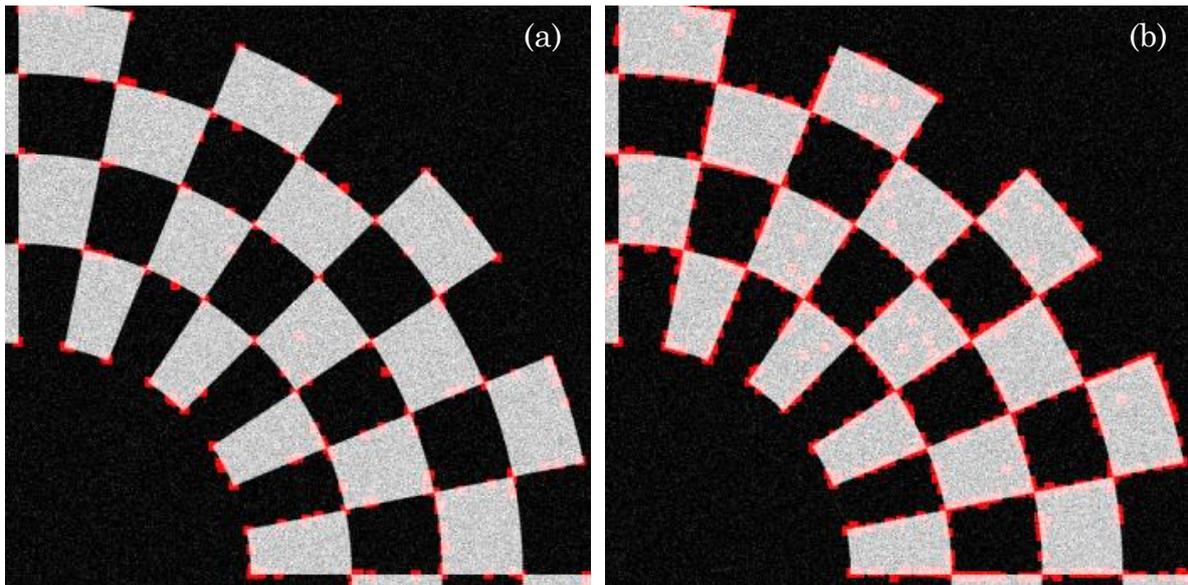


Figure 3.31 : Résultat de détection de coins dans la seconde image synthétique bruitée. (a) L'algorithme proposé. (b) Harris standard.

Tableau 3-5 : Taux d'évaluation de la seconde image synthétique (bruit gaussien)

| | Approche proposée | Harris standard |
|-----|-------------------|-----------------|
| FRR | 0% | 6.66% |
| FAR | 68.53% | 92.02% |
| TER | 52.12% | 85.28% |

Ensuite, nous simulons une image floue qui pourrait être le résultat du mouvement de la caméra, avec un angle de 0 degrés dans le sens antihoraire. Ce qui correspond à un mouvement horizontal de neuf pixels. Les résultats sont les suivants :

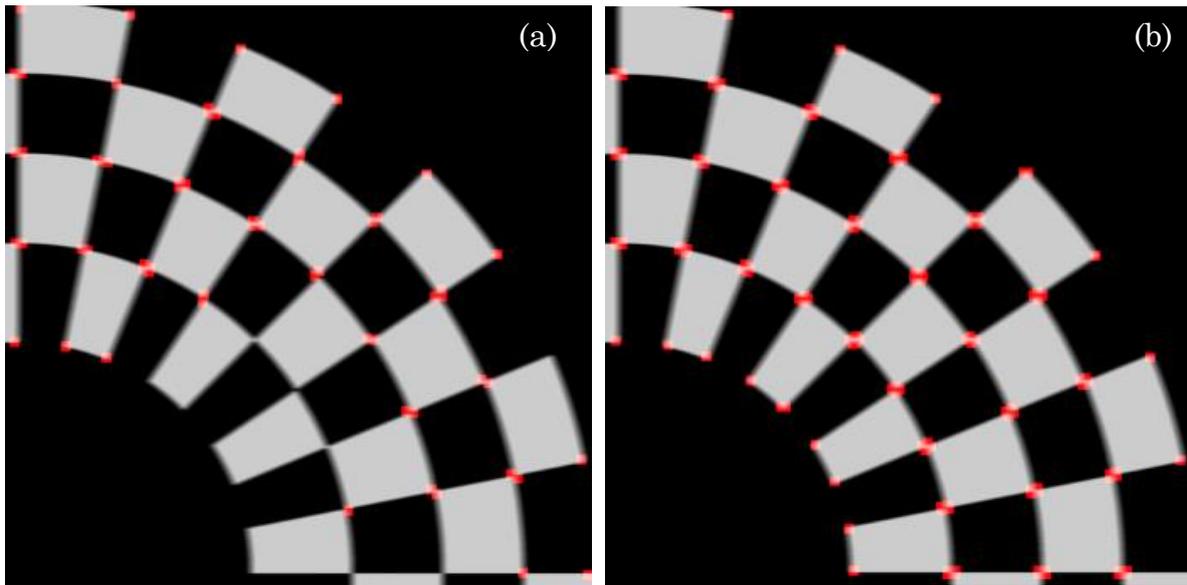


Figure 3.32 : Résultat de détection de coins dans la seconde image synthétique floue. (a) L'algorithme proposé. (b) Harris standard.

Tableau 3-6 : Taux d'évaluation de la seconde image synthétique (floue).

| | Approche proposée | Harris standard |
|-----|-------------------|-----------------|
| FRR | 24.44% | 4.44% |
| FAR | 44.26% | 67.18% |
| TER | 35.84% | 51.13% |

L'évaluation de la localisation des coins détectés est effectuée en mesurant la distance entre la position des vrais coins et ceux détectés par chaque détecteur. Considérons deux pixels $p(u_p, v_p)$ et $q(u_q, v_q)$, cette distance euclidienne est calculée selon l'équation suivante :

$$d = \sqrt{(u_p - u_q)^2 + (v_p - v_q)^2} \quad 3.66$$

Un bloc de neuf coins de l'image synthétique a été considéré, les résultats obtenus sont les suivants :

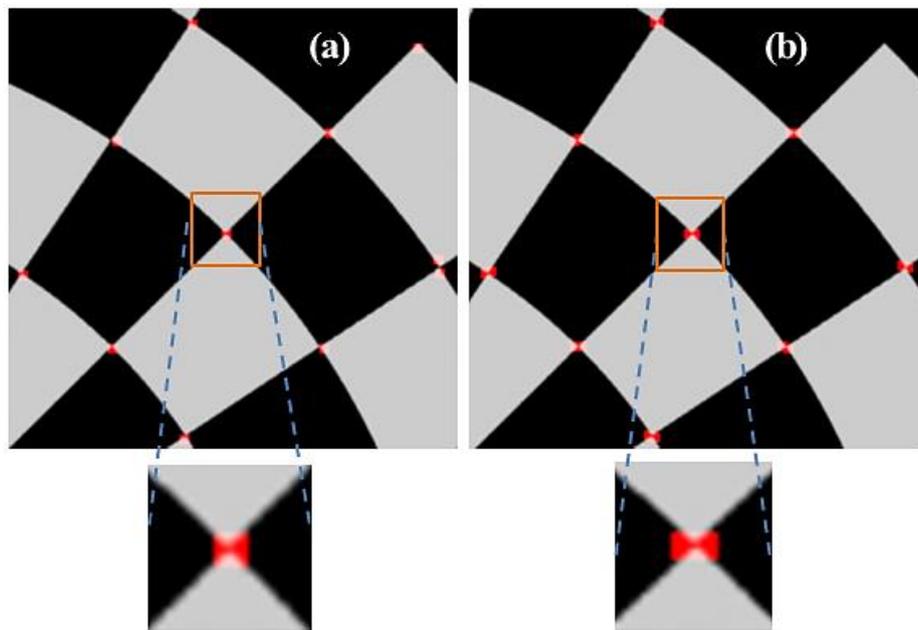


Figure 3.33 : Résultats d'évaluation de la localisation des coins détectés. (a) L'approche proposée. (b) Harris standard.

Tableau 3-7 : Evaluation de la localisation des coins détectés par mesure de la distance euclidienne.

| | Approche proposée | Harris standard |
|-----|-------------------|-----------------|
| d | 2.15 | 2.96 |

Nous pouvons remarquer en observant le tableau 3-3 que les taux FRR et FAR de détecteur standard de Harris sont plus élevés que ceux de notre approche. Ceci peut être expliqué par le fait que le détecteur standard tend à détecter une grande quantité de doubles coins (FAR supérieur), tout en ignorant certains vrais coins (FRR supérieur).

Ces mêmes observations ont été confirmées lors de l'exécution des deux méthodes sur une seconde image omnidirectionnelle synthétique de caractéristiques différentes.

L'influence de l'utilisation de noyaux de détection de contours non basé sur le gradient, était très perceptible lors du traitement d'une image bruitée. Les résultats présentés dans le tableau 3-5 ont prouvé que l'approche proposée est plus stable en présence du bruit.

Considérant le TER (taux global d'erreur), dans le cas d'une image floue (figure 3.32) l'approche proposée demeure fournir de meilleurs résultats que l'approche classique. Mais son FRR (taux des faux rejets) est relativement élevé, cela signifie que certains coins réels ont été manquées.

D'après le tableau 3-7, on peut voir l'amélioration relative de la précision de localisation des coins détectés par la méthode proposée, par respect du voisinage correcte des pixels traités.

En deuxième phase, autres expérimentations ont été conduites sur des images catadioptriques réelles. Une scène d'intérieur, incluant un damier en noir et blanc a été acquise avec une caméra catadioptrique, dans le but d'évaluer le détecteur dans des conditions réelles d'utilisation. Les résultats obtenus sont présentés sur les figures ci-dessous :

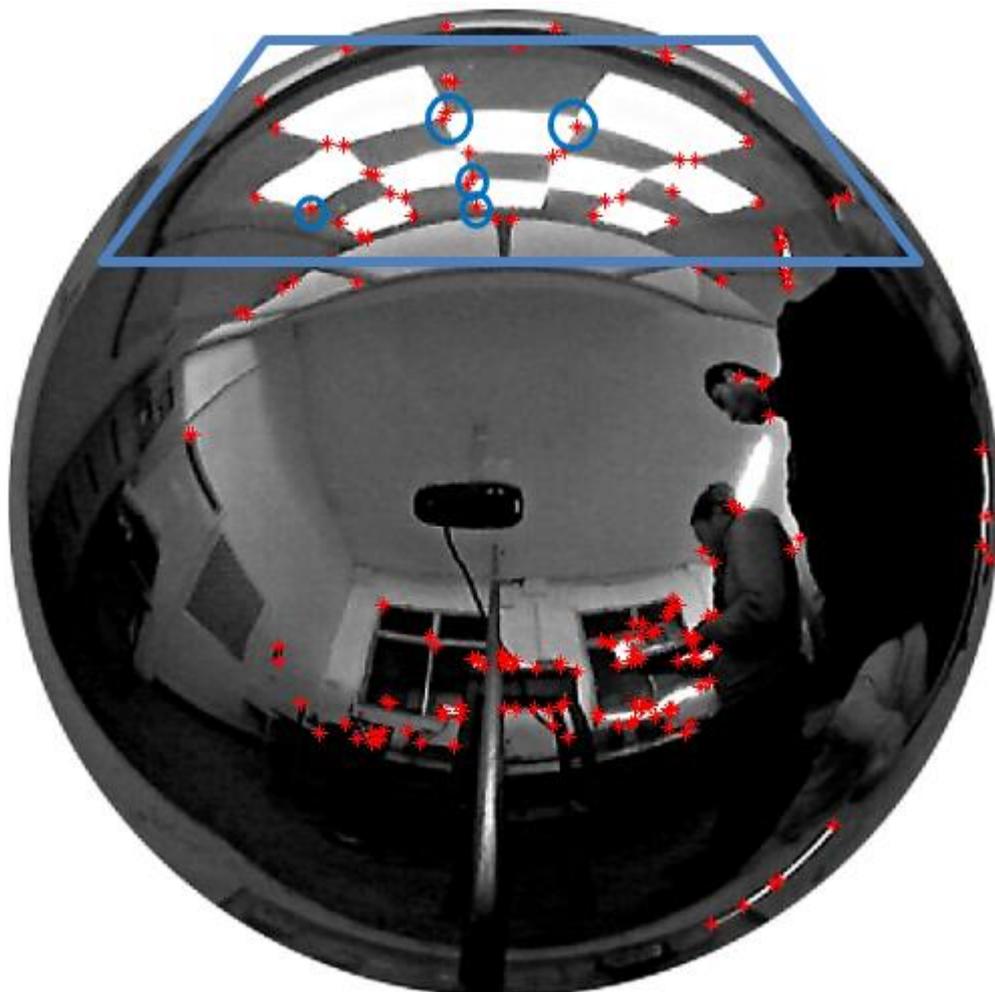


Figure 3.34 : Résultats de détection de coins sur une image réelle en utilisant l'algorithme proposé.

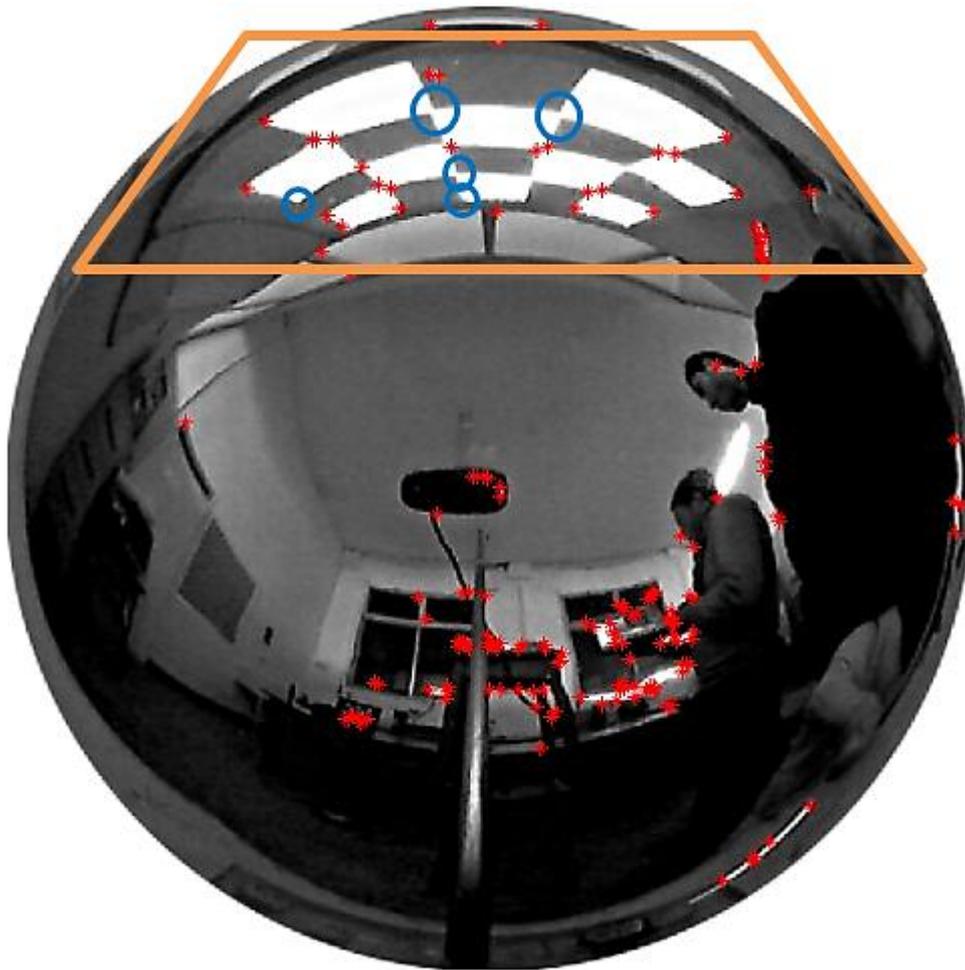


Figure 3.35 : Résultats de détection de coins sur une image réelle en utilisant l'algorithme standard de Harris.

Les expériences menées sur des images catadioptriques réelles (figure 3.34, figure 3.35) confirment que la robustesse de l'approche proposée face au bruit n'est pas simplement due à un effet de lissage intrinsèque; voyant que dans le cas d'une scène réelle d'un damier, cinq vrais coins (encerclés en bleu) ont été manqués par l'approche classique, alors qu'ils ont été détectés par l'approche adaptée. En d'autres termes l'efficacité du détecteur proposé contre les diverses formes de bruit, n'affecte pas son taux de faux rejets et offre toujours un bon compromis entre les critères caractérisant un bon algorithme de détection de coins comme ils ont été mentionnés précédemment.

3.5 Conclusion

Les images omnidirectionnelles présentent d'importantes distorsions radiales et des propriétés géométriques spécifiques issues de leur processus d'acquisition. De nombreux chercheurs se sont intéressés aux problèmes du traitement et d'interprétation de ce type d'images. Des auteurs ont penché pour appliquer les mêmes traitements classiques développés pour les images perspectives, sur les images omnidirectionnelles. Cette approche aboutit à des résultats insatisfaisants, puisque la géométrie de formation de ces images n'est pas prise en compte.

D'autres auteurs ont préféré d'adapter les traitements classiques. L'introduction du modèle de projection unifié, et de la projection stéréographique dans le domaine de la vision omnidirectionnelle, a permis de mettre en évidence un espace géométrique de représentation des images catadioptriques plus adéquat, qui est la sphère d'équivalence. Nous avons présenté dans ce chapitre, l'élaboration de nouveaux outils de traitement des images sphériques. Il s'agit d'un opérateur de détection de contours, basé sur modèle sphérique de charges électriques virtuelles. L'évaluation comparative aux autres opérateurs classiques, a prouvé clairement l'amélioration apportée par l'approche proposée. Nous nous sommes basé sur ce fondement pour adapter l'algorithme de détection des coins de Harris aux images omnidirectionnelles. La comparaison quantitative des critères de bonne détection et bonne localisation a relevé l'efficacité de l'algorithme proposé.

4. La vision omnidirectionnelle pour les systèmes autonomes

4.1 Introduction

L'exploitation des capteurs de vision est de plus en plus répandue dans différents domaines d'applications (vidéosurveillance, robotique mobile, véhicules intelligents, etc...). Les caméras sont en effet des sources d'informations généreusement plus riches que les télémètres et plus discrètes. La vision artificielle cependant nécessite des traitements plus élaborés et plus adaptatifs.

Grâce à son champ de vision omnidirectionnel, une caméra catadioptrique est beaucoup plus avantageuse pour certains types d'applications comme la localisation et le suivi visuel d'une cible en mouvement. Dans ce chapitre nous proposons l'utilisation des systèmes de vision omnidirectionnels pour améliorer les procédures de détection et de suivi des cibles mobiles. Nous présentons d'abord un robot mobile, doté d'un système de vision artificielle omnidirectionnelle, capable de détecter et de suivre une cible en mouvement de façon autonome. Ensuite nous présentons un suiveur solaire équipé d'une caméra catadioptrique, ce dispositif permet l'orientation optimale de panneaux photovoltaïques, dans le but d'augmenter leur rendement énergétique.

Il y a environ deux décennies, l'idée d'avoir des robots se déplaçant comme les humains ou faire tâches humanoïde était une simple pensée futuriste. Aujourd'hui, la nécessité pour les applications robotiques mobiles pratiques prend de l'importance. La robotique autonome implique différents domaines de la recherche qui nécessitent une attention particulière individuelle, mais il y a des capacités qui pourraient être considérées comme générales et fondamentales quelle que soit la tâche particulière. Ces tâches incluent le suivi, la localisation, la navigation, l'évitement d'obstacles, et la planification de parcours. Pour effectuer ces tâches, des capteurs doivent être utilisés pour percevoir l'environnement du robot. Les sonars, les télémètres laser, les modules GPS, accéléromètres, et les caméras sont les capteurs les plus éminents. Les Sonars [119] et les télémètres laser [120] ont fait l'objet de beaucoup travaux de recherche et en général les tâches fondamentales mentionnées précédemment peuvent être traitées avec de très bonnes performances dans des environnements intérieurs. Il y a un consensus que la vision devrait jouer un rôle clé dans les systèmes robotiques autonomes, et les quelques dernières années ont prouvé que les caméras sont souvent un bon choix. La quantité d'information qui peut être extraite à partir d'images est extrêmement riche, et un système de vision nous donne la possibilité de viser non seulement les compétences premières telles que la localisation, mais la conscience spatiale et une compréhension plus générale de la scène. Plus récemment, les caméras omnidirectionnelles ont été utilisées pour la navigation. Cette idée a beaucoup de points en commun avec la biologie où la majorité des insectes et arthropodes bénéficient d'un large champ de vision [121]. L'intégration de la vision

omnidirectionnelle en robotique mobile, a ouvert un large éventail de recherche. Dans [122] Aihara et al., ont utilisé une séquence d'images omnidirectionnelles auto-corrélées pour déterminer la position globale. L'auto-localisation des robots mobiles a été étudiée dans [123], l'acquisition des données a été faite à l'aide d'une caméra couleur omnidirectionnelle, ces données sont traitées ensuite par un réseau de neurones artificiel, de sorte que le robot pourrait connaître son environnement. Delahoche [124] a présenté un système qui a fusionné l'odométrie et les données d'images omnidirectionnelles en utilisant un filtre de Kalman étendu. La position des balises dans le repère absolu de l'environnement ont été connus a priori.

4.2 Conception et réalisation d'un robot mobile

Compte tenu de cette motivation, nous avons choisi de mettre l'accent sur cette discipline dans cette thèse, en développant de nouvelles méthodes pour l'utilisation des capteurs de vision omnidirectionnelle pour la navigation. Pour cela nous avons conçu et réalisé un robot mobile nommé ESCALADE360, qui nous a servi comme plateforme de teste et d'expérimentation des méthodes développées.

Le développement du robot en question a passé par trois principales phases de conception. En décembre 2011, une première version d'ESCALADE360 a été réalisée (figure 4.1). Il s'agissait d'une plateforme mécanique à chenilles, avec deux moteurs DC 12V à réducteur assurant les mouvements principaux, et deux autres moteurs similaires, responsables des mouvements des deux flippers installés en avant de la plateforme. Le tout étant alimenté par deux batteries de 12V 9Ah. La plateforme pouvait être contrôlée par une commande filaire à relais.

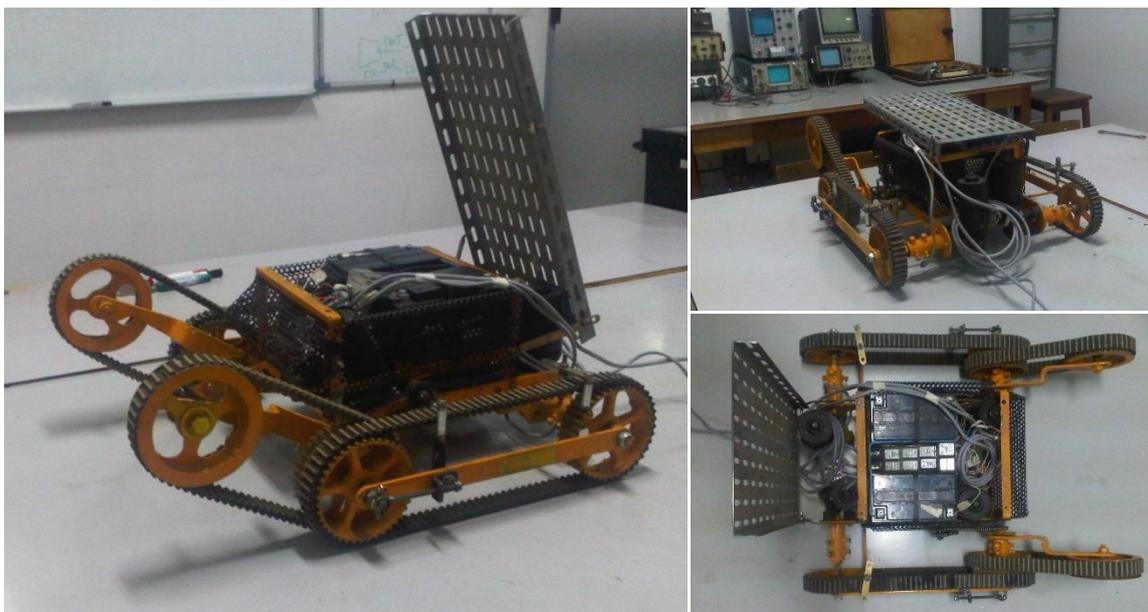


Figure 4.1 : Le robot ESCALADE360 v1.0

La configuration mécanique à chenilles assure plus d'adhérence sur les terrains déformés. Tant qu'aux flippers, ils permettent de franchir les petits obstacles. Parmi les différentes configurations possibles pour les robots à chenilles ³¹(figure 4.2), nous avons opté pour le concept "PackBot" qui assure une plus grande mobilité et flexibilité. Ce concept a été développé à l'origine par la Société *iRobot* en 2000 pour des robots démineurs, et a été largement adopté par la suite, notamment pour les robots de sauvetages et les robots monte-escalier.

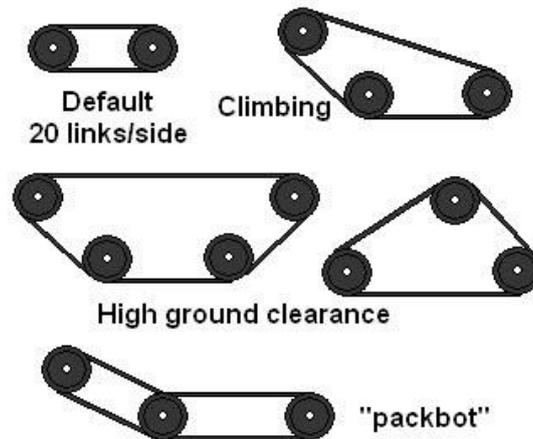


Figure 4.2 : Différentes configurations pour les robots à chenilles.

En février 2012, une deuxième version d'ESCALADE360 a été réalisée (figure 4.3).



Figure 4.3 : ESCALADE360 v2.0

³¹ Source de figure : [160]

La plateforme avait été allongée pour lui donner la capacité d'escalader les rampes et les escaliers. Elle a été menue d'un ordinateur à processeur *Intel® pentium® 4* alimenté par un onduleur 12VDC-220VAC, d'un système de vision omnistéréo à caméra unique, ainsi que d'un circuit de commande à base de microcontrôleur assurant le contrôle des moteurs à travers six relais. La communication entre l'unité de traitement et le circuit de commande consistait en un bus parallèle de 8 bits. Le robot ESCALADE360 v2.0 pouvait naviguer de façon autonome, comme il pouvait être commandé par réseau local Wifi.

La troisième version (actuelle) du robot a été réalisée en avril 2013 (figure 4.4). L'unité de traitement avait été remplacée par un ordinateur portable à processeur *Intel® Core™ i5*, cela a permis d'optimiser l'autonomie ainsi que le poids du robot, qui pouvait embarquer l'un des deux systèmes de vision omnidirectionnelle. Le premier étant un système omnistéréo binoculaire, le deuxième consistait en un système catadioptrique monoculaire. Des améliorations ont été faites sur le châssis du robot ainsi que son circuit de commande qui utilise désormais un double pont H de transistors de puissance pour commander les quatre moteurs au lieu des relais, tout en étant relié à l'ordinateur via une liaison RS232. La navigation peut être autonome, commandée par Wifi ou par internet.

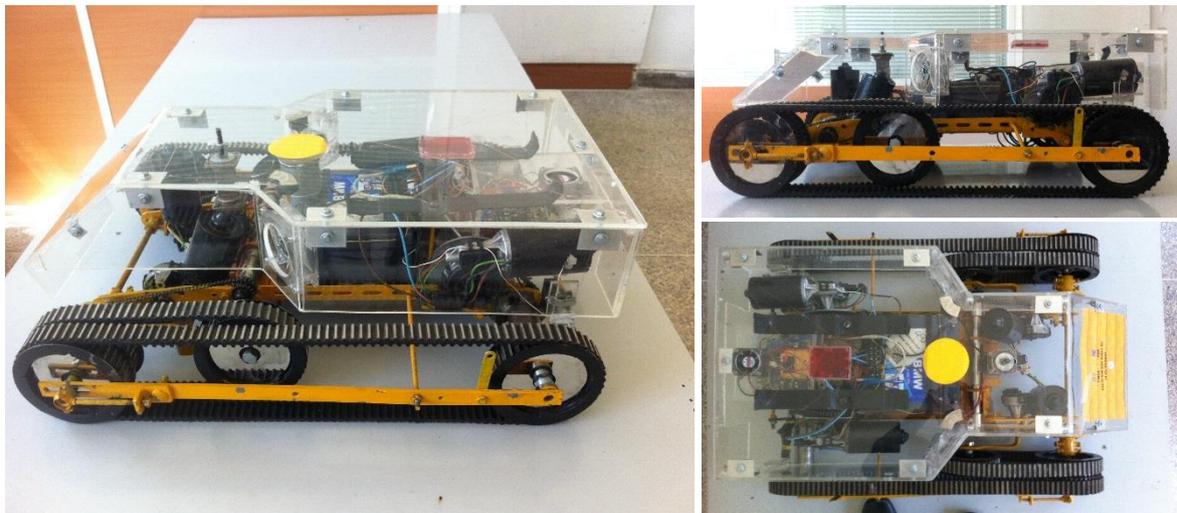


Figure 4.4 : ESCALADE360 v3.0

4.2.1 L'électronique

Le système électronique permettant de gérer le robot se compose de trois modules principaux : un module d'acquisition qui consiste en un capteur de vision omnidirectionnelle, une unité de traitement, et un système de contrôle.

Le module d'acquisition : Nous avons conçu trois types de systèmes d'acquisition omnidirectionnels qui peuvent être embarqués sur la plateforme robotique. Le premier est un capteur omnistéréo monoculaire à *baseline* variable, présenté dans le deuxième chapitre. Le deuxième est un système omnistéréo

équivalent au précédent, mais binoculaire et à *baseline* fixe, et finalement une caméra omnidirectionnelle monoculaire.

Le premier système (figure 4.5.a) peut être utilisé comme système de vision omnidirectionnel 2D en temps réel, comme il peut fonctionner tel un système de stéréovision (3D), dans ce cas l'acquisition ne peut se faire en temps réel, vu que le capteur a besoin d'un certain temps pour acquérir la scène à partir de deux points de vue distants. Ce capteur est plus adéquat pour les scènes statiques. Le mouvement de l'ensemble caméra-miroir le long de l'axe vertical est assuré par un vérin mécanique, motorisé par un moteur DC 12V. Les deux positions extrêmes, basse et haute, du système omnistéréo sont bridées électroniquement par des capteurs de fin de course.



Figure 4.5 : (a) Capteur de vision omnistéréo à baseline variable. (b) Capteur omnistéréo à baseline fixe.

Le second capteur de vision artificielle (figure 4.6 figure 4.5.b), est un système omnistéréo à deux cameras CCD, chacune orientée vers un miroir hémisphérique. La géométrie ainsi que les formules de triangulation, de ce système sont équivalentes au précédent à la différence que selon cette configuration, ce capteur

peut opérer en temps réel, vu que l'acquisition de la paire d'image omnidirectionnelle peut se faire instantanément.

Le troisième capteur (figure 4.6) consiste simplement en une caméra catadioptrique monoculaire à miroir hémisphérique ; un système plus compacte mais n'offrant que l'information en 2D. On peut remarquer que ces trois systèmes de vision proposés peuvent être orientés soit vers le bas, ce qui privilège la zone inférieure de scène notamment le sol. Soit vers le haut ce qui favorise la zone supérieure de la scène. Le choix de l'une de ces configurations dépendra de l'application visée et de l'emplacement des objets à manœuvrer par le robot. Il est à noter que les cameras utilisées sont toutes connectées à l'unité du traitement par le biais d'une liaison USB 2.0.



Figure 4.6 : Capteur de vision omnidirectionnelle monoculaire.

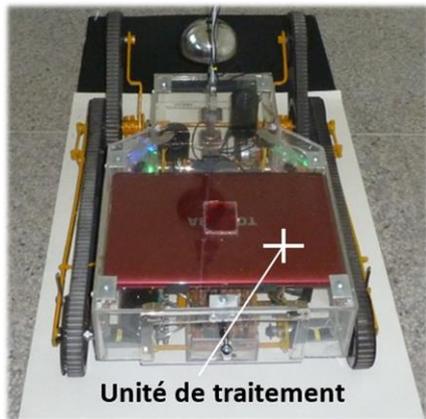


Figure 4.7 : Unité de traitement du robot ESCALADE360

Le module de traitement : L'unité de traitement est un ordinateur embarqué sur le robot, chargé de la gestion d'acquisition des images omnidirectionnelles, de leur traitement, ainsi que de la prise de décisions. Etant muni d'un concentrateur (*HUB*) USB externe, l'ordinateur peut être connecté à plusieurs caméras parallèlement. Cette unité est alimentée par sa propre batterie. Les décisions prises sont traduites en un langage de communication entre l'ordinateur et le circuit de commande, sous formes d'instructions transmises via une liaison RS232. Ces instructions permettent de gérer les mouvements principaux du robot, des flippers, du système omnistéréo, etc...

Le système de contrôle : Les instructions émises par l'unité de traitement sont traitées par un circuit de commande (figure 4.7). Le noyau de ce système est un microcontrôleur de type *RISC* à 16 bits. Chaque instruction reçue correspond à une action qui doit être faite. Parmi ces actions : activer les moteurs en marche avant, en marche arrière, faire tourner le robot à droite ou à gauche, actionner les flippers vers le haut ou vers le bas, et ainsi de suite.

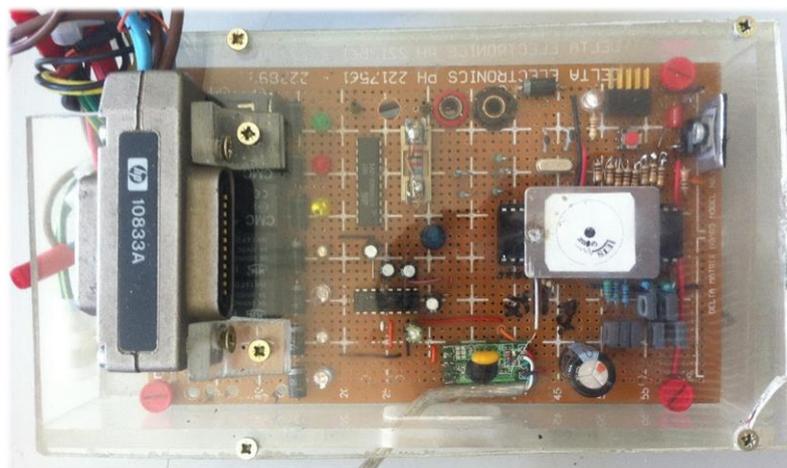


Figure 4.8 : Circuit de commande du robot ESCALADE360.

Le microcontrôleur agit sur les mouvements des moteurs par le biais d'un double pont H de transistors de puissance. Deux de ces ponts sont dédiés aux deux moteurs principaux agissant sur les chenilles. Le troisième est réservé aux deux moteurs couplés qui permettent de mouvoir les flippers du robot. Le moteur du système omnistéréo à caméra mobile est géré par deux relais. L'alimentation du robot en énergie est assurée par une batterie étanche rechargeable de 12V d'une capacité de 19Ah. La tension délivrée au système de contrôle (figure 4.9) est

régulée et stabilisée. Afin d'assurer au robot développé une autonomie en énergie, nous avons pensé à l'équiper d'un module d'énergie renouvelable optimisé. Le module en question consiste en un panneau photovoltaïque embarqué sur un suiveur solaire. Cette configuration qui sera détaillée dans la section 4.3 permet de garantir un meilleur rendement en minimisant l'angle d'incidence du rayonnement solaire.

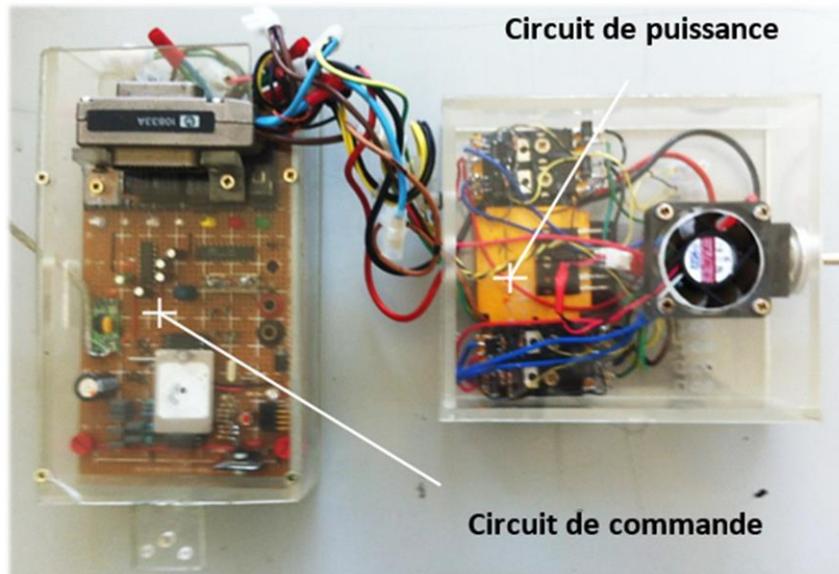


Figure 4.9 : Système de contrôle du robot ESCALADE360.

4.2.2 La mécanique

Les composantes sont rassemblées autour d'un châssis en acier, mobile sur huit roues dentées articulées par quarts chaînes de transmission en caoutchouc à double face dentées. Les quatre roues primaires sont entraînés par deux moteurs électriques BOSCH Réf : "0 390 246 310" à réducteur mécanique à roue et vis sans fin (figure 4.10). Ce type de moteur peut fonctionner à trois différentes vitesses par changement du brochage du moteur.



Figure 4.10 : Moteur électrique DC 12 V.

La combinaison des deux permutations possibles du sens de rotation de ces deux moteurs primaires (figure 4.11) permet d'assurer les quatre mouvements principaux du robot, qui sont les deux rotations sens horaire et antihoraires, ainsi que les deux translations vers l'avant et vers l'arrière.

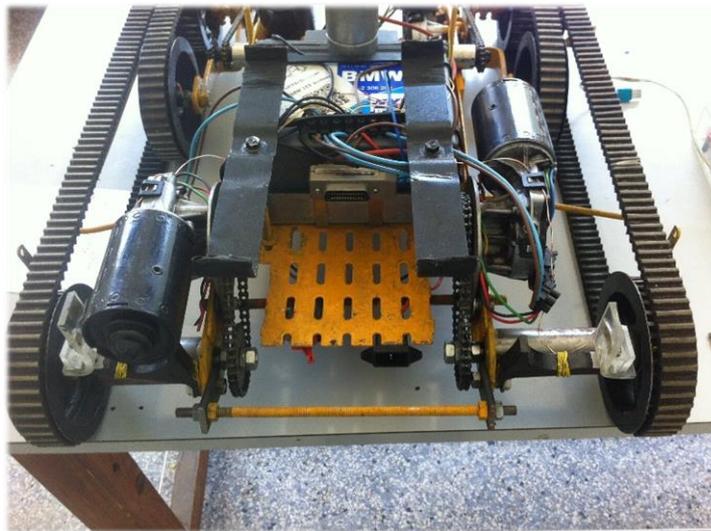


Figure 4.11 : Les moteurs primaires du robot.

La transmission du mouvement de rotation des deux moteurs aux roues arrière est assurée par deux chaînes de distribution entraînant deux axes logés dans des moyeux à roulement.

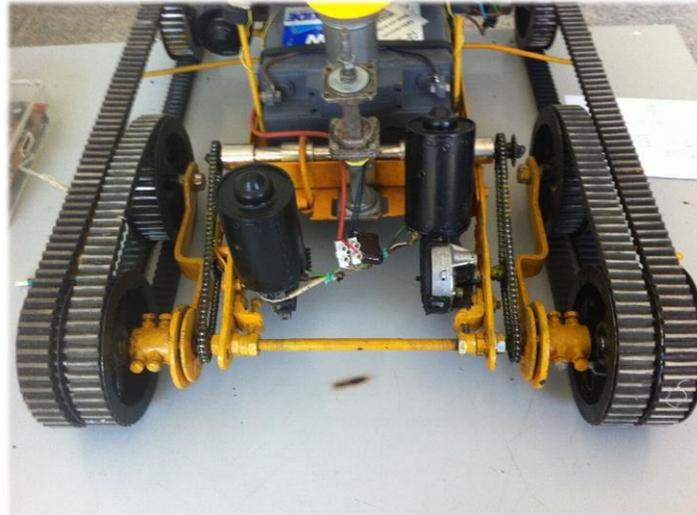


Figure 4.12 : Mécanisme des flippers du robot ESCALADE360.

Les flippers du robot sont actionnés par deux autres moteurs secondaires similaires couplés entre eux (figure 4.12). Les deux roues intérieures avant sont solidaires à ceux extérieures avant et servent de pivot pour les chaînes de transmission équipant les flippers lesquelles peuvent tourner autour de l'axe pivot, d'un angle pouvant atteindre 360° .

L'utilité de ce mécanisme s'explique d'un côté par la possibilité d'agir sur la surface de contact du robot avec le terrain selon quatre positions principales des flippers (figure 4.13). La première position réduit la surface de contact des chenilles à

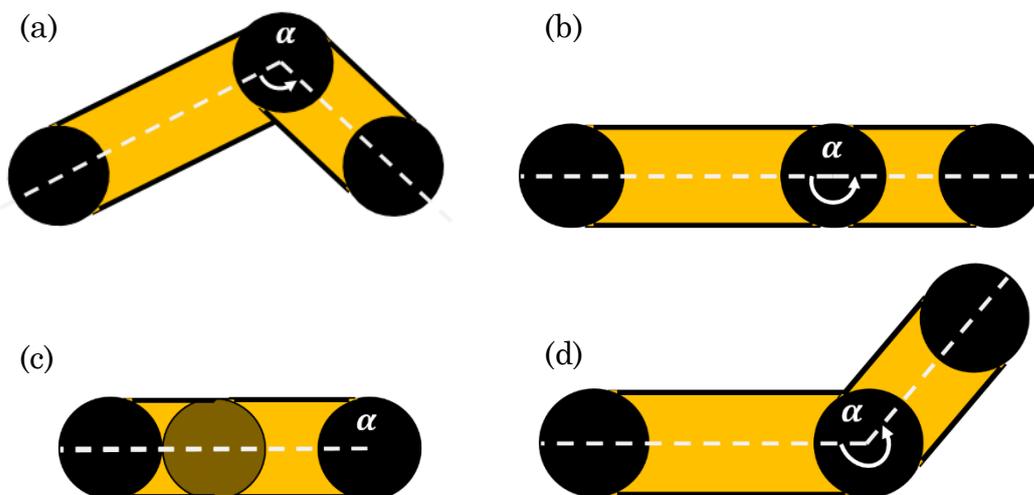


Figure 4.13 : Les quatre positions principales des flippers. (a) Quatre points de contact, $\alpha \in]0^\circ, 180^\circ[$. (b) Contact allongé, $\alpha = 180^\circ$. (c) Contact condensé $\alpha = 0^\circ$. (d) Contact réduit. $\alpha \in]180^\circ, 360^\circ[$.

quatre points, ce qui a pour effet de diminuer l'adhérence ainsi que les frottements. La deuxième position allonge le contact au maximum, la troisième condense le contact et restreint la longueur du robot au minimum. La dernière position élimine le contact des chenilles secondaires avec le sol. D'un autre côté les flippers permettent une meilleure manœuvre du robot lors de son déplacement sur des terrains accidentés, ainsi que la capacité de franchir les rampes et les petits obstacles. Autre avantage que présente ce concept est celui de pouvoir réaliser un mouvement de rotation autour de l'axe de révolution vertical passant par le centre de gravité du robot, ce qui lui confère une plus grande mobilité et agilité dans des environnements à espace réduit ou encombré.

Les dimensions du robot sont détaillées sur les figures ci-dessous.

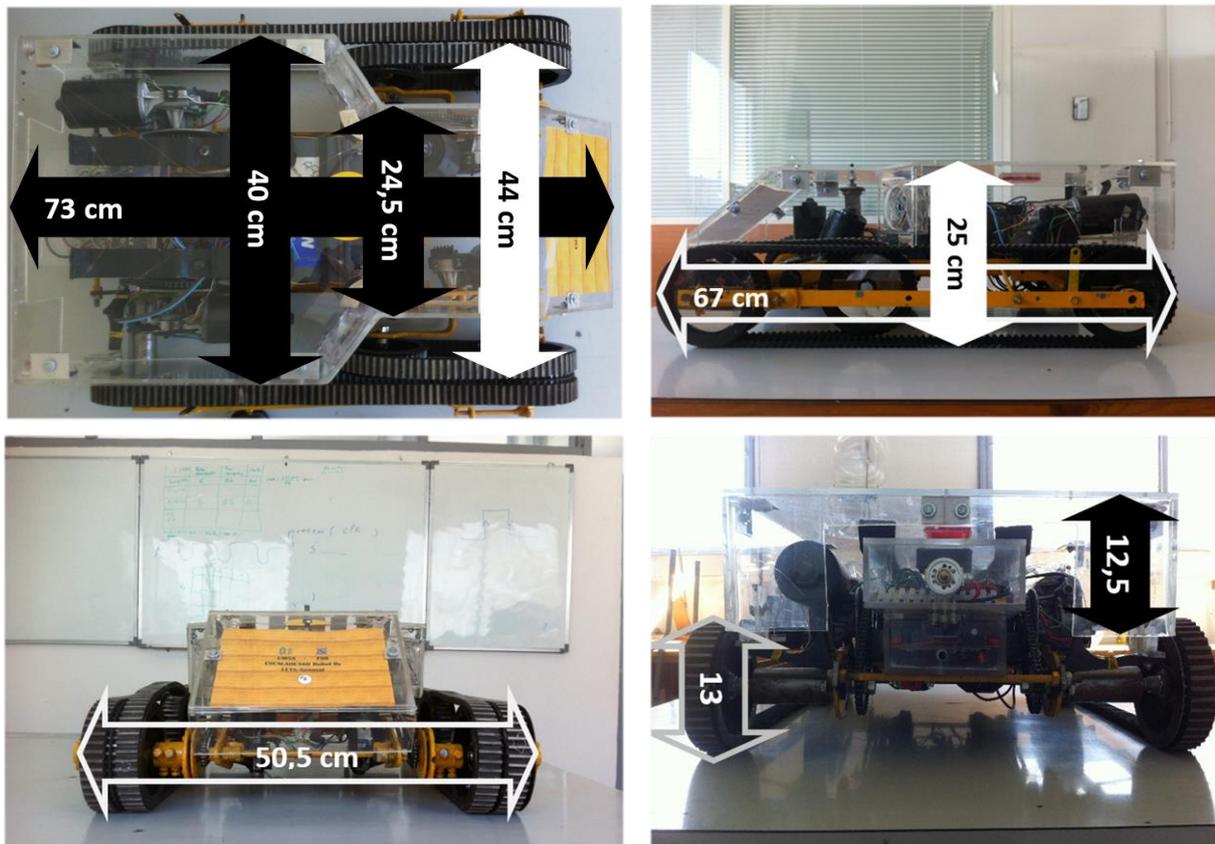


Figure 4.14 : Les dimensions du robot ESCALADE360.

4.2.3 Application : suivi d'une cible en mouvement

Le suivi visuel est un thème prévalant dans le domaine de la vision artificielle. Son utilisation est souvent indispensable pour des applications telles que la surveillance, la robotique ou la réalité augmentée. La littérature compte de nombreuses méthodes de suivi, La plupart de ces méthodes extraient les caractéristiques des cibles basées sur des indices simples telles que la couleur, la

texture, la forme ou le contour. Dans [125] La méthode de suivi visuel développée est basée sur les caractéristiques couleurs de l'objet à suivre par une caméra catadioptrique. Pour remédier aux problèmes de changement d'illumination les auteurs ont jugé utile d'opter pour l'espace de représentation des couleurs "*transformed RGB*" présenté dans [126]. L'utilisation des caméras catadioptriques offre l'avantage de réduire considérablement le nombre de capteurs de vision nécessaires pour réaliser cette tâche.

Sous le même thème, nous avons développé un algorithme de détection et de suivi d'une cible en mouvement en utilisant un système de vision omnistéréo embarqué sur un robot mobile. L'algorithme se base sur une modélisation colorimétrique de l'objet à suivre.

Représentation colorimétrique de l'objet cible : L'incommodité principale de tout algorithme de détection ou de suivi d'objet basé sur les indices couleur, est sa sensibilité aux variations de l'éclairage. En effet la représentation d'une couleur par ses coordonnées dans un espace colorimétrique tel que l'espace RGB, $L^*a^*b^*$, ou HSV inclue implicitement de l'information sur l'illumination de la couleur. Selon l'espace choisi, cette information est souvent corrélée à celle de la chrominance de la couleur. Mais il est à noter que certains espaces de représentation sont plus discriminants que d'autre. L'intensité d'une couleur se calcule en fonction de ses trois coordonnées rouge, vert et bleu dans l'espace RGB à titre d'exemple. Toute variation d'éclairage affecte les trois coordonnées R, G et B de la couleur simultanément. D'autres espaces de représentation essaient de séparer l'information sur l'intensité de la couleur de l'information sur sa chrominance. Dans l'espace HSV, la chrominance est représentée par les deux coordonnées H (*Hue*) qui la teinte de la couleur, et S (*Saturation*) qui désigne sa saturation. L'intensité est représentée par la troisième composante, qui est la valeur V.

La détection d'un objet par la primitive couleur se base essentiellement sur la détermination des bornes des intervalles délimitant les coordonnées de sa couleur à différentes conditions d'éclairage dans un espace de représentation donné. Malheureusement, les espaces précédemment cités ne présentent pas une robustesse satisfaisante face au problème de la variation d'éclairage de la scène dans des conditions réelles d'utilisation. Cela se traduit par des fausses détections ou par des faux rejets. Par conséquent, nous avons choisi de travailler sur un espace modifié de représentation de la couleur, basé sur l'espace HSV.

Pour déterminer les caractéristiques colorimétriques de l'objet cible, il est initialement placé dans un environnement à faible éclairage (figure 4.15.a). Une

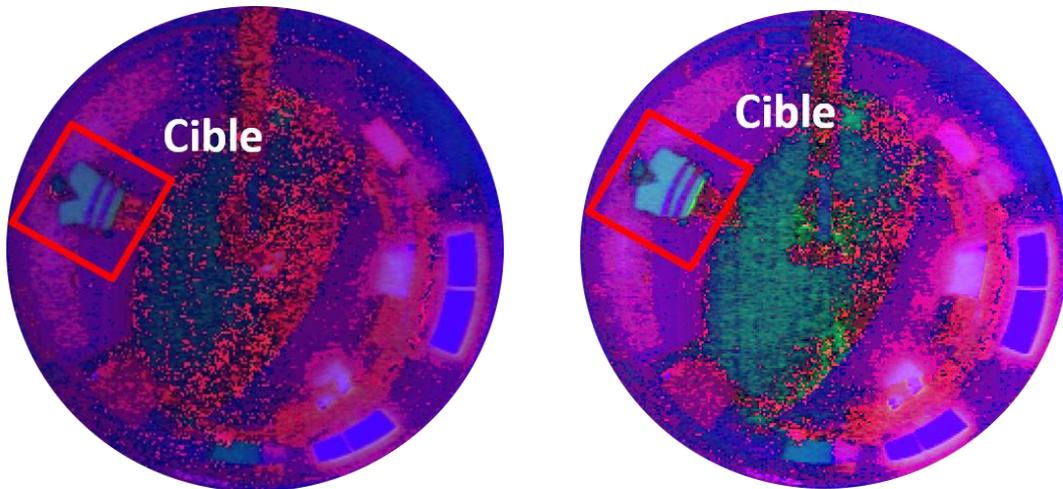


Figure 4.15 : La cible suivie. (a) Image catadioptrique HSV à faible éclairage. (b) Image catadioptrique HSV à fort éclairage.

fois l'image catadioptrique est acquise les coordonnées (h_i, s_i, v_i) avec $i = 1, 2 \dots 5$ de cinq pixels de même couleur de l'objet cible sont enregistrés. Le même procédé est exécuté dans des conditions d'un éclairage fort, où cinq autres coordonnées (H_i, S_i, V_i) sont prises en compte (figure 4.15.b).

A partir de cette étape la couleur de chaque pixel sera désormais représentée par trois composantes (T_1, T_2, T_3) avec :

$$\begin{cases} T_1 = \frac{h}{s*v} \\ T_2 = \frac{s}{h*v} \\ T_3 = \frac{v}{s*h} \end{cases} \quad 4.1$$

Une moyenne des composantes de couleur des cinq pixels enregistrés à différentes conditions d'éclairage est ensuite calculée :

$$\begin{cases} T'_1 = \frac{1}{5} \sum_{i=1}^5 \frac{h_i}{s_i*v_i} \\ T'_2 = \frac{1}{5} \sum_{i=1}^5 \frac{s_i}{h_i*v_i} \\ T'_3 = \frac{1}{5} \sum_{i=1}^5 \frac{v_i}{s_i*h_i} \end{cases} \quad \text{et} \quad \begin{cases} T''_1 = \frac{1}{5} \sum_{i=1}^5 \frac{H_i}{S_i*V_i} \\ T''_2 = \frac{1}{5} \sum_{i=1}^5 \frac{S_i}{H_i*V_i} \\ T''_3 = \frac{1}{5} \sum_{i=1}^5 \frac{V_i}{S_i*H_i} \end{cases} \quad 4.2$$

Une seconde moyenne globale de ces composantes est calculée comme suit :

$$\begin{aligned} T_1 &= (T'_1 + T''_1)/2 \\ T_2 &= (T'_2 + T''_2)/2 \\ T_3 &= (T'_3 + T''_3)/2 \end{aligned} \quad 4.3$$

T'_1 , T'_2 et T'_3 représentent les composantes de la couleur à faible conditions d'éclairage, tant dit que T''_1 , T''_2 et T''_3 sont les composantes à fortes conditions d'éclairage. T_1 , T_2 , et T_3 sont les composantes moyennes qui seront utilisées pour la détection de l'objet.

Détection de l'objet cible : Une fois les paramètres de couleur de référence sont définis, l'étape suivante permettra au robot de détecter et naviguer vers la cible d'une façon autonome. Les images catadioptriques acquises sont parcourues en temps réel. Pour chaque pixel p_i de l'image nous calculons ses trois composantes colorimétriques λ_{i1} , λ_{i2} , et λ_{i3} données ci-dessous, où $i = 1, 2, \dots, N$ avec N le nombre total de pixels dans l'image.

$$\lambda_{i1} = \frac{h_i}{s_i * v_i}, \lambda_{i2} = \frac{s_i}{h_i * v_i}, \lambda_{i3} = \frac{v_i}{s_i * h_i} \quad 4.4$$

Si les trois conditions suivantes sont respectées par un pixel p_i de l'image, il sera considéré comme un point de la cible choisie :

$$\begin{cases} |\lambda_{i1} - T_1| < \alpha \\ |\lambda_{i2} - T_2| < \beta \\ |\lambda_{i3} - T_3| < \gamma \end{cases} \quad 4.5$$

Avec : $\alpha = |T'_1 - T''_1|$; $\beta = |T'_2 - T''_2|$; $\gamma = |T'_3 - T''_3|$

Pour chaque pixel détecté de l'objet cible, son angle d'orientation θ_i par rapport au repère du robot est calculé. Le barycentre de la cible observée est ensuite déterminé comme suit :

$$\theta_g = \sum_{i=1}^n \theta_i / n \quad 4.6$$

Où n est le nombre total de pixels cibles détectés.

A partir de cette étape le robot peut s'orienter vers l'objet et naviguer vers sa direction comme illustré sur la figure 4.16.

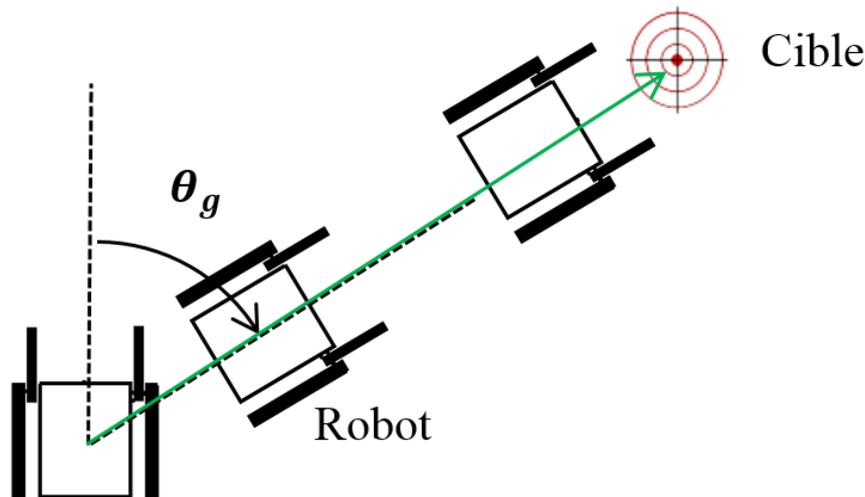


Figure 4.16 : Détection et orientation du robot vers la cible suivie.

Estimation de la profondeur : Le robot étant équipé d'un système de vision omnistéréo, il lui est possible d'estimer la profondeur de l'objet suivi par la même triangulation présentée dans la section 2.4.3 :

$$D = \varepsilon + \frac{(h-\lambda) \times \tan(2\varphi_1) \times \tan(2\varphi_2)}{\tan(2\varphi_1) - \tan(2\varphi_2)} \quad 4.7$$

L'estimation de la profondeur est une étape qui se fait juste après la détermination de l'orientation de la cible. Une fois la distance séparant l'objet du robot est calculée, la condition d'arrêt du robot est désormais définie. Notons que l'algorithme de suivi opère en temps réel ; l'orientation de la cible ainsi que la condition d'arrêt sont rafraîchis à chaque image (*frame*) traitée du flux d'acquisition.

Résultats : Comme expérience, une personne portant une couleur bien distincte des objets de la scène se déplace aux alentours du robot (figure 4.17). Ce dernier doit suivre la personne en mouvement tant qu'elle est perceptible. Les expériences ont été menées dans des conditions d'éclairage naturel et artificiel. Nous avons remarqué de forts écarts des résultats de détection entre les méthodes basées sur la représentation colorimétrique classique et la méthode développée basée sur l'espace HSV modifié. Des changements importants dans l'illumination de la cible au cours du temps sont également notables, ainsi qu'aux effets du reflet. Dans ces expériences, les méthodes de détection conventionnelles ne sont pas susceptibles de suivre correctement l'objet durant la totalité de l'expérience.

Pour mettre en exergue la précision du système de vision omnistéréo une série de tests ont été menés. La deuxième phase d'expérimentation consistait à étudier la variation de l'erreur générée par le système en fonction de la variation de la profondeur. Pour une longueur fixe h de la ligne de base, l'erreur est évaluée pour différentes distances D . Le tableau 4-1 rapporte les résultats obtenus.

Tableau 4-1 : Variation de l'erreur en fonction de D , $h=0.425m$.

| h (m) | 0.425 | | | |
|------------------------|-------------|------|-------------|------|
| Profondeur réelle (m) | 1 | 1.97 | 2.95 | 3.97 |
| Profondeur estimée (m) | 0.92 | 2.05 | 2.96 | 3.91 |
| Erreur (m) | 0.08 | 0.08 | 0.01 | 0.06 |
| Erreur (%) | 8.00 | 4.05 | 0.33 | 1.51 |
| Erreur moyenne (%) | 3.47 | | | |

Nous pouvons déduire à partir des résultats obtenus, qu'à une longueur de la ligne de base fixe et égale 0.425m, l'erreur sur la profondeur ne dépasse pas 8%, avec une valeur moyenne de 3.47%, ce qui reste généralement admissible. Ces résultats affirment qu'à chaque distance D donnée, correspond une unique valeur optimale

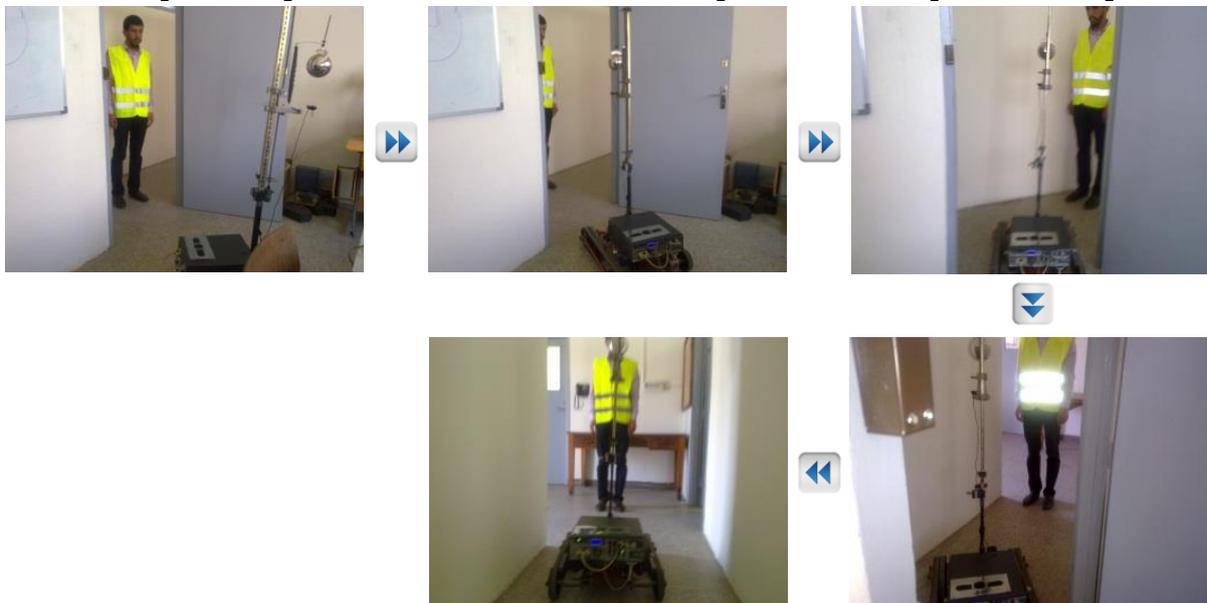


Figure 4.17 : Expérience de suivi autonome d'une cible en mouvement

de h qui minimise l'erreur générée ; en fixant h à une valeur de 0.475m on constate que l'erreur chute significativement lorsque la distance D évaluée est égale à 2.95m, car 0.475m correspond d'après le tableau 4-1 à l'optimum de h pour une distance de 3.18m.

4.3 Conception et réalisation d'un suiveur solaire

L'énergie solaire est l'une des sources d'énergie renouvelable les plus prometteuses et à croissance rapide. Cette technologie essentielle vise à remplacer la consommation finie et coûteuse de combustibles fossiles, par la conversion gratuite de la lumière naturelle en énergie électrique renouvelable et propre. De nombreuses études de recherche novatrices sont menées dans ce domaine, en utilisant des technologies d'analyse avancées afin de diminuer le coût par watt, d'accroître l'adoption, d'améliorer le rendement, et aussi essayer d'élargir le marché grâce à des applications innovantes.

Les suiveurs solaires sont des systèmes qui orientent automatiquement les collecteurs solaires comme les panneaux photovoltaïques plats, panneaux photovoltaïques à concentration (CPV) ou les systèmes solaires thermiques à concentration (CSP) vers le soleil. Ces systèmes augmentent la production d'énergie par l'optimisation de l'angle d'incidence entre le rayonnement solaire et des collecteurs solaires entrants [127]. Plusieurs concepts de suiveurs solaires sont disponibles. En général, ils sont classés en deux grandes catégories : les suiveurs passifs et ceux actifs. Les suiveurs passifs sont ceux qui n'impliquent pas d'appareil électrique, ils sont généralement basés sur l'expansion thermique de la matière (comme le fréon) quant aux suiveurs actifs, ils sont fondées sur des dispositifs électroniques. Nous pouvons également faire la distinction entre ces systèmes de suivi en ce qui concerne leurs degrés de liberté mécanique. Ceux qui ont un degré de liberté sont appelés suiveurs mono-axes tels que les suiveurs à axe unique horizontaux (HSAT), suiveurs mono-axes verticaux (VSAT) ou suiveurs mono-axes inclinés (TSAT). Les suiveurs à deux axes ont deux degrés de liberté, ce qui les rend plus efficaces. Comme a été confirmé dans [128], la production de l'énergie solaire peut être augmentée de 41% dans les régions de moyenne latitude lors de l'utilisation de systèmes de suivi dual-axes. Salah Abdallah dans [129] a réalisé une étude comparative de l'effet d'utilisation des systèmes de suivi du soleil, sur les caractéristiques courant-tension et la production d'électricité des panneaux photovoltaïques plats. Les résultats expérimentaux de son étude montrent que la puissance augmente de 15% dans le cas d'un suiveur à axe nord-sud, 34% pour un axe est-ouest, 37,5% pour les systèmes à axe vertical, et atteint 44% lors de l'utilisation d'un suiveur à double axe.

Plusieurs techniques de suivi du soleil ont été développées et évaluées dans la littérature. Tous ont des avantages et des inconvénients. Par exemple, les suiveurs passives sont très simples et ne consomment pas d'énergie électrique, mais sont relativement moins efficace et peu opérationnels à des températures basses. De bonnes critiques des moyens de suivi du soleil ont été présentés dans [130] et [131] où les inconvénients et les avantages de chacun ont été discutés. Classiquement les suiveurs solaires exploitent des photo-capteurs (figure 4.18) pour déterminer l'orientation optimale vers le soleil. Différentes configurations ont été développées et utilisées dans la littérature [132] [133] [134] [135] [136]. Cependant, ces techniques classiques présentent plusieurs inconvénients comme le manque de précision, un champ limité de suivi qui ne convient pas pour les régions de la terre de haute latitude, et la sensibilité aux conditions climatiques.

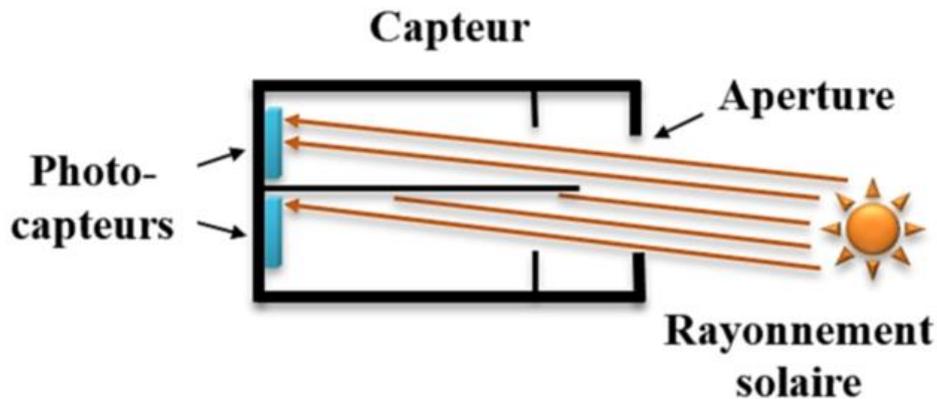


Figure 4.18 : Photo-capteur à balance de luminosité

Les photo-capteurs utilisent l'effet d'ombrage pour déterminer l'orientation optimale vers le soleil. Lorsque le rayonnement solaire n'est pas suffisamment intense pour créer assez de différence de luminosité, ils deviennent inopérants. Il convient de noter que ces capteurs ont un champ étroit de suivi de 180° au maximum. Quelques solutions ont été présentées dans la littérature pour étendre le champ de suivi. Dans [137] les auteurs ont présenté une conception simple d'un système de suivi du soleil avec un champ de suivi de 360° , malheureusement, le concept proposé est un suiveur mono-axes, ce qui a limité sa performance.

Récemment, une nouvelle tendance de recherche tente d'exploiter la vision artificielle dans le domaine des énergies renouvelables. Dans [138] les auteurs ont développé un système de commande automatique de correction et compensation du positionnement d'héliostat en utilisant des techniques de vision directionnelle artificielle. Dans cette même perspective nous avons proposé dans [139] un système autonome de suivi du soleil à deux axes, basé sur le traitement de l'image et les techniques de vision omnidirectionnelle. Le champ de vision du dispositif d'acquisition couvre 360° à l'horizontale (θ) et de 200° à la verticale (φ). La nécessité d'un large champ de vision est justifiée étant donné que le chemin de l'azimut solaire dépasse 120° dans de nombreuses régions de la terre. À Rabat (Maroc), par exemple, la trajectoire du soleil varie de 120° à 240° d'Est en Ouest de l'hiver à l'été. (figure 4.19).

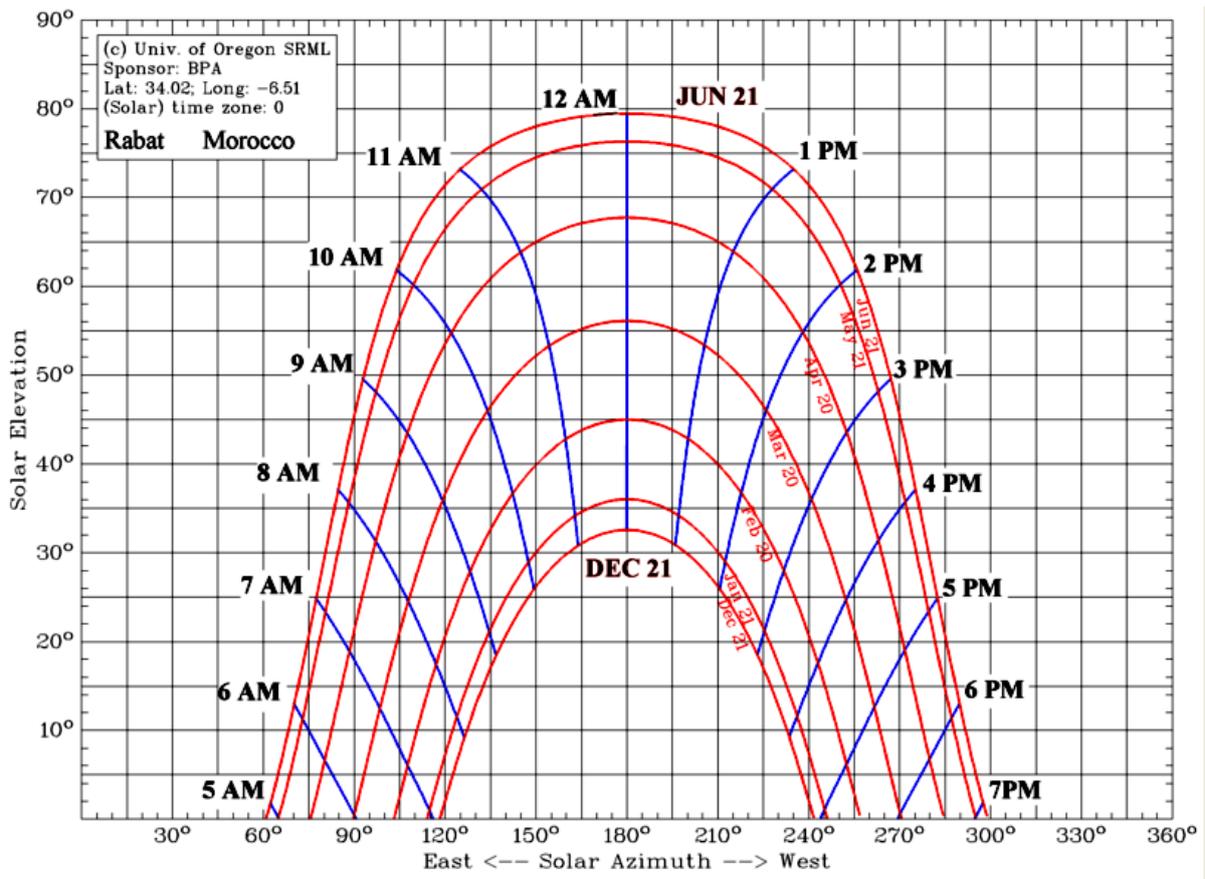


Figure 4.19 : Carte de la trajectoire solaire de Rabat (Maroc).

Par conséquent, en offrant un large champ de vision, une caméra omnidirectionnelle peut être très utile pour la détection et le suivi du soleil. En outre, contrairement aux capteurs classiques, les caméras catadioptriques sont en

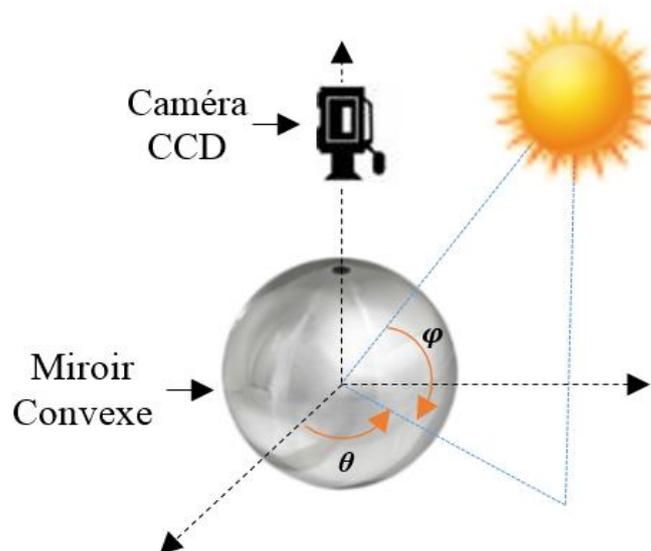


Figure 4.20 : Illustration de la localisation du soleil avec une caméra catadioptrique.

mesure de fournir des informations précises sur la position du soleil à tout moment de la journée. Pratiquement, la position du soleil est pourvue en coordonnées sphériques (θ, φ) extraites des images omnidirectionnelles acquises, comme le montre la figure 4.20.

4.3.1 Le suiveur solaire omnidirectionnel : concept et contrôle

Pour étudier les différences, les avantages et les points faibles du concept du suiveur solaire proposé, un prototype expérimental nommé Helianthus360 a été conçu et réalisé. Il se compose de quatre modules principaux : l'unité d'acquisition représentée par la caméra catadioptrique, l'unité de commande, le mécanisme de rotation à double axes, et un panneau photovoltaïque. La figure 4.21 ci-dessous illustre l'architecture de Helianthus360. Ce suiveur peut atteindre 360 degrés de mouvement d'azimut et 90 degrés d'élévation.

Le module d'acquisition de la position du soleil est une caméra catadioptrique composée d'une caméra CCD standard et d'un miroir sphérique d'un rayon de 30 mm. Ce système de vision fournit des images quasi-hémisphériques du ciel, comme le montre la figure 4.24. Ces images sont analysées par l'unité de traitement pour extraire les coordonnées sphériques du spectre solaire, associée au repère système catadioptrique. Une fois que l'action corrective est générée, l'unité de traitement envoie des instructions au circuit de commande (figure 4.23) via une liaison RS232.

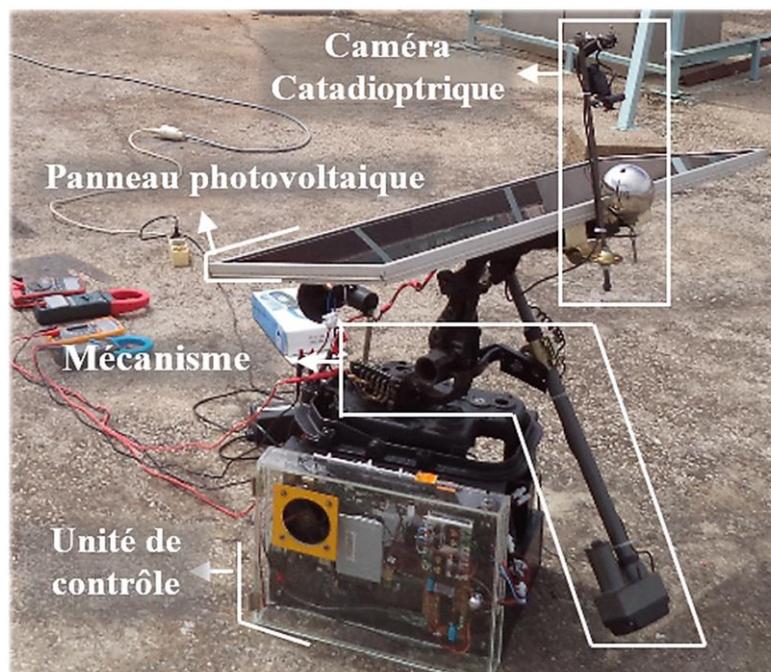


Figure 4.21 : Suiveur solaire Helianthus360.

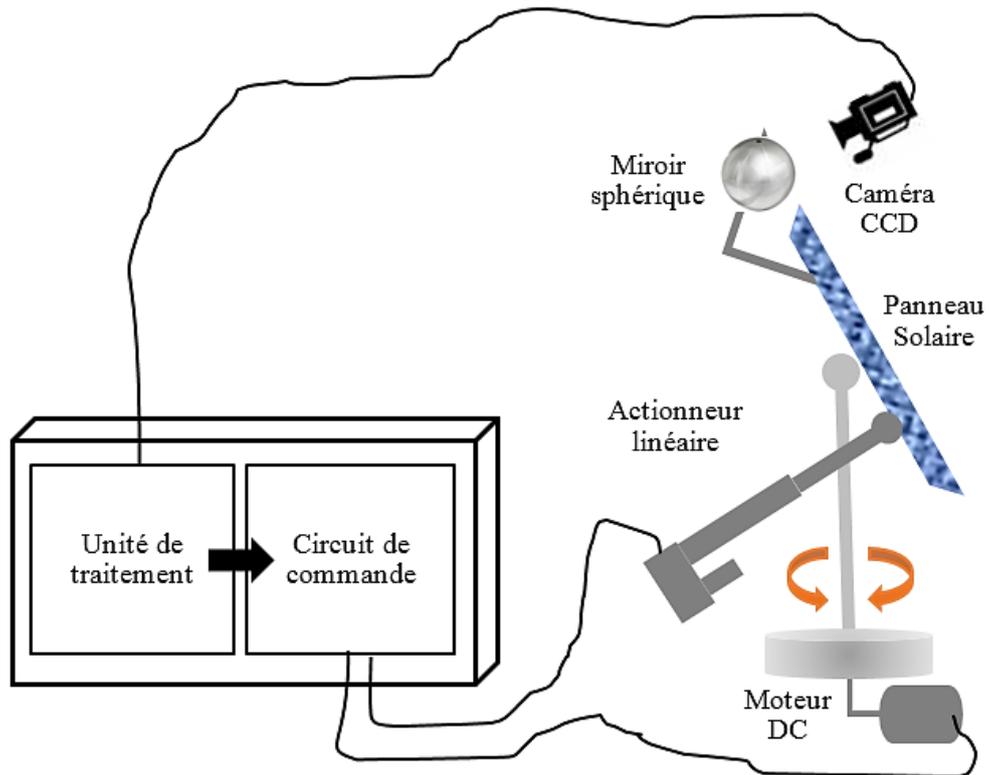


Figure 4.22 : L'architecture du suiveur solaire proposé.

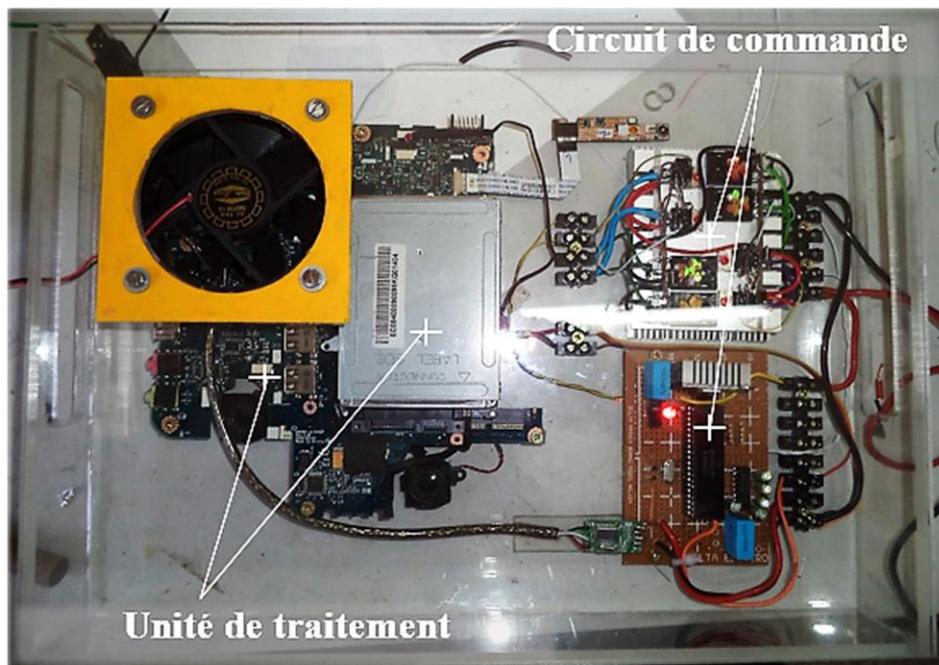


Figure 4.23 : Unité de contrôle du suiveur solaire Helianthus360.

Celui-ci traite les actions de contrôle et de conduit les moteurs à réaliser les mouvements désirés. Le circuit de commande est composé essentiellement d'un double pont H de transistors de puissance, et d'un microcontrôleur. Ce circuit se charge d'un rôle intermédiaire entre l'unité de traitement et les actionneurs mécaniques.

Comme décrit précédemment, la procédure de détection et suivi du soleil proposée, est basée principalement sur des algorithmes de traitement d'image. En effet, une fois l'image omnidirectionnelle est acquise, il est convertie de l'espace de couleur RGB vers l'espace HSV, donc l'image sera représentée par trois matrices ; la teinte (H), la saturation (S) et la valeur (V). Selon la matrice V qui représente la luminosité des pixels de l'image, L'étape suivante consiste à déterminer la valeur maximale de l'intensité des pixels. Ensuite, une étape de seuillage est réalisée afin de générer une image binaire (figure 4.24.b), qui caractérise le spectre solaire, ce procédé est schématisé sur la figure 4.24.c.

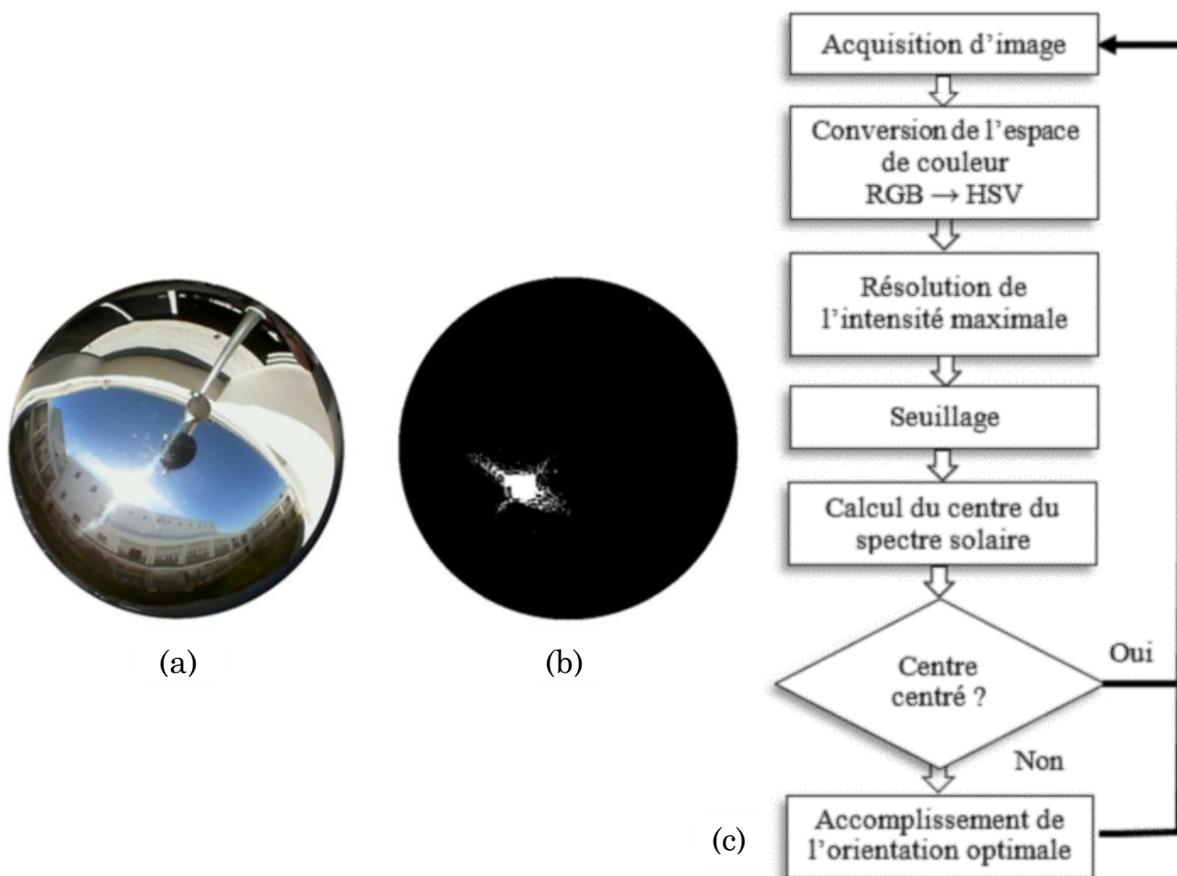


Figure 4.24 (a) Image omnidirectionnelle échantillon d'un ciel partiel. (b) L'image omnidirectionnelle après seuillage. (c) Schématisation de la procédure de suivi du soleil.

Les coordonnées du barycentre du spectre solaire sont calculées selon les équations suivantes :

$$\theta_c = \sum_{i=1}^n \theta_i / n \quad 4.8$$

$$\varphi_c = \sum_{i=1}^n \varphi_i / n \quad 4.9$$

θ_i et φ_i sont les coordonnées sphériques de chaque pixel du spectre solaire dans l'image binaire selon le repère du système imageur catadioptrique (voir figure 4.20). Quand n est le nombre total de ces pixels. Désormais, la position du soleil est identifiée par les coordonnées θ_c et φ_c du barycentre. L'idée fondamentale de la procédure de suivi du soleil est de réaliser les mouvements mécaniques nécessaires, simultanément selon l'azimut et l'élévation pour ramener le barycentre du spectre solaire au centre de l'image omnidirectionnelle. Ce qui correspond à l'orientation optimale du panneau vers le rayonnement solaire. Etant donné que l'axe optique principal de la caméra omnidirectionnelle est calibré pour être perpendiculaire au plan du panneau.

4.3.2 Validation expérimentale

Afin d'évaluer l'efficacité du suiveur solaire présenté, une étude expérimentale a été menée. Le but de cette étude est de comparer la production d'énergie d'un panneau photovoltaïque embarqué sur le suiveur solaire à un autre panneau fixe. L'étude expérimentale avait débuté à partir du 17 Juillet 2013 au 20 Septembre 2013. Les deux panneaux photovoltaïques comparés sont identiques et génèrent une tension de 18V, avec une puissance crête de 12W. Pour éviter l'influence du niveau de charge des batteries sur les mesures du courant et de la tension, les panneaux ont été reliés directement à une charge résistive surestimé de 15W. Le panneau fixe a été orienté comme convenu vers le sud avec une élévation de 34°, ce qui correspond à la latitude de Rabat. Les mesures courant-tension ont été prises pour chaque panneau dans les mêmes conditions. La figure 4.25 montre les mesures obtenues le 17 Juillet 2013 de 11h42 à 17h41 pendant un jour partiellement nuageux. Alors que la figure 4.26 représente les résultats obtenus le 20 Septembre 2013, journée ensoleillée, de 09h30 à 17h30.

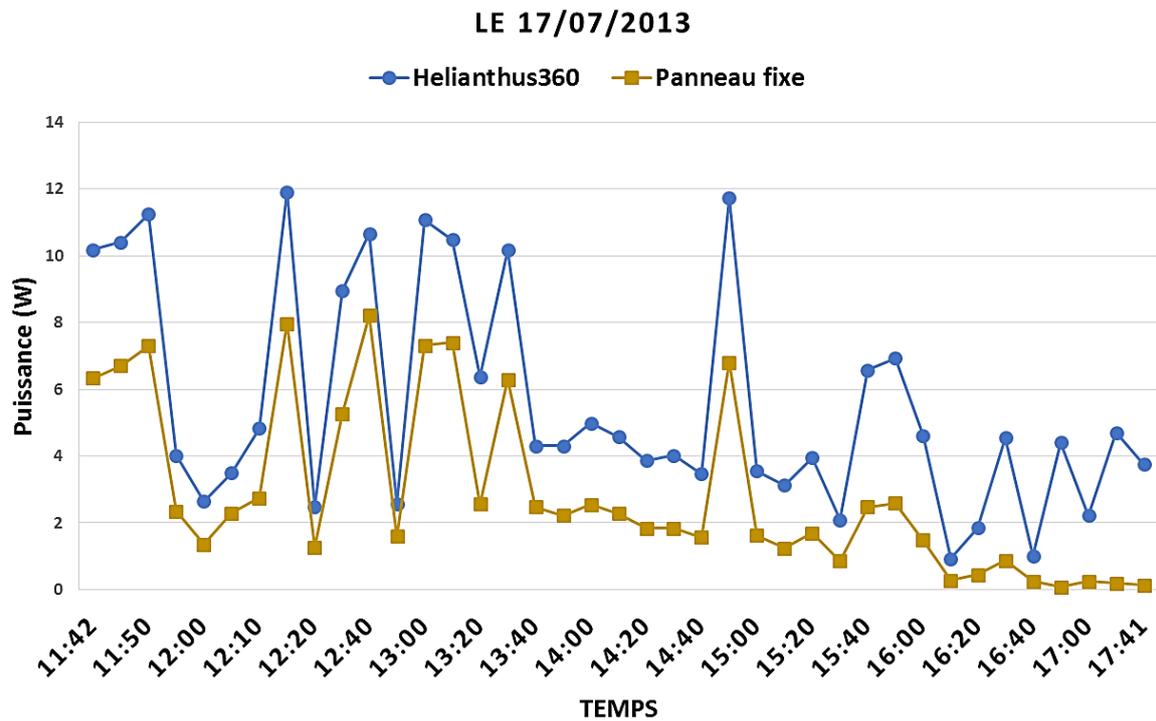


Figure 4.25 : Les mesures de puissance du 17 Juillet 2013.

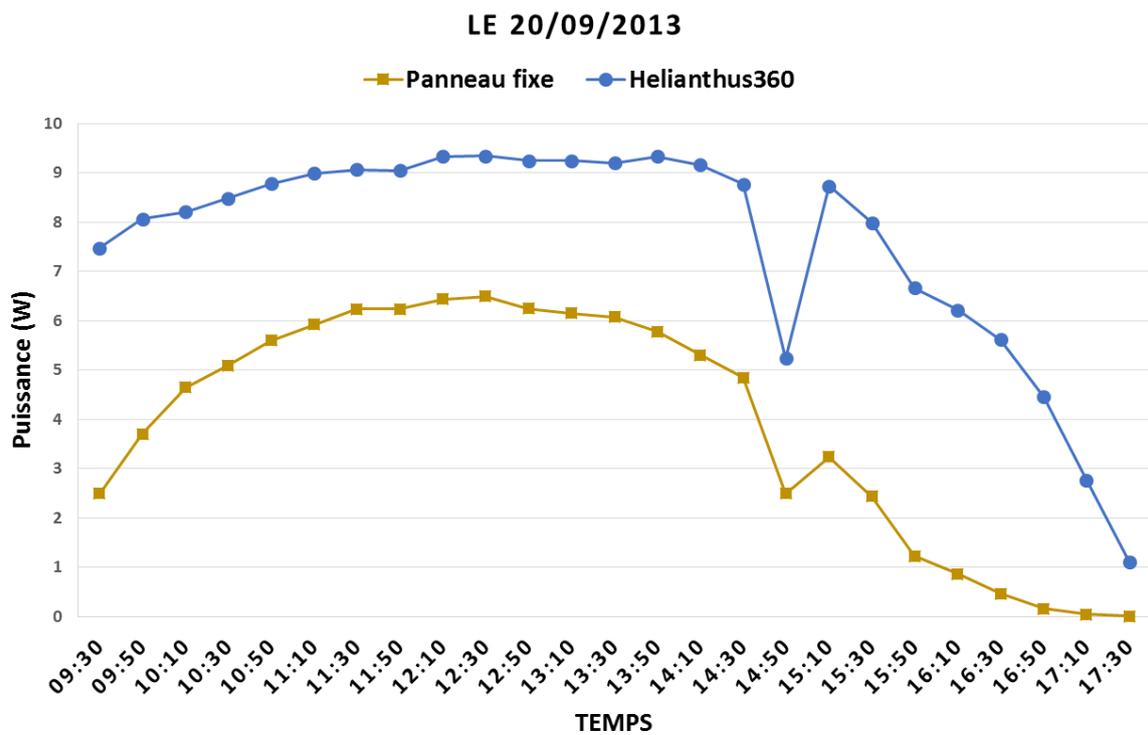


Figure 4.26 : Les mesures de puissance du 20 Septembre 2013.

Le gain de production d'énergie E_h du suiveur solaire étudié est donné par l'équation suivante :

$$E_h = \frac{P_h - P_f}{P_f} \quad 4.10$$

Où P_h est la puissance générée par Helianthus360, P_f est la puissance de panneau fixe.

On constate les mesures prises que la puissance électrique délivré par le suiveur solaire est constamment supérieure à celle du panneau fixe. En fait, au cours du premier jour de l'expérimentation (figure 4.27.a) le gain minimal de puissance de Helianthus360 était de 30% avec un gain journalier moyen de 92%. L'effet d'un ciel clair était remarquable durant la deuxième journée de l'étude expérimentale (figure 4.27.b) où le minimum du gain E_h était de 43,79% et le gain journalier moyen a atteint 94%.

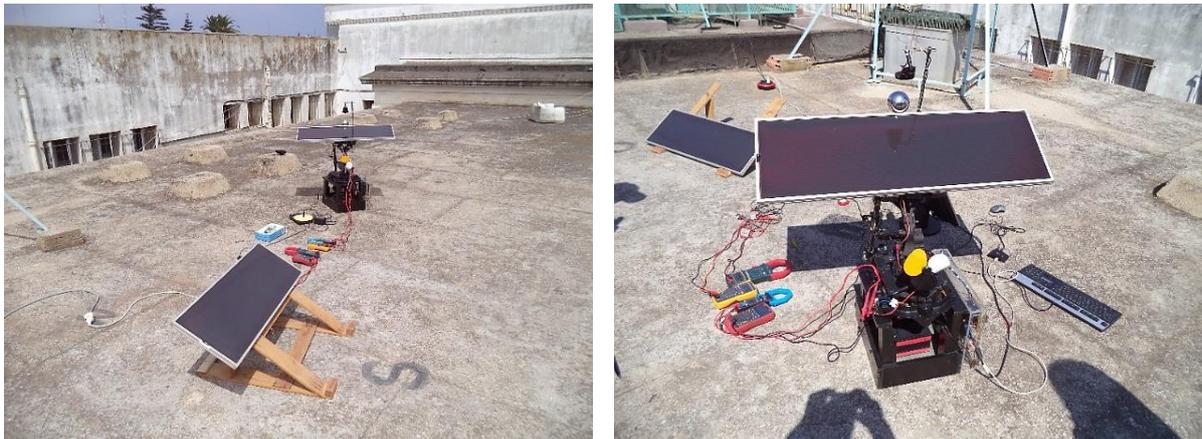


Figure 4.27 : (a) Première expérience 17 Juillet 2013. L'enseuillement total journalier: 25.7 MJ/m² (b) Deuxième expérience 20 septembre 2013. L'enseuillement total journalier: 21 MJ/m².

4.3.3 Discussion des résultats

À partir des résultats expérimentaux présentés, on peut voir que le gain en production d'énergie en utilisant le système de poursuite du soleil proposé est considérable. En effet ce gain est augmenté de 30% au moins et peut atteindre 94%. Il est bien connu que chaque suiveur solaire actif consomme de l'énergie comme il optimise sa production. Cette consommation d'énergie est dissipée par les unités de traitement et les actionneurs mécaniques. Dans le cas du système proposé, la consommation de l'unité de traitement est d'environ 12W, et même si le suivi est effectué en temps réel, les deux moteurs à courant continu sont actionnés uniquement pendant environ 3 secondes toutes les 30 minutes. Donc leur consommation d'énergie est négligeable. Cependant, dans le but d'une étude comparative précise et objective, l'unité de traitement et les moteurs ont été alimentés par une source d'énergie externe. Il est évident que le système proposé

n'est efficace que si l'augmentation de la production d'énergie est supérieure à la consommation interne d'énergie, un compromis qui doit être assumé lors de la phase de dimensionnement et conception du système. Le système proposé peut être amélioré en employant un petit ordinateur de bord avec une consommation d'énergie moindre.

Bien sûr, ce suiveur solaire sera d'un grand intérêt dans le cas d'une unité de production d'énergie à l'échelle industrielle. En effet, dans ce cas, l'investissement initial est sans doute important, et nécessite une rentabilité maximale. Pratiquement, le concept proposé peut être étendue afin d'optimiser la production d'une centrale d'énergie solaire photovoltaïque. Eu égard à la technique de suivi, qui combine des modules distincts à savoir de l'unité d'acquisition représenté par la caméra catadioptrique, le module de traitement, et le module mécanique, l'extension de son utilisation dans le cas de plusieurs panneaux n'a besoin que d'un unique module d'acquisition et de traitement, et une adaptation du système mécanique. Cette adaptation peut être effectuée selon deux concepts : le premier est de relier plusieurs panneaux solaires mécaniquement, dans ce cas, seuls deux moteurs seront nécessaires pour assurer le mouvement d'élévation et de l'azimut qui sont les deux degrés de liberté nécessaires. Le second concept que nous proposons est de contrôler les panneaux solaires séparément avec deux actionneurs mécaniques montés sur chaque panneau. Il convient de préciser que le choix d'un concept à un autre dépend des caractéristiques de la centrale solaire prévue, ces caractéristiques s'appliquent à la puissance unitaire des panneaux sélectionnés, la taille, le poids et la géométrie du parc.

Les unités mobiles telles que les robots (figure 4.28), les caravanes, des unités médicales mobiles, et les bateaux équipés de panneaux solaires, peuvent certainement tirer de grands avantages de l'utilisation du suiveur solaire développé. En effet, il a trois principaux avantages en termes d'optimisation de l'efficacité de la production d'énergie, par rapport à d'autres techniques de suivi : la première est que le processus de suivi ne nécessite aucune initialisation supervisée, grâce au large champ de vision de son module d'acquisition, la position du soleil peut être située à n'importe quel moment. Le deuxième avantage à noter, c'est que le processus de suivi se fait indépendamment des coordonnées spatiotemporelles. Le dernier concerne l'impact remarquable sur l'efficacité énergétique du système proposé à l'issue d'un suivi en temps réel.



Figure 4.28 : Robot mobile équipé d'un suiveur solaire.

4.4 Conclusion

Ce chapitre a concerné l'intégration des capteurs de vision omnidirectionnelle aux systèmes autonomes. Nous avons étudié l'intégration de tels systèmes de vision artificielle dans le domaine de la robotique mobile. Le robot prototype développé nous a permis de vérifier la validité des algorithmes de traitement et de stéréovision développés. Nous pouvons conclure que les systèmes de vision catadioptriques permettent de réduire considérablement la complexité de conception électronique et algorithmique des systèmes autonomes, et de fournir une alternative très riche, et plus économique.

Afin d'assurer au robot développé une autonomie en énergie, nous avons pensé à l'équiper d'un module d'énergie renouvelable propre et efficace. Le système proposé est un suiveur solaire destiné à améliorer le rendement énergétique d'une installation photovoltaïque, en optimisant l'orientation des panneaux vers le soleil. L'utilisation d'une caméra catadioptrique a permis d'élargir considérablement le champ de suivi du spectre solaire. À la différence des suiveurs classiques le processus du suivi est indépendant de la référence spatiotemporelle du site. L'utilisation des techniques de traitement des images omnidirectionnelles a réduit la sensibilité du système aux conditions climatiques. En effet le spectre solaire étant partiellement occlus, ou même dans le cas d'un ciel nuageux, le suiveur pourra toujours orienter les panneaux vers la zone la plus éclairci du ciel, tant qu'il

en a une. L'effet de la précision du système était notable sur le rendement énergétique d'une installation photovoltaïque. Nous avons noté un gain en puissance qui peut atteindre 94%. L'extension d'un tel concept à des installations moyennes de production d'énergie solaire, ou même à échelle industrielle sera sans doute amplement bénéfique.

Conclusions

Dans cette thèse nous nous sommes intéressés à la stéréovision omnidirectionnelle ainsi qu'au traitement des images catadioptriques. La thèse commence par une introduction générique à la vision omnidirectionnelle, y compris une description des techniques utilisées pour l'obtention de larges champs de vision. L'étude montre que les caméras catadioptriques offrent un compromis optimal entre le coût et les performances.

Nous présentons une classification des technologies des capteurs d'acquisition 3D. Parmi lesquelles, les systèmes actifs et ceux passifs. L'un des apports de cette étude est l'évaluation comparative des caméras catadioptriques et des systèmes omnistéréo qui sont classés selon les critères suivants: la technologie de fabrication, la précision d'estimation de la profondeur, la résolution, la propriété du centre de projection unique, et l'isotropie. L'une des principales contributions de cette thèse est l'introduction d'un modèle formel d'un nouveau capteur omnidirectionnel avec des capacités de perception de profondeur. Il consiste en un capteur omnistéréo, à camera unique et à disposition verticale. Cette configuration avait été adoptée vu les avantages qu'elle présente par rapport aux autres, en termes d'occlusions et de géométrie épipolaire. La précision du système avait été évaluée expérimentalement. Un concept à longueur de ligne de base ajustable tel que celui proposé, offre l'avantage d'agir sur la précision du capteur en fonction des conditions de son utilisation.

Traiter les images catadioptriques par des outils conçus pour les images planes classiques aboutit à des erreurs, puisque la géométrie de leur formation n'est pas prise en compte. En effet, une image catadioptrique présente une résolution non-uniforme et de fortes distorsions radiales. Comme contribution aux traitements de moyen-niveau de ce type d'images, nous proposons de projeter l'image catadioptrique sur un espace adéquat, qui est la sphère d'équivalence, et de procéder aux traitements de l'image obtenue par des outils adaptés. Nous avons développé un détecteur de contours basé sur un modèle sphérique de charges électriques virtuelles, ainsi qu'une adaptation du détecteur de coins de Harris aux images omnidirectionnelles. L'évaluation comparative aux autres opérateurs classiques, a prouvé clairement l'amélioration apportée par l'approche proposée.

Les techniques de vision omnidirectionnelle recueillies ont été finalement utilisées dans deux applications. La première application concernait l'intégration de tels systèmes de vision dans le domaine de la robotique mobile. La conception et la réalisation d'un robot prototype, nous a permis de vérifier la validité des algorithmes de traitement et de stéréovision développés. La deuxième application consistait à concevoir un suiveur solaire à caméra catadioptrique. En effet l'intégration de cette technologie a permis d'élargir considérablement le champ de suivi du spectre solaire. L'évaluation expérimentale du système a validé l'impact de sa précision sur l'amélioration du rendement énergétique.

Références

- [1] N. Franceschini, J. M. Pichon et C. & Blanes, «Real time visuomotor control: from flies to robots,» *In Advanced Robotics, 1991. Robots in Unstructured Environments', 91 ICAR., Fifth International Conference on (pp. 931-935). IEEE., 1991.*
- [2] H. Afshari, L. Jacques, L. Bagnato, A. Schmid, P. Vandergheynst et Y. & Leblebici, «The PANOPTIC camera: a plenoptic sensor with real-time omnidirectional capability,» *Journal of Signal Processing Systems, 70(3), 305-328, 2013.*
- [3] A. Basu et S. Licardie., «Multi-camera networks: Eyes from eyes,» *In: IEEE/RSJ, Intelligent Robots and Systems. Vol. 3. pp. 1822 –1828, 1993.*
- [4] D. Slater, «Panoramic Photography with Fisheye Lenses,» *Panorama, Vol. 13, No. 2, 3, International Association of Panoramic Photographers, 1996.*
- [5] S. Peleg et M. & Ben-Ezra, «Stereo panorama with a single camera,» *In Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on. (Vol. 1). IEEE, 1999.*
- [6] C. Fermuller, Y. Aloimonos, P. Baker, R. Pless et J. N. a. B. Stuart, «Multi-camera networks: Eyes from eyes,» *In: Third Workshop on Omnidirectional Vision. pp. 11–18, 2000.*
- [7] R. Patil, P. E. Rybski, T. Kanade et M. M. & Veloso, «People detection and tracking in high resolution panoramic video mosaic,» *In Intelligent Robots and Systems, (IROS 2004). Proceedings. IEEE/RSJ International Conference , 2004.*
- [8] C. Geyer et a. K. Daniilidis, «A unifying theory for central panoramic systems and practical applications,» *In: IEEE Sixth European Conference on Computer Vision. pp. 445–461, 2000.*
- [9] D. Rees, «Panoramic television viewing system,» *United States Patent (3,505,465), 1970.*
- [10] Y. Yagi et a. S. Kawato, «Panoramic scene analysis with conic projection,» *In: IEEE/RSJ International Conference on Intelligent Robots and Systems. Vol. 1.pp. 181 – 187, 1990.*
- [11] N. Winters, J. Gaspar, G. Lacey et J. Santos-Victor, «Omni-directional vision for robot navigation,» *Proceedings of IEEE Workshop on Omnidirectional Vision, pp. 2128, 2000.*

- [12] A. Rituerto, L. Puig et J. J. Guerrero, «Comparison of omnidirectional and conventional monocular systems for visual slam,» *10th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras OMNIVIS 2010, Zaragoza, Spain.*, 2010.
- [13] M. I. Skolnik, « Introduction to Radar,» *Radar Handbook*, 1990.
- [14] M. Kobilarov, G. Sukhatme, J. Hyams et P. & Batavia, «People tracking and following with mobile robot using an omnidirectional camera and a laser.,» *ICRA 2006. Proceedings 2006 IEEE International Conference on Robotics and Automation*, 2006.
- [15] S. Thrun, M. Bennewitz, W. Burgard, A. Cremers, F. Dellaert, D. Fox, D. Haehnel, N. R. C. Rosenberg, J. Schulte et and D. Schulz, «Minerva: A second generation mobile tour-guide robot,» *in Proc. of the IEEE International Conference on Robotics and Automation*.
- [16] J. F. Engelberger, «Health-care robotics goes commercial: The helpmate experience,» *in Robotica, vol. 11,1993 , pp. 517–523*, 1993.
- [17] L. Davis, V. Philomin et a. R. Duraiswami, «Tracking humans from a moving platform,» *in ICPR '00: Proceedings of the International Conference on Pattern Recognition (ICPR'00)-Volume 4. Washington, DC, USA: IEEE Computer Society, 2000, p. 4171*], 2000.
- [18] J. [Miura, Y. Negishi et Y. & Shirai, «Mobile robot map generation by integrating omnidirectional stereo and laser range finder.,» *In Intelligent Robots and Systems, 2002. IEEE/RSJ International Conference on (Vol. 1, pp. 250-255. IEEE. , 2002*.
- [19] D. Anguelov, C. Dulong, D. Filip, C. Frueh, S. Lafon, R. Lyon et J. ... & Weaver, «Google street view: Capturing the world at street level.,» *Computer, 43(6), 32-38*], 2010.
- [20] P. J. Besl, «Active, optical range imaging sensors.,» *Machine vision and applications, 1(2), 127-152. , 1988*.
- [21] S. M. Nejad et S. Olyaei, «Comparison of TOF, FMCW and Phase-Shift Laser Range-Finding Methods by Simulation and Measurement,» *Quarterly Journal of Technology & Education Vol.1, No.1 Autumn 2006 , 2006 .*
- [22] V. Frati et D. & Prattichizzo, «Using Kinect for hand tracking and rendering in wearable haptics,» *In World Haptics Conference (WHC), 2011 IEEE (pp. 317-321). IEEE, 2011*.

- [23] O. Romain, C. GASTAUD et &. GARDA, «Un capteur de vision multi spectrale omnidirectionnelle: conception, calibrations et exploitation,» 2005.
- [24] R. Orghidan, E. M. Mouaddib et J. & Salvi, «Omnidirectional depth computation from a single image,» *In Robotics and Automation. ICRA 2005. Proceedings of the 2005 IEEE International Conference on* (pp. 1222-1227). *IEEE*, 2005.
- [25] R. Orghidan, J. Salvi et E. M. & Mouaddib, « Calibration of a structured light-based stereo catadioptric sensor.,» *In Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on* (Vol. 7, pp. 70-70). *IEEE*, 2003.
- [26] S. C. B. Yi et N. & Ahuja, «Real-time omni-directional distance measurement with active panoramic vision,» *International Journal of Control Automation and Systems*, 5(2), 184, 2007.
- [27] R. A. Jarvis, « A perspective on range finding techniques for computer vision.,» *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (2), 122-139. , 1983.
- [28] F. Blais, M. Picard et G. Godin, «Accurate 3D acquisition of freely moving objects,» *Proceedings. 2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004*, pp.422-429 .
- [29] http://hds.leica-geosystems.com/downloads123/hds/hds/HDS7000/brochures-datasheet/HDS7000_DAT_en.pdf.
- [30] M. Pollefeys, L. Gool, V. M., V. F., C. K. et J. and Tops, «Visual Modeling with a Hand-Held Camera,» *International Journal of Computer Vision (IJCV)*, vol. 59, pp. 207-232, 2004 , 2004 .
- [31] C. Geyer et K. Daniilidis, «Structure and motion from uncalibrated catadioptric views.,» *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (Vol. 1, pp. I-279). *IEEE*, 2001.
- [32] D. Strelow, J. Mishler, S. Singh et H. & Herman, «Extending shape-from-motion to noncentral omnidirectional cameras,» *Intelligent Robots and Systems, 2001. Proceedings. 2001 IEEE/RSJ International Conference on* (Vol. 4, pp. 2086-2092). *IEEE*, 2001.
- [33] B. Lucas et a. T. Kanade., «An iterative image registration technique with an application to stereo vision,» *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, 1981.

- [34] K. Wong et F. Ernst, «Master thesis “Single Image Depth-from-Defocus,» *Delft university of Technology & Philips Natlab Research, Eindhoven, The Netherlands.],* 2004.
- [35] A. P. Pentland, «Depth of Scene from Depth of Field,» *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, No.4, Page(s) 523-531. ,* 1987.
- [36] F. Cozman et E. Krotkov, «Depth from scattering,» *IEEE Computer society conference on Computer Vision and Pattern Recognition, Proceedings, Pages: 801–806 ,* 1997.
- [37] B. Horn, «Obtaining Shape From Shading Information,» *The Psychology of Computer Vision, vol. 4, pp. 115-155,* 1975.
- [38] Z. Zhang, «Determining the Epipolar Geometry and its Uncertainty : A Review,» *Rapport N° 2927, Juillet 1996,* 1996.
- [39] R. Hartley et A. Zisserman, «Multiple View Geometry in Computer Vision,» *Cambridge University Press, July 31, 2000 ,* 2000.
- [40] D. Scharstein et R. & Szeliski, «A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,» *International journal of computer vision, 47(1-3), 7-42,* 2002.
- [41] S. K. Nayar, «Sphereo: Determining depth using two specular spheres and a single camera,» *In 1988 Robotics Conferences (pp. 245-254) International Society for Optics and Photonics,* 1989.
- [42] K. C. Ng et e. al., «An integrated surveillance system—human tracking and view synthesis using multiple omni-directional vision sensors,» *Image and Vision Computing 22.7 (2004): 551-561,* 2004.
- [43] T. Gandhi et M. M. & Trivedi, «Vehicle mounted wide FOV stereo for traffic and pedestrian detection,» *Image Processing, 2005. ICIP 2005. IEEE International Conference on (Vol. 2, pp. II-121). IEEE,* 2005.
- [44] J. Shimamura, N. Yokoya, H. Takemura et K. & Yamazawa, «Construction of an immersive mixed environment using an omnidirectional stereo image sensor,» *Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on (pp. 62-69). IEEE,* 2000.
- [45] H. Koyasu, J. Miura et Y. & Shirai, «Recognizing moving obstacles for robot navigation using real-time omnidirectional stereo vision,» *image, 2(b2), 1,* 2002.
- [46] J. Miura, Y. Negishi et Y. & Shirai, «Mobile robot map generation by integrating omnidirectional stereo and laser range finder,» *Intelligent*

Robots and Systems, 2002. IEEE/RSJ International Conference on (Vol. 1, pp. 250-255). IEEE, 2002.

- [47] Z. Zhu, K. D. Rajasekar, E. M. Riseman et A. R. & Hanson, «Panoramic virtual stereo vision of cooperative mobile robots for localizing 3d moving objects,» *In Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on (pp. 29-36). IEEE, 2000.*
- [48] N. Winters, J. Gaspar, G. Lacey et J. & Santos-Victor, «Omni-directional vision for robot navigation,» *In Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on (pp. 21-28). IEEE, 2000.*
- [49] S. Nayar et S. Baker., «A theory of single-viewpoint catadioptric image formation,» *International Journal of Computer Vision, 35(2) :175–196, 1999.*
- [50] S. Baker et S. K. & Nayar, «A theory of catadioptric image formation,» *In Computer Vision, 1998. Sixth International Conference on (pp. 35-42). IEEE, 1998.*
- [51] S. Baker et S. K. & Nayar, «A theory of catadioptric image formation,» *Computer Vision, 1998. Sixth International Conference on (pp. 35-42). IEEE, 1998.*
- [52] D. Southwell, B. Vandegriend et A. & Basu, «A conical mirror pipeline inspection system,» *In Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on (Vol. 4, pp. 3253-3258). IEEE, 1996.*
- [53] S. S. Lin et R. & Bajcsy, «Single cone mirror omni-directional stereo,» 2001.
- [54] S. S. Lin et R. & Bajcsy, «High resolution catadioptric omni-directional stereo sensor for robot vision.,» *In Robotics and Automation, 2003. Proceedings. ICRA'03. IEEE International Conference on (Vol. 2, pp. 1694-1699). IEEE, 2003.*
- [55] L. Spacek, «A catadioptric sensor with multiple viewpoints,» *Robotics and Autonomous Systems, 51(1), 3-15. , 2005.*
- [56] Z. Zhu, S. Yang, G. Xu, X. Lin et D. & Shi, «Fast road classification and orientation estimation using omni-view images and neural networks,» *Image Processing, IEEE Transactions on, 7(8), 1182-1197, 1998.*
- [57] D. Southwell, A. Basu, M. Fiala et J. & Reyda, «Panoramic stereo,» *In Pattern Recognition, 1996., Proceedings of the 13th International Conference on (Vol. 1, pp. 378-382). IEEE, 1996.*

- [58] M. Fiala et A. & Basu, «Panoramic stereo reconstruction using non-SVP optics,» *Computer Vision and Image Understanding*, 98(3), 363-397, 2005.
- [59] M. Ollis, H. Herman et S. & Singh, «Analysis and design of panoramic stereo vision using equi-angular pixel cameras,» *Carnegie Mellon University, The Robotics Institute*, 1999.
- [60] J. Gluckman et S. K. & Nayar, «Ego-motion and omnidirectional cameras,» *Computer Vision, 1998. Sixth International Conference on (pp. 999-1005). IEEE*, 1998.
- [61] P. Sturm, «A method for 3d reconstruction of piecewise planar objects from single panoramic images,» *In Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on (pp. 119-126). IEEE*, 2000.
- [62] Z. Zhu, «Omnidirectional stereo vision,» *Proceedings of the Workshop on Omnidirectional Vision, Budapest, Hungary*, 2001.
- [63] C. Salinas et e. al., «Catadioptric panoramic stereovision for humanoid robots,» *Robotica 30.5 (2012): 799-811*, 2012.
- [64] J. H. Kim et M. J. & Chung, «SLAM with omni-directional stereo vision sensor.,» *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on (Vol. 1, pp. 442-447). IEEE*, 2003.
- [65] M. Schönbein, H. Rapp et M. & Lauer, «Panoramic 3D Reconstruction with Three Catadioptric Cameras,» *Intelligent Autonomous Systems 12 (pp. 345-353). Springer Berlin Heidelberg*, 2013.
- [66] S. Goto et e. al., «3D environment measurement using binocular stereo and motion stereo by mobile robot with omnidirectional stereo camera,» *Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on. IEEE*, 2011.
- [67] Z. Zhu, K. D. Rajasekar, E. M. Riseman et A. R. & Hanson, «Panoramic virtual stereo vision of cooperative mobile robots for localizing 3d moving objects,» *In Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on (pp. 29-36). IEEE*, 2000.
- [68] Z. Zhu, K. D. Rajasekar et a. A. R. H. E. M. Riseman, «3D localization of multiple moving people by an omnidirectional stereo system of cooperative mobile robots,» *Technical Report TR #00-14, Computer Science Dept., UMASS-Amherst, March, 2000*, 2000.
- [69] J. H. Kim et M. J. & Chung, «SLAM with omni-directional stereo vision sensor,» *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings.*

2003 *IEEE/RSJ International Conference on (Vol. 1, pp. 442-447). IEEE, 2003.*

- [70] M. Schönbein, H. Rapp et M. & Lauer, «Panoramic 3D Reconstruction with Three Catadioptric Cameras,» *Intelligent Autonomous Systems 12 (pp. 345-353). Springer Berlin Heidelberg, 2013.*
- [71] Z. Zhu, K. D. Rajasekar, E. M. Riseman et A. R. & Hanson, «Panoramic virtual stereo vision of cooperative mobile robots for localizing 3d moving objects,» *Omnidirectional Vision, 2000. Proceedings. IEEE Workshop on (pp. 29-36). IEEE, 2000.*
- [72] L. Spacek et a. C. Burbridge, «Instantaneous robot self-localization and motion estimation with omnidirectional vision,» *Robotics and Autonomous Systems 55.9 (2007): 667-674, 2007.*
- [73] J. Gluckman, S. K. Nayar et K. J. & Thoresz, « Real-time omnidirectional and panoramic stereo,» *Proc. of DARPA Image Understanding Workshop (Vol. 1, pp. 299-303), 1998.*
- [74] T. Yi-ping et e. al., «Design of vertically aligned binocular omnistereo vision sensor,» *EURASIP Journal on Image and Video Processing 2010 , 2010.*
- [75] J. Gluckman, S. K. Nayar et K. J. & Thoresz, «Real-time omnidirectional and panoramic stereo,» *Proc. of DARPA Image Understanding Workshop (Vol. 1, pp. 299-303), 1998.*
- [76] L. Spacek et a. C. Burbridge, «Instantaneous robot self-localization and motion estimation with omnidirectional vision,» *Robotics and Autonomous Systems 55.9 (2007): 667-674, 2007.*
- [77] L. Su, C. Luo et a. F. Zhu, «Obtaining obstacle information by an omnidirectional stereo vision system,» *Information Acquisition, 2006 IEEE International Conference on. IEEE, 2006, 2006.*
- [78] D. Southwell et e. al., «Panoramic stereo,» *Pattern Recognition, 1996, Proceedings of the 13th International Conference on. Vol. 1. IEEE, 1996, 1996.*
- [79] E. L. Cabral, J. C. d. Souza et a. M. C. Hunold, «Omnidirectional stereo vision with a hyperbolic double lobed mirror,» *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. Vol. 1. IEEE, 2004.*

- [80] D. [Southwell, A. Basu, M. Fiala et J. & Reyda, «Panoramic stereo,» *Pattern Recognition, 1996., Proceedings of the 13th International Conference on (Vol. 1, pp. 378-382). IEEE, 1996.*
- [81] M. Fiala et A. & Basu, «Panoramic stereo reconstruction using non-SVP optics,» *Computer Vision and Image Understanding, 98(3), 363-397, 2005.*
- [82] S. Yi et a. N. Ahuja, «An omnidirectional stereo vision system using a single camera,» *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on. Vol. 4. IEEE, 2006.*
- [83] K. C. Ng et e. al., «An integrated surveillance system—human tracking and view synthesis using multiple omni-directional vision sensors,» *Image and Vision Computing 22.7 (2004): 551-561, 2004.*
- [84] S. Morita, K. Yamazawa et a. N. Yokoya, «Networked video surveillance using multiple omnidirectional cameras.»,» *Computational Intelligence in Robotics and Automation, 2003. Proceedings. 2003 IEEE International Symposium on. Vol. 3. IEEE, 2003.*
- [85] W. Shimizuhira et a. Y. Maeda, «Self-localization method used multiple omnidirectional vision system,» *SICE 2003 Annual Conference. Vol. 1. IEEE, 2003.*
- [86] Y. Taguchi et e. al., «Axial-cones: modeling spherical catadioptric cameras for wide-angle light field rendering,» *ACM Transactions on Graphics (TOG) 29.6 (2010): 172., 2010.*
- [87] R. Sagawa et e. al., «Compound catadioptric stereo sensor for omnidirectional object detection,» *Intelligent Robots and Systems, 2004.(IROS 2004) Proceedings. 2004 IEEE/RSJ International Conference on. Vol. 3. IEEE, 2004.*
- [88] G. Caron, E. M. Mouaddib et a. E. Marchand, «3D model based tracking for omnidirectional vision: A new spherical approach,» *Robotics and Autonomous Systems 60.8 (2012): 1056-1068, 2012.*
- [89] Z. El Kadmiri, O. El Kadmiri et L. & MASMOUDI, «DEPTH ESTIMATION FOR MOBILE ROBOT USING SINGLE OMNIDIRECTIONAL CAMERA SYSTEM,» *Journal of Theoretical and Applied Information Technology, 44(1), 2012.*
- [90] R. A. Hicks et R. & Bajcsy, «Catadioptric sensors that approximate wide-angle perspective projections,» *In Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on (Vol. 1, pp. 545-551). IEEE, 2000.*

- [91] J. Gaspar, C. Deccó, J. Okamoto Jr et J. & Santos-Victor, «Constant resolution omnidirectional cameras,» *In Omnidirectional Vision, 2002. Proceedings. Third Workshop on (pp. 27-34). IEEE, 2002.*
- [92] J. Salvi, X. Armangué et J. & Batlle, «A comparative review of camera calibrating methods with accuracy evaluation,» *Pattern recognition, 35(7), 1617-1635, 2002.*
- [93] W. K. Wong, J. T. Y. Liew, C. K. Loo et W. K. & Wong, «Omnidirectional surveillance system for digital home security.,» *In Signal Acquisition and Processing, 2009. ICSAP 2009. International Conference on (pp. 8-12). IEEE, 2009.*
- [94] X. Zhihui, C. Wang et Z. & Maojun, «Catadioptric Omni-directional Stereo Vision and Its Applications in Moving Objects Detection,» *Computer Vision, Xiong Zhihui (Ed.), ISBN: 978-953-7619-21-3, InTech, DOI: 10.5772/6159, 2008.*
- [95] J. Lei, X. Du, Y. F. Zhu et J. L. & Liu, «Unwrapping and stereo rectification for omnidirectional images,» *Journal of Zhejiang University SCIENCE A, 10(8), 1125-1139, 2009.*
- [96] N. S. Chong, Y. H. Kho et M. L. D. & Wong, «Custom aspect ratio correction for unwrapped omnidirectional view images,» *Computers & Electrical Engineering, 2013.*
- [97] G. Cielniak, M. Miladinovic, D. Hammarin, L. Goranson, A. Lilienthal et T. & Duckett, «Appearance-based tracking of persons with an omnidirectional vision sensor,» *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Confe, 2003.*
- [98] N. Vlassis, Y. Motomura, I. A. Hara et & M. T. H., «Edge-based features from omnidirectional images for robot localization,» *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on (Vol. 2, pp. 1579-1584), 2001.*
- [99] O. El Kadmiri et L. & Masmoudi, «An omnidirectional image unwrapping approach,» *International Conference on Multimedia Computing and Systems (ICMCS), 2011 International Conference on (pp. 1-4). IEEE, (2011, April).*
- [100] S. O. et Comby F., «Variable structuring element based fuzzy morphological operations for single viewpoint omnidirectional images,» *Pattern Recognition, volume 40, number 12, 3578-3596, 2007 .*

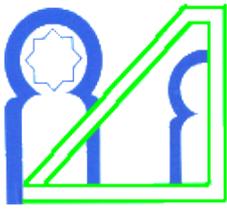
- [101] J. F., L. K., C. F. et Strauss O., «Non-Additive Approach for Gradient-Based Edge Detection.» *IEEE International Conference on Image Processing, volume 3, III49-III52*, 2007.
- [102] J. Cruz-Mota, I. Bogdanova, B. Paquier, M. Bierlaire et J. P. & Thiran, «Scale invariant feature transform on the sphere: Theory and applications.» *International journal of computer vision*, 98(2), 217-241, 2012.
- [103] C. Geyer et K. & Daniilidis, «A unifying theory for central panoramic systems and practical implications.» *In Computer Vision—ECCV 2000 (pp. 445-461)*., 2000.
- [104] X. Ying et H. & Zha, «Using sphere images for calibrating fisheye cameras under the unified imaging model of the central catadioptric and fisheye cameras.» *In Pattern Recognition, 2006. ICPR 2006. 18th International Conference on Vol. 1*, 2006.
- [105] R. B. S. Lin., «True single view point cone mirror omni-directional catadioptric system.» *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on (Vol. 2, pp. 102-107)*. IEEE, 2001.
- [106] S. Derrien et K.G. Konolige, «Approximating a Single Viewpoint in Panoramic Imaging Devices.» *IEEE International Conference on Robotic and Automation*, 2000.
- [107] Z. D. Xiang et & G. X. X., «Noncentral catadioptric camera calibration using a generalized unified model.» *Optics letters*, 38(9), 1367-1369, 2013.
- [108] R. A. Hicks et R. & Bajcsy, «Catadioptric sensors that approximate wide-angle perspective projections.» *In Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on (Vol. 1, pp. 545-551)* , 2000.
- [109] O. EL KADMIRI, Z. EL KADMIRI et L. & MASMOUDI, «A SPHERICAL ELECTROSTATIC MODEL EDGE DETECTOR FOR OMNIDIRECTIONAL IMAGES.» *Journal of Theoretical and Applied Information Technology*, 51(3)., 2013.
- [110] J. Fram et E.S. Deutsch, «On the quantitative evaluation of edge detection schemes and their comparisons with human performance.» *IEEE Trans.Comput. C-24 (6) (1975) 616–627* , 1975.
- [111] E. S. Deutsch et J. R. & Fram, «A quantitative study of the orientation bias of some edge detector schemes.» *IEEE Transactions on Computers*, 27(3), 205-213, 1978.

- [112] K. Rohr, «Modelling and identification of characteristic intensity variations,» *Image and Vision Computing*, 10(2), 66-76, 1992.
- [113] A. Guiducci, « Corner characterization by differential geometry techniques,» *Pattern Recognition Letters*, 8(5), 311-318, 1988.
- [114] Z. Zheng, H. Wang et E. & Khwang Teoh, «Analysis of gray level corner detection,» *Pattern Recognition Letters*, 20(2), 149-162, 1999.
- [115] A. Singh et M. Shneier, «Gray level corner detection a generalization and a robust real time implementation,» *Computer Vision, Graphics, and Image Processing* 51, 54±69 , 1990.
- [116] P. R. Beaudet, «Rotationally invariant image operators,» *Proceedings of the International Joint Conference on Pattern Recognition (Vol. 579, pp. 579-583)*, 1978.
- [117] H. Moravec, «Towards automatic visual obstacle avoidance,» *Fifth International Joint Conference On Arti@cial Intelligent*, p. 584, 1977.
- [118] C. Harris et M. & Stephens, «A combined corner and edge detector,» *In Alvey vision conference (Vol. 15, p. 50)*, 1988.
- [119] O. Wijk et a. H. I. Christensen., «Localization and navigation of a mobile robot using natural point landmarks extracted from sonar data,» *Robotics and Autonomous Systems*, 31(1):31–42, 2000.
- [120] D. Hahnel, W. Burgard, D. Fox et a. S. Thrun., «An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements,» *In Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*, 2003.
- [121] M. Srinivasan, K.Weber et a. S. Venkatesh., «From living eyes to seeing machines,» *Oxford University Press*, 1997.
- [122] N. Aihara, H. Iwasa, N.Yokoya et a. H. Takemura., «Memorybased self-localization using omnidirectional images,» *In ICPR'98, pages 1799–1803*, 1998.
- [123] R. Cassinis, D. Grana et a. A. Rizzi., «Self-localization using an omnidirectional image sensor,» *In 4th Int. Symp. on Intelligent Robotic Systems (SIRS'96), pages 215–222, Lisbon, Portugal*, 1996.
- [124] L. Delahoche, C. P´egard, B. Marhic et a. P. Vasseur, «A navigation system based on an omnidirectional vision sensor,» *In Proc. of the Int. Conf. on Intelligent Robotics and Systems 1997 (IROS'97), pages 718–724, Grenoble, France*, 1997.

- [125] F. Rameau, D. Sidibé, C. Demonceaux et D. & Fofi, «Une approche performante de suivi visuel pour les caméras catadioptriques,» *In Actes de la conférence RFIA*, 2012.
- [126] K. E. A. v. d. Sande, T. Gevers et a. C. G. M. Snoek., «Evaluating color descriptors for object and scene recognition,» *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596, 2010.
- [127] K. Kabidi, M. N. Bargach et R. Tadili., «Generation of solar irradiance on inclined surface using along terms measurements,» *Physical and Chemical News*, vol. 68. April 2013, 1892, pp.17–24, 2013.
- [128] R. Neville, «Solar energy collector orientation and tracking mode,» *Solar energy 20 (1)*, 7–11J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73, 1978.
- [129] S. Abdallah, «The effect of using sun tracking systems on the voltage-current characteristics and power generation of flat plate photovoltaics,» *Energ. Convers. Manage.* 2004, 45, 1671-1679, 2004.
- [130] H. Mousazadeh, A. Keyhani, A. Javadi, H. Mobli, K. Abrinia et A. Sharifi, «A review of principle and sun-tracking methods for maximizing solar systems output,» *Renewable and Sustainable Energy Reviews*, Volume 13, Issue 8, 2009.
- [131] C.-Y. Lee, P.-C. Chou, C.-M. Chiang et C.-F. Lin, «Sun Tracking Systems: A Review,» *Sensors 9*, no. 5: 3875-3890, 2009.
- [132] B. P, T. M et R. P., «Evaluation of power output for fixed and step tracking photovoltaic arrays,» *Solar Energy 1986;37(2):147–63*, 1986.
- [133] B. M, G. A et B. U., «Two-stage concentrator permitting concentration factors up to 300X with one-axis tracking,» *Solar Energy 1996;56(3):285–300*, 1996.
- [134] S. S et R. N., «A shadow method for automatic tracking,» *Solar Energy 1986;37(3):245–7*, 1986.
- [135] Z. R et L. D., «Design and construction of a sun tracker,» *Solar Energy 1984;33(314):369–72*, 1984.
- [136] P. Milea, O. Oltu, M. Dragulinescu et M. Dascalu, «Optimizing solar panel energetic efficiency using an automatic tracking microdetector,» *Proceedings of the WSEAS International Conference on Renewable Energy Sources*, Arcachon, France, 2007.

- [137] V. Poulek et M. Libra, «A Very Simple Solar Tracker for Space and Terrestrial Applications,» *Solar Energy Materials and Solar Cells*, 60, 2, (2000), 99-103, 2000.
- [138] M. Berenguel, F. R. Rubio, A. Valverde, P. J. Lara, M. R. Arahál, E. F. Camacho et M. Lopez, «An artificial vision-based control system for automatic heliostat positioning offset correction in a central receiver solar power plant,» *Solar Energy, Volume 7*, 2004.
- [139] Z. E. KADMIRI, O. E. KADMIRI, L. MASMOUDI et M. N. BARGACH, «A Novel Solar Tracker Based On Omnidirectional Computer Vision,» *accepted for publication*.
- [140] http://upload.wikimedia.org/wikipedia/commons/5/56/Thru_The_Eyes_Of_Ruby_%28the_fly%29_%288219315716%29.jpg.
- [141] <http://scienceblogs.com/deepseanews/wp-content/blogs.dir/414/files/2012/04/i-2bb4136f0f1303f215d35c1cf185d044-giganto.jpg>.
- [142] <http://scienceblogs.com/deepseanews/wp-content/blogs.dir/414/files/2012/04/i-2bb4136f0f1303f215d35c1cf185d044-giganto.jpg>.
- [143] <http://www.viewplus.co.jp/product/system/astro.html>.
- [144] <http://filmmakeriq.com/wp-content/uploads/2012/10/Female-Thiodina-puterpera-Jumping-Spider.jpg>.
- [145] <http://www.wildlife-art.net/images/kimberley131.jpg>.
- [146] <http://www.panoramio.com/photo/87859692>.
- [147] <http://www.peleng8.com/samyang-8mm.html>.
- [148] <http://www.sandydan.com/photo/wide/fish/ftest5.jpg>.
- [149] <http://sandydan.com/photo/wide/sphere/sphere2/index.htm>.
- [150] http://upload.wikimedia.org/wikipedia/commons/8/80/Archimedes-Mirror_by_Giulio_Parigi.jpg.
- [151] <http://www.independent.co.uk/life-style/gadgets-and-tech/news/google-ordered-to-delete-data-collected-by-street-view-cars-8668974.html>.
- [152] <http://www.lunkwill.org/cv/>.
- [153] <http://www.xbox.com/en-US/Kinect/>.

- [154] <http://www.engadget.com/2010/11/08/visualized-kinect-night-vision-lots-and-lots-and-lots-of-do/>.
- [155] http://farm3.static.flickr.com/2671/3964074600_0f2999a030_m.jpg.
- [156] A. M. Loh et R. & Hartley, «Shape from Non-homogeneous, Non-stationary, Anisotropic, Perspective Texture,» *BMVC*, 2005.
- [157] E. Prados et O. & Faugeras, «A generic and provably convergent shape-from-shading method for orthographic and pinhole cameras,» *International Journal of Computer Vision*, 65(1-2), 97-125, 2005.
- [158] D. Roxo, N. Gonçalves et J. P. & Barreto, «Perspective Shape from Shading for Wide-FOV Near-Lighting Endoscopes,» *Pattern Recognition and Image Analysis (pp. 21-30)*. Springer Berlin Heidelberg, 2013.
- [159] <http://www.logitech.com/fr-be/support/hd-webcam-c310>.
- [160] <http://www.robotshop.com/media/files/images/design/johnny-robot-track-configurations.jpg>.
- [161] <http://www.riegl.com/nc/products/mobile-canning/produktdetail/product/scanner/22/>.
- [162] <http://www.retronaut.com/2013/03/self-portrait-in-spherical-mirror-m-c-escher/>.



UNIVERSITE MOHAMMED V-AGDAL

FACULTE DES SCIENCES

Rabat

DOCTORAT

Résumé de la Thèse

Spécialité : Électronique Et Traitement Des Images

Laboratoire: **Laboratoire d'Electronique et de Traitement du Signal/Géomatique.**

Responsable de la structure de recherche : **Lhoussaine MASMOUDI, PES.**

Période d'accréditation : **2012/2016.**

Titre de la thèse : **Contribution à la stéréovision omnidirectionnelle et au traitement des images catadioptriques : Application aux systèmes autonomes.**

Prénom, Nom : **Omar El KADMIRI.**

Résumé : La vision par ordinateur est une discipline qui vise doter les ordinateurs d'un sens de perception et d'analyse d'image semblable à celui de l'homme. La perception visuelle artificielle peut être grandement améliorée quand un grand champ de vision est disponible.

Cette thèse traite de l'utilisation des caméras omnidirectionnelles comme un moyen d'élargir le champ de vision des systèmes de vision artificielle. La perception visuelle de la profondeur (3D) par le biais de configurations omnistéréo, et les algorithmes de traitement adaptés aux images catadioptriques, sont les principaux sujets étudiés.

Tout d'abord une étude des systèmes de vision omnidirectionnelle 3D est menée. Elle met en évidence les principales approches pour obtenir l'information sur la profondeur et fournit des indications précieuses sur le choix de la configuration en fonction des besoins de l'application. Ensuite, la conception d'un capteur omnistéréo est adressée ; nous présentons une nouvelle configuration du capteur proposé basé une caméra catadioptrique unique, et dédié à la robotique mobile. Des expérimentations sur la précision d'estimation de la profondeur ont été menées pour valider la nouvelle configuration.

Les images catadioptriques présentent diverses propriétés géométriques particulières, telles que la résolution non-uniforme et de fortes distorsions radiales. L'application des algorithmes de traitement classiques à ce type d'images se trouve limité en termes de performances. Dans ce sens, de nouveaux algorithmes adaptés à la géométrie sphérique de ces images ont été développés.

Les techniques de vision omnidirectionnelle artificielle recueillies ont été finalement exploitées dans deux applications réelles. La première concerne l'intégration des caméras catadioptriques à un robot mobile. La seconde porte sur la conception d'un suiveur solaire, à base d'une caméra catadioptrique.

Les résultats obtenus confirment que l'adoption de tels capteurs pour les systèmes autonomes offre plus de performances et de flexibilité en regards aux capteurs classiques.

Mots-clefs : stéréovision omnidirectionnelle, caméra catadioptrique, traitement d'images, systèmes autonomes

Faculté des Sciences, 4 Avenue Ibn Battouta B.P. 1014 RP, Rabat – Maroc

Tel +212 (0) 37 77 18 34/35/38, Fax : +212 (0) 37 77 42 61, <http://www.fsr.ac.ma>

