



**HAL**  
open science

# Conception d'un capteur de stéréovision omnidirectionnelle: architecture, étalonnage et applications à la reconstruction de scènes 3D

Nicolas Ragot

► **To cite this version:**

Nicolas Ragot. Conception d'un capteur de stéréovision omnidirectionnelle: architecture, étalonnage et applications à la reconstruction de scènes 3D. Sciences de l'ingénieur [physics]. Université de Rouen, 2009. Français. NNT: . tel-00417963

**HAL Id: tel-00417963**

**<https://theses.hal.science/tel-00417963>**

Submitted on 17 Sep 2009

**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.



# Conception d'un capteur de stéréovision omnidirectionnelle : architecture, étalonnage et applications à la reconstruction de scènes 3D

## THÈSE

présentée et soutenue publiquement le 07 Avril 2009

pour l'obtention du

**Doctorat de l'Université de Rouen**

**Discipline : Génie Informatique, Automatique et Traitement du signal**

**Spécialité : Vision par ordinateur**

par

Nicolas RAGOT

### Composition du jury

<i>Président :</i>	M. Claude PEGARD	Professeur à l'Université de Picardie Jules Verne
<i>Rapporteurs :</i>	M. Ryad BENOSMAN	Maître de Conférences HDR à l'Université Paris VI
	M. Olivier STRAUSS	Maître de Conférences HDR à l'Université Montpellier II
<i>Examineurs :</i>	M. Pierre BLAZEVIC	Professeur à l'Université de Saint Quentin en Yvelines
	M. Bélahcène MAZARI	Professeur à l'ESIGELEC-IRSEEM, Directeur de thèse
	M. Xavier SAVATIER	Enseignant-Chercheur à l'ESIGELEC-IRSEEM, Encadrant
<i>Invité :</i>	M. Jean-Yves ERTAUD	Enseignant-Chercheur à l'ESIGELEC-IRSEEM, Encadrant

## Remerciements

Bien qu'apparaissant au début du manuscrit, cette page est chronologiquement la dernière écrite. Elle revêt de fait un caractère particulier puisqu'elle clôture une longue aventure... Elle permet également de remercier les différentes personnes ayant collaboré de près ou de loin à mes travaux de thèse.

Je tiens tout d'abord à remercier M. Claude GUILLERMET, Directeur Général de l'ESIGELEC de m'avoir permis d'effectuer cette thèse. J'exprime ma profonde reconnaissance à M. Bélahcène MAZARI, Directeur de l'IRSEEM pour m'avoir accueilli au sein du laboratoire et pour avoir accepté d'être mon directeur de thèse. Qu'il reçoive mes sincères remerciements pour les conseils et encouragements qu'il m'a prodigués et pour la confiance qu'il m'a témoignée.

Je remercie chaleureusement M. Olivier STRAUSS de l'Université de Montpellier II pour avoir accepté d'expertiser mon manuscrit. Je salue le regard avisé qu'il a porté sur mon travail et la justesse des remarques qu'il m'a formulées. Elles m'ont été particulièrement utiles pour parfaire mon mémoire. Mes sincères remerciements vont également à M. Ryad BENOSMAN de l'Université Paris VI pour avoir accepté d'être le rapporteur de ma thèse. Je lui témoigne ma reconnaissance pour sa lecture approfondie du mémoire et ses conseils qui ne peuvent que m'encourager dans mes recherches futures.

J'adresse également mes remerciements à M. Claude PEGARD de l'Université de Picardie Jules Verne pour avoir présidé mon jury de thèse et M. Pierre BLAZEVIC de l'Université de Saint Quentin en Yvelines pour avoir accepté d'en être l'examineur.

Je souhaite ensuite adresser ma profonde gratitude à mes deux encadrants, M. Jean-Yves ERTAUD, responsable du département Génie Electrique et Informatique Industrielle à l'ESIGELEC et M. Xavier SAVATIER, responsable du pôle Instrumentation et Systèmes de l'IRSEEM. Ces quelques lignes sont un faible hommage en comparaison du soutien et du dévouement dont ils ont fait preuve. Sans leur aide et leur encouragement, ce travail n'aurait sans doute jamais abouti...

Un grand merci aux membres de l'équipe IIS : à Rémi mon collègue de bureau, pour nos échanges fructueux, à Romain pour ses codes C et ses excellentes idées pour remplir le cahier de laboratoire (... la photoconsistance spatiotemporelle), à Jean-François pour ses codes PovRay et à Pierre pour les « vendredis-restos ». J'ai également une pensée amicale pour Fengchun et Mohammed qui ont rejoint l'équipe depuis peu, pour les stagiaires

(notamment Miguel) qui ont collaboré à mes travaux et pour tous les autres thésards qui débutent ou qui finissent leur doctorat. Je leur souhaite bon courage.

J'adresse également mes remerciements à l'ensemble du personnel de l'ESIGELEC et de l'IRSEEM pour leur aide et leur gentillesse. Merci en particulier à Anne, Christine, Fabrice, Vincent et une pensée particulière pour Claude mon binôme de TP.

Un grand merci à Nicole, ma belle-maman, et ma Maman pour le travail si fastidieux mais tellement important de relecture des fautes d'orthographe. Merci également à Ben, mon expert matheux, pour ses conseils et ses corrections.

Je terminerai ces remerciements en mentionnant les personnes qui tiennent une place particulière dans mon coeur : mes Parents sans qui tout ceci ne serait pas. Mille mercis ne suffiraient pas à leur exprimer mon affection et ma reconnaissance. Ma femme Célia qui m'a accompagné dans cette aventure. Son soutien, ses encouragements, sa patience et sa gentillesse ont été le plus beau témoignage d'amour qui m'ait été fait. Encore mille fois merci...

*A ma femme Célia,  
A mon fils Quentin,  
A mes Parents,  
A la mémoire de ma Grand-Mère et de mon Grand-Père*

# Table des matières

<b>Table des figures</b>	<b>viii</b>
<b>Liste des tableaux</b>	<b>xii</b>
<b>Introduction générale</b>	<b>1</b>
<b>Conventions et notations</b>	<b>5</b>
<b>1 Capteurs de perception de scènes 3D</b>	<b>7</b>
1.1 Introduction . . . . .	8
1.2 Perception 3D par capteurs actifs : télémétrie et triangulation . . . . .	10
1.2.1 Télémétrie . . . . .	10
1.2.2 Triangulation active en lumière structurée . . . . .	11
1.2.3 Synthèse . . . . .	13
1.3 Perception 3D par capteurs passifs : la vision . . . . .	16
1.3.1 Capteurs de vision directionnelle . . . . .	17
1.3.2 Capteurs de vision omnidirectionnelle . . . . .	22
1.3.3 Synthèse . . . . .	29

---

1.4	Capteurs de stéréovision omnidirectionnelle . . . . .	32
1.4.1	Quelques exemples de capteurs . . . . .	34
1.4.2	Géométrie épipolaire . . . . .	34
1.5	Conclusions . . . . .	42
<b>2</b>	<b>Solution matérielle et étalonnage paramétrique</b>	<b>44</b>
2.1	Introduction . . . . .	45
2.2	Proposition d'une solution matérielle . . . . .	47
2.2.1	Caractéristiques matérielles et techniques . . . . .	47
2.2.2	Réglage du point de vue unique . . . . .	49
2.3	Étalonnage de capteurs catadioptriques . . . . .	52
2.3.1	Modèles de capteurs centraux et non-centraux . . . . .	53
2.3.2	Méthodes d'étalonnage . . . . .	61
2.3.3	Mesures de la précision d'un étalonnage . . . . .	64
2.4	Étalonnage paramétrique par un modèle ad hoc . . . . .	66
2.4.1	Modèles de projection et rétro-projection . . . . .	67
2.4.2	Estimation des paramètres . . . . .	68
2.5	La mire d'étalonnage . . . . .	71
2.5.1	Caractéristiques techniques . . . . .	71
2.5.2	Extraction des points d'intérêt . . . . .	72
2.5.3	Dispositif d'aide à l'étalonnage . . . . .	74
2.6	Résultats . . . . .	78

---

2.6.1	Environnement de simulation et protocole d'étalonnage . . . . .	78
2.6.2	Estimation des paramètres . . . . .	79
2.7	Discussion . . . . .	82
2.8	Conclusions . . . . .	84
<b>3</b>	<b>Etalonnage non-paramétrique et conception d'une mire adaptée</b>	<b>86</b>
3.1	Introduction . . . . .	87
3.2	Etalonnage non-paramétrique par un modèle discret . . . . .	88
3.2.1	Principe . . . . .	88
3.2.2	Projection . . . . .	90
3.2.3	Rétro-projection . . . . .	93
3.3	Conception d'une mire adaptée . . . . .	97
3.3.1	Nombre de points d'intérêt par cercle directeur . . . . .	97
3.3.2	Distribution verticale des cercles directeurs sur le cylindre . . . . .	99
3.4	Evaluation en simulation et sur le système réel . . . . .	102
3.4.1	Protocoles expérimentaux et méthodes de validation . . . . .	102
3.4.2	Résultats . . . . .	105
3.5	Discussion . . . . .	115
3.6	Conclusions . . . . .	121

---

<b>4 Applications à la reconstruction de scènes 3D</b>	<b>122</b>
4.1 Introduction . . . . .	123
4.2 Reconstruction 3D volumétrique pour le capteur de stéréovision omnidirectionnelle . . . . .	125
4.2.1 Principe général . . . . .	125
4.2.2 Applications au capteur de stéréovision omnidirectionnelle . . . . .	127
4.3 Résultats . . . . .	132
4.3.1 Reconstructions 3D statiques pour une scène réelle . . . . .	132
4.3.2 Reconstruction 3D dynamique par déplacements successifs du capteur pour une scène de synthèse . . . . .	137
4.4 Discussion . . . . .	138
4.5 Conclusions . . . . .	141
<b>Conclusions et perspectives</b>	<b>142</b>
<b>Bibliographie</b>	<b>145</b>
<b>A Modèle sténopé</b>	<b>160</b>
<b>B Modèle orthographique</b>	<b>163</b>
<b>C Algorithme de Levenberg-Marquardt</b>	<b>165</b>
<b>D Décomposition en valeurs singulières</b>	<b>167</b>
<b>E Normes de Minkowski</b>	<b>168</b>

# Table des figures

1	Chaîne de traitement pour un processus de numérisation 3D . . . . .	2
1.1	Technologie de capteurs par télémétrie . . . . .	12
1.2	Technologie de capteurs par lumière structurée . . . . .	14
1.3	Champs d’observation des capteurs de vision . . . . .	16
1.4	Le modèle sténopé . . . . .	17
1.5	Le modèle de projection orthographique . . . . .	19
1.6	Exemples de capteurs cylindrique et sphérique . . . . .	23
1.7	Exemples de capteurs à réseaux de caméras . . . . .	24
1.8	Exemple de capteurs à objectif grand angle . . . . .	25
1.9	Exemple de capteurs catadioptriques . . . . .	25
1.10	La contrainte de point de vue unique . . . . .	26
1.11	Des capteurs catadioptriques à projection centrale . . . . .	28
1.12	D’autres capteurs catadioptriques à projection centrale . . . . .	29
1.13	Courbes caustiques pour des capteurs à projection non-centrale . . . . .	30
1.14	Principe de la stéréovision . . . . .	32
1.15	Géométrie épipolaire pour des capteurs perspectifs . . . . .	35

---

1.16	Géométrie épipolaire pour des capteurs catadioptriques . . . . .	37
1.17	Géométrie épipolaire de capteurs catadioptriques en configuration coaxiale	38
1.18	Projections cylindrique et sphérique d'une image omnidirectionnelle . . . .	41
2.1	Étalonnage d'un capteur de vision : principe général . . . . .	45
2.2	Détails techniques du capteur hypercatadioptrique utilisé . . . . .	48
2.3	Architecture matérielle : photo et schéma détaillé . . . . .	49
2.4	Réglage du point de vue unique pour un capteur hypercatadioptrique par la méthode de Zivkovic & Booij . . . . .	50
2.5	Projection d'un point 3D par des modèles ad hoc de transformation entre l'espace 3D et le plan image . . . . .	55
2.6	Projection d'un point 3D par le modèle de la sphère d'équivalence . . . . .	58
2.7	Modèle générique de transformation entre l'espace 3D et le plan image . .	61
2.8	Étalonnage par l'empreinte du miroir . . . . .	63
2.9	Projection d'un point 3D pour un capteur hypercatadioptrique . . . . .	67
2.10	La mire d'étalonnage et différentes vues d'un point d'intérêt 3D . . . . .	72
2.11	Exemple d'une image d'étalonnage et détail d'un ovoïde . . . . .	73
2.12	Schéma du dispositif d'aide à l'étalonnage . . . . .	75
2.13	Interface graphique pour une aide à l'étalonnage . . . . .	76
2.14	Carte électronique pour le pilotage des marqueurs lumineux 3D . . . . .	77
2.15	Photo de la mire d'étalonnage . . . . .	77
2.16	Environnement d'étalonnage sous POV-Ray . . . . .	79
2.17	Points 2D extraits et estimés pour le vecteur initial des paramètres . . . .	80
2.18	Points 2D extraits et estimés pour le vecteur optimal des paramètres . . .	81

---

3.1	Relations entre l'espace 3D et le plan image pour un étalonnage non-paramétrique . . . . .	88
3.2	Représentation de l'espace des variables $(\theta, \Phi, \rho)$ . . . . .	91
3.3	Projection des points $\mathbf{V}$ et du point $\widehat{\mathbf{V}}$ dans le plan $(\theta\mathbf{O}\Phi)$ . . . . .	92
3.4	Représentation du plan passant par les points $\{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\}$ . . . . .	93
3.5	Représentation de l'espace des variables $(\theta, \rho, \Phi)$ . . . . .	94
3.6	Projection des points $\mathbf{V}$ et du point $\widehat{\mathbf{V}}$ dans le plan $(\theta\mathbf{O}\rho)$ . . . . .	95
3.7	Représentation du plan passant par les points $\{\mathbf{V}'_1, \mathbf{V}'_2, \mathbf{V}'_3\}$ . . . . .	96
3.8	Répartition uniforme des cercles directeurs en 3D . . . . .	100
3.9	Répartition adaptée des points d'intérêt 3D . . . . .	101
3.10	Schémas illustrant les erreurs de rétro-projection et de reconstruction . . .	104
3.11	Maillage pour une distribution régulière des points d'intérêt 3D obtenu pour un environnement simulé . . . . .	105
3.12	Projection des trois cercles 3D dans l'image omnidirectionnelle . . . . .	106
3.13	Erreur de rétro-projection pour une répartition uniforme des points d'intérêt 3D . . . . .	107
3.14	Erreur de reconstruction en millimètre estimée par intersection des lignes de vue des deux capteurs pour une distribution régulière des points d'intérêt	108
3.15	Erreur de reconstruction en pourcent estimée par intersection des lignes de vue des deux capteurs pour une distribution régulière des points d'intérêt .	108
3.16	Maillage pour une distribution régulière des points d'intérêt 3D obtenu pour le système réel . . . . .	109
3.17	Erreur de rétro-projection pour une répartition uniforme des points d'intérêt 3D . . . . .	110
3.18	Vue du maillage obtenu en simulation et correspondant à une répartition adaptée des points d'intérêt 3D sur la mire d'étalonnage . . . . .	111

---

3.19	Evolution de l'erreur de projection en fonction du nombre de cercles 3D . . .	111
3.20	Erreur de rétro-projection pour une distribution adaptée des points d'intérêt 3D . . . . .	113
3.21	Erreur de reconstruction en millimètre estimée par intersection des lignes de vue des deux capteurs pour une distribution adaptée des points d'intérêt	114
3.22	Erreur de reconstruction en pourcent estimée par intersection des lignes de vue des deux capteurs pour une distribution adaptée des points d'intérêt . . .	114
3.23	Projections théorique et estimée des ovoïdes . . . . .	117
3.24	Estimation de la pose entre la mire et le capteur pour un étalonnage non-paramétrique . . . . .	119
4.1	Photo-consistance d'un voxel : principe général . . . . .	126
4.2	Reconstruction 3D volumétrique pour le capteur de stéréovision . . . . .	128
4.3	Reconstruction 3D volumétrique pour le capteur de stéréovision . . . . .	130
4.4	Reconstruction 3D statique par projection du centre de gravité du voxel . . .	133
4.5	Reconstruction 3D statique avec suppression des voxels de luminance élevée	134
4.6	Portion du volume de reconstruction découpé le plan $(\mathbf{Y}_W \mathbf{O}_W \mathbf{Z}_W)$ . . . . .	134
4.7	Vue en coupe dans le plan $(\mathbf{Y}_W \mathbf{O}_W \mathbf{Z}_W)$ du volume de reconstruction . . . . .	135
4.8	Reconstruction 3D statique pour le capteur stéréoscopique . . . . .	136
4.9	Reconstruction 3D dynamique pour le capteur stéréoscopique . . . . .	137
A.1	Le modèle sténopé détaillé . . . . .	160
B.1	Le modèle de projection orthographique détaillé . . . . .	163

# Liste des tableaux

2.1	Paramètres caractéristiques pour le modèle de la sphère d'équivalence . . .	57
2.2	Valeurs théoriques et estimées des paramètres pour l'étalonnage paramétrique	82
3.1	Evolution de l'erreur moyenne de projection en fonction du nombre de couples point 3D-point 2D pour une génératrice du cylindre d'étalonnage .	112
3.2	Diamètres des projections des cercles 3D mesurés entre des barycentres d'ovoïdes diamétralement opposés . . . . .	117

# Introduction générale

La **numérisation tridimensionnelle** (ou numérisation 3D) d'objets ou d'environnements est une thématique scientifique et technologique qui connaît depuis une vingtaine d'années un essor important. L'objectif recherché est d'obtenir, à l'aide d'un ensemble de données collectées, une représentation 3D d'un objet ou plus, d'une scène, similaire à celle que l'on obtiendrait par perception visuelle.

La numérisation 3D regroupe deux principaux aspects : l'acquisition et la modélisation. Une définition simplifiée proposée par Laurent [81] mentionne que l'acquisition s'opère par une interface matérielle à l'aide de capteurs et la modélisation par une interface logicielle à l'aide d'algorithmes. Par abus de langage, il ne sera fait aucune distinction dans ce document entre les parties matérielles et logicielles et les termes *numérisation 3D*, *modélisation 3D*, *reconstruction 3D*, etc. renverront à une même notion : celle d'obtenir un rendu en trois dimensions d'un objet ou d'un environnement. Une chaîne de traitement, dont une illustration simplifiée est proposée en figure 1, fait apparaître les phases d'acquisition des données et leur modélisation.

Les domaines d'applications utilisant des techniques de numérisation 3D sont nombreux et vont de la reproduction d'objets, aux activités éducatives ou ludiques, en passant par l'analyse scientifique ou historique. Les objectifs sont donc variés et dépendent étroitement des applications visées [81] :

→ **pour des applications industrielles**, la numérisation 3D concerne principalement le « *reverse engineering* » (rétro-conception ou rétro-ingénierie). Le principe repose sur l'obtention d'un nuage de points d'un objet scanné numériquement ou palpé mécaniquement. Le nuage de points, préalablement maillé afin de reconstruire les surfaces, est ensuite exporté dans des logiciels de Conception Assistée par Ordinateur (CAO) pour être analysé, modifié et adapté à son environnement. Pour l'automobile ou l'aéronautique par exemple, une partie des processus de validation passe par la réalisation de maquettes physiques, suivie d'une phase de rétroconception afin d'améliorer, de

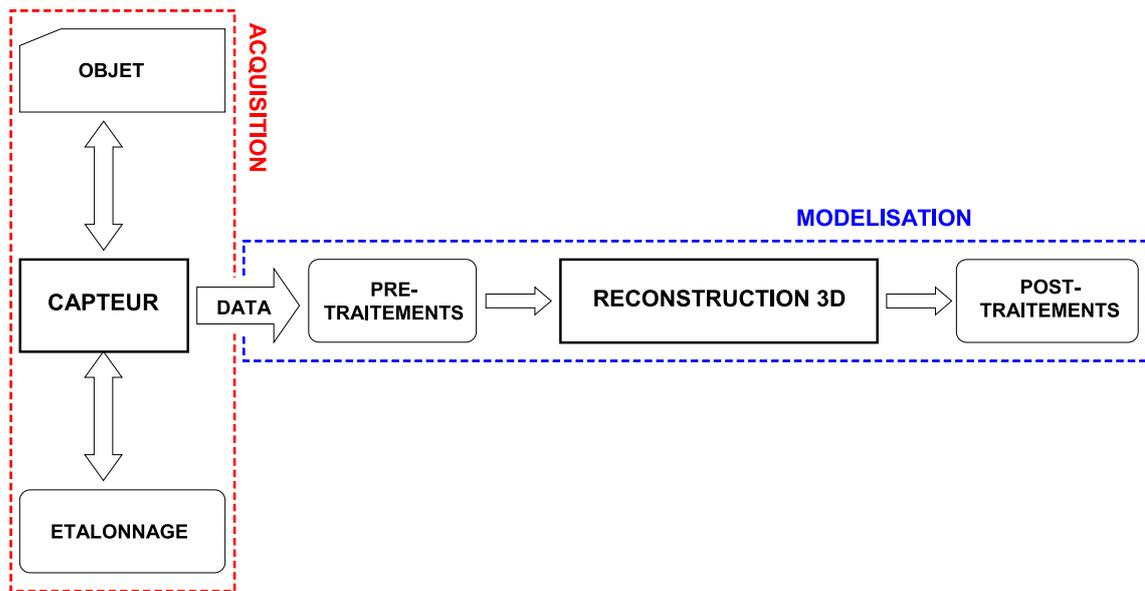


FIG. 1 – Synoptique simplifié d'une chaîne de traitement pour un processus de numérisation 3D.

modifier le procédé initial et d'intégrer virtuellement des éléments supplémentaires. La numérisation d'objets pour le contrôle dimensionnel sur des lignes de production fait aussi partie des applications bénéficiant des apports de la numérisation 3D. On peut également mentionner des applications d'aide à la saisie d'objets pour des bras manipulateurs et des applications en métrologie ;

- **pour des applications de type cartographie 3D**, il s'agit par exemple d'obtenir des représentations 3D de villes afin de faciliter la recherche d'un lieu ou d'en faire une visite virtuelle. Il peut s'agir également de visites virtuelles de bâtiments ou de lieux culturels, de projets d'aménagement du territoire, etc. La cartographie 3D concerne également la navigation autonome de robots mobiles. L'objectif est alors de pouvoir rendre un robot totalement autonome de sorte qu'il puisse éviter des obstacles et planifier ses trajectoires. Cette problématique suscite un grand intérêt car les enjeux économiques sont importants : pour la robotique autonome militaire, urbaine ou aérienne, ou pour la surveillance environnementale et l'intervention après des catastrophes ou encore pour la robotique d'aide aux handicapés (ex. : chaise roulante, assistance aux mal-voyants, etc.) ;
- **pour le domaine médical**, la numérisation 3D est dédiée à l'obtention de modèles géométriques 3D d'organes ou plus généralement du corps humain. Dans le cas de la modélisation 3D d'organes, il s'agit entre autres, de suivre l'évolution de pathologies ou d'obtenir des échographies 3D de nourrissons. Pour la modélisation 3D du corps,

---

certaines problématiques concernent la biomécanique. Il s'agit de pouvoir visualiser et mesurer la répartition des efforts appliqués en des points particuliers du corps humain [135]. D'autres encore concernent la reconstruction tridimensionnelle d'un corps humain à partir de photographies de tranches d'un cadavre<sup>1</sup>.

Il existe beaucoup d'autres applications pour lesquelles la numérisation 3D est d'un intérêt significatif. Il serait difficile de les référencer de manière exhaustive. Citons néanmoins des applications liées à la conservation du patrimoine qui utilisent la modélisation 3D pour la numérisation d'objets d'art et de documents [35]. Mentionnons également la modélisation 3D de végétaux pour des problématiques en recherche agronomique [132]. Signalons encore des applications de reconstruction 3D en volcanologie pour l'étude des déformations du sol et la surveillance des éruptions [27].

Les travaux de thèse proposés dans ce document s'inscrivent dans un programme de recherche européen Interreg IIIA **OmniViss** associant l'Institut de Recherche en Systèmes Électroniques EMbarqués et le département d'Électronique de l'université du Kent en Grande-Bretagne pour la **conception et l'évaluation d'un outil de reconstruction et d'analyse de scènes 3D inconnues a priori**. Ce programme avait pour objectif le déploiement d'une solution technologique (matérielle et algorithmique) répondant au problème de reconstruction 3D d'environnements sans connaissance a priori. Deux types d'applications devaient être considérés : la première concernait la reconstruction 3D d'environnements pour un capteur en position fixe. Il s'agissait, par exemple, de l'analyse du trafic routier aux abords d'un rond-point, le capteur étant localisé en son centre. La seconde portait sur la reconstruction 3D d'environnements pour un capteur en mouvement. L'application principalement visée était dans ce cas la navigation autonome de robots mobiles.

Notre contribution concerne donc le développement d'un capteur permettant de répondre favorablement aux applications visées par le programme *OmniViss*. Les premiers travaux ont porté sur la définition d'une solution technologique adaptée aux applications envisagées. Un deuxième axe d'étude s'est intéressé à l'étalonnage du capteur. Les dernières réflexions ont porté sur les algorithmes pour la reconstruction de scènes 3D.

→ le **CHAPITRE 1** dresse un état de l'art des dispositifs de perception dédiée à la reconstruction 3D d'environnements, en détaillant plus particulièrement les systèmes catadioptriques pour leur capacité à fournir une vue à 360° en azimut instantanément [125], [126] ;

---

<sup>1</sup>[http://www.nlm.nih.gov/research/visible/visible\\_human.html](http://www.nlm.nih.gov/research/visible/visible_human.html)

- 
- le **CHAPITRE 2** présente d'une part la solution matérielle envisagée pour des applications en reconstruction de scènes 3D, et s'intéresse d'autre part à l'étalonnage du capteur. A ce titre, un état de l'art détaille les modèles mathématiques et les méthodes d'étalonnage pour des capteurs catadioptriques. Puis, une méthode classique d'étalonnage paramétrique est ensuite proposée. Les résultats obtenus sont présentés et la discussion soulève la nécessité d'une méthode d'étalonnage alternative qui puisse s'affranchir des contraintes liées au modèle.
  - le **CHAPITRE 3** détaille notre contribution pour l'étalonnage de capteurs catadioptriques. Il s'agit d'une méthode non-paramétrique qui considère le capteur sous forme d'une boîte noire. Un modèle discret de transformation est utilisé qui définit les relations entre l'espace 3D et le plan image dans un nouvel espace de représentation et pour un nombre fini de couples point 3D-point 2D. Un maillage sur le nouvel espace des variables est donc défini, tel qu'à un couple point 3D-point 2D correspond un point dans cet espace.

La précision de l'étalonnage est donc fonction de la densité du maillage, c'est-à-dire au nombre de points définissant le maillage. Elle est également fonction des positions des points d'intérêt 3D sur la mire. Une méthode permettant de définir une répartition adaptée des points d'intérêt 3D est ensuite détaillée. La méthode d'étalonnage non-paramétrique est évaluée sur des images de synthèse et pour des images réelles [128], [127];
  - le **CHAPITRE 4** s'intéresse aux problématiques de reconstruction de scènes 3D. Les travaux se sont orientés vers l'implémentation et l'évaluation d'une méthode de modélisation 3D alternative à l'appariement de primitives 2D et triangulation. Il s'agit d'une méthode volumétrique qui modélise la scène sous forme d'un tableau 3D dont chaque cellule est un voxel. Une méthode de reconstruction volumétrique statique est implémentée pour une acquisition simultanée du capteur. Une méthode de reconstruction volumétrique dynamique est évaluée pour des déplacements simples du capteur. Les résultats sont illustrés pour des images réelles et de synthèse [129];
  - une **CONCLUSION ET PERSPECTIVES** reprend les points clés soulevés et propose des perspectives aux problèmes traités.

# Conventions et notations

## Conventions typographiques

- une citation est notée entre guillemets « ». Elle est suivie de la référence bibliographique dont elle provient : « une étape incontournable [...] » [83] ;
- un mot anglais est noté entre guillemets « » avec une police en italique : « *reverse engineering* » ;
- les mots mis en relief sont notés en caractère **gras**.

## Notations mathématiques

- $a$  : un scalaire ;
- $\mathbf{t}$ ,  $A$  : respectivement un vecteur colonne par défaut (en minuscule gras) et une matrice (en majuscule romaine). Les dimensions des matrices ou des vecteurs se déduisent généralement du contexte. Elles peuvent néanmoins être précisées en indice (ex :  $A_{3 \times 3}$ ) ;
- $\mathbf{x} \cdot \mathbf{y}$ ,  $\mathbf{x} \wedge \mathbf{y}$  : respectivement le produit scalaire et le produit vectoriel des vecteurs  $\mathbf{x}$  et  $\mathbf{y}$  ;
- $\mathbf{t}^\top$  : le transposée de  $\mathbf{t}$ ,  $^\top$  étant le symbole de transposition (un vecteur ligne) ;
- $A^{-1}$  : la matrice inverse de  $A$  ;
- $\mathfrak{R}$  : un référentiel (en majuscule gothique) de l'espace bi ou tridimensionnel. Une lettre en indice précise la nature du référentiel :  $\mathfrak{R}_W$  référentiel  $W$ (M)onde,  $\mathfrak{R}_M$  référentiel Miroir,  $\mathfrak{R}_P$  référentiel Pixel, etc ;
- $\mathbf{q} = (u, v)^\top$ ,  $\mathbf{Q} = (X, Y, Z)^\top$  : respectivement un point 2D et un point 3D. Une lettre en indice peut éventuellement préciser la nature du référentiel dans lequel le point est exprimé afin d'éviter toute ambiguïté :  $\mathbf{Q}_W = (X_W, Y_W, Z_W)^\top$  ;
- $\hat{\mathbf{q}} = (\hat{u}, \hat{v})^\top$  : l'estimé de  $\mathbf{q}$  ;
- $g, g^*$  : respectivement les fonctions de projection et rétro-projection ;

- 
- $\mathcal{L}, \ell$  : respectivement une droite en 3D et une droite en 2D ;
  - Un plan est désigné par la lettre grecque 'pi' en majuscule ou en minuscule :  $\Pi, \pi$  ;
  - Les ensembles sont notés entre accolades :  $\{, \}$ . Les indices sont notés  $i, j, m, n$  ou  $p$  et permettent de sélectionner un élément particulier parmi un ensemble. Par exemple un ensemble de points 3D sera noté  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_i, \dots, \mathbf{Q}_I\}$  avec  $i \in \{1, \dots, I\}$  et  $\mathbf{Q}_i$  étant le  $i^{\text{ème}}$  point 3D.  
Dans un même paragraphe, un indice ne fait référence qu'à un unique ensemble. Pour deux paragraphes distincts, un même indice peut qualifier deux ensembles de nature différente.
  - Les entités géométriques (sphère, cercle en 3D ou en 2D, volume, etc.) sont notées en écriture sans empattement. Les formes géométriques 3D sont en majuscule et celles en 2D sont en minuscule : une sphère  $S$ , un cercle en 3D  $C$ , un volume  $V$ , un cercle en 2D  $c$ , etc.

# Chapitre 1

## Capteurs de perception de scènes 3D

### Sommaire

---

<b>1.1</b>	<b>Introduction</b>	<b>8</b>
<b>1.2</b>	<b>Perception 3D par capteurs actifs : télémétrie et triangulation</b>	<b>10</b>
1.2.1	Télémétrie	10
1.2.2	Triangulation active en lumière structurée	11
1.2.3	Synthèse	13
<b>1.3</b>	<b>Perception 3D par capteurs passifs : la vision</b>	<b>16</b>
1.3.1	Capteurs de vision directionnelle	17
1.3.2	Capteurs de vision omnidirectionnelle	22
1.3.3	Synthèse	29
<b>1.4</b>	<b>Capteurs de stéréovision omnidirectionnelle</b>	<b>32</b>
1.4.1	Quelques exemples de capteurs	34
1.4.2	Géométrie épipolaire	34
<b>1.5</b>	<b>Conclusions</b>	<b>42</b>

---

## 1.1 Introduction

Le principal objectif du programme *OmniViss* est le développement d'une architecture matérielle pour la numérisation 3D de scènes. Le capteur peut être en position fixe : le but est alors de détecter et d'analyser le mouvement d'objets présents dans la scène. Le capteur est mobile : le but recherché est alors de permettre à un véhicule non habité de reconstruire en trois dimensions son environnement pour y naviguer de façon autonome. Ces deux applications, bien qu'un dénominateur commun soit la reconstruction de scènes 3D, s'adressent à des domaines applicatifs différents. Les exigences techniques quant à la réalisation de la solution matérielle devront donc prendre en compte les spécificités de ces deux applications. Idéalement, le système développé devra permettre :

- une acquisition 3D non intrusive (discrétion, furtivité) ;
- une observation large de l'environnement pour une numérisation 3D temps-réel ;
- afin de rester compatible avec une solution industrialisable, les technologies bas-coût seront privilégiées.

Les technologies d'acquisition de données 3D font appel à des procédés très différents selon qu'elles se destinent au contrôle géométrique, à la réalisation de modèles numériques ou à l'imagerie médicale. On distingue principalement deux catégories [113] :

- les techniques de **numérisation 3D à contact** qui utilisent essentiellement des capteurs mécaniques de type palpeurs à suivi de forme ou à déclenchement [30]. Ces capteurs sont fréquemment utilisés dans les machines à mesurer tridimensionnelles ou dans les machines à commande numérique. Les précisions de mesure permises sont importantes, de l'ordre du micromètre. Ces capteurs sont néanmoins sujets à des temps d'acquisition longs ce qui limite les dimensions des volumes à numériser ;
- les techniques de **numérisation 3D sans contact** [82], [31] qui utilisent des procédés de mesure par télémétrie, triangulation, tomographie ou des technologies optiques. La reconstruction par tomographie utilise les propriétés d'émission ou d'absorption de corps irradiés, l'objectif étant d'estimer l'image de l'émetteur à partir des projections obtenues. Elle est principalement utilisée en imagerie médicale mais trouve un intérêt certain en milieu industriel pour des applications en contrôle non-destructif. Les techniques par procédés optiques sont l'interférométrie, la variation d'astigmatisme, les systèmes confocaux, la défocalisation dynamique, la conoscopie ou la chromatie. Les temps de numérisation de ces procédés de mesure sont souvent longs et les portées souvent faibles.

---

Au vu des applications envisagées pour le programme *OmniViss*, l'état de l'art des systèmes de perception 3D est volontairement axé sur les technologies permettant la numérisation 3D d'environnements. Le lecteur souhaitant un inventaire exhaustif des techniques de numérisation 3D à contact ou sans contact peut se référer aux articles de Laurgeau [82] et Charron [29], [30], [31]. Les technologies de capteurs d'acquisition 3D détaillées dans les paragraphes suivants concernent donc les capteurs de numérisation 3D sans contact **actifs** ou **passifs**. Ces deux technologies se différencient principalement par le fait que les capteurs interagissent avec la cible, par émission d'un signal ou non.

Le **CHAPITRE 1** est structuré de la manière suivante : une première partie traite des capteurs d'acquisition 3D actifs, tel que les télémètres ou les capteurs actifs en lumière structurée. Une deuxième partie s'intéresse aux capteurs passifs, c'est-à-dire aux systèmes de vision, en particulier ceux qui possèdent un champ d'observation de 360° en azimut. Une troisième partie s'intéresse aux capteurs stéréoscopiques et détaille la solution matérielle retenue pour les applications du programme *OmniViss*. Une synthèse clôt ce chapitre et reprend les points clés abordés au cours des différents paragraphes.

## 1.2 Perception 3D par capteurs actifs : télémétrie et triangulation

Les capteurs actifs émettent un signal pour mesurer la distance entre un capteur et une cible observée. On distingue principalement les capteurs à **télémétrie** et les capteurs **actifs par triangulation**. Les capteurs à télémétrie utilisent l'énergie d'une onde incidente rétrodiffusée<sup>2</sup> par une cible pour mesurer la distance entre le capteur et l'objet observé. Les signaux émis peuvent être des ondes électromagnétiques lumineuses opérant dans le visible, le proche infrarouge ou l'ultraviolet, des ondes électromagnétiques radio-fréquences ou des ondes mécaniques acoustiques. Le capteur peut fonctionner en régime impulsionnel ou en régime permanent sinusoïdal à fréquence fixe ou variable. En régime impulsionnel, la mesure de distance est basée sur le calcul du temps de parcours d'un signal entre son émission et sa réception. En régime permanent sinusoïdal, la mesure de distance est soit une mesure de déphasage, soit une différence de fréquence entre l'onde émise et l'onde reçue. Les capteurs actifs par triangulation sont composés d'une source lumineuse et d'un capteur vidéo. Cette méthode de mesure de distance est dite en lumière structurée [48], [116]. Dans sa forme la plus simple, la lumière structurée est un unique rayon. Les projections de plans lumineux ou de patrons bidimensionnels sont également utilisées.

Cette section débute par une revue des technologies d'acquisition 3D par télémétrie. Les capteurs utilisant le principe de la mesure du temps de vol, de la mesure du déphasage et de la mesure de la fréquence sont détaillés. Puis, les systèmes d'acquisition par projection d'un motif lumineux sont présentés. Cette section se termine par une synthèse de ces deux technologies en termes d'avantages et d'inconvénients par rapport aux exigences des applications du programme *OmniViss*.

### 1.2.1 Télémétrie

#### 1.2.1.1 Mesure du temps de vol

La **télémétrie par mesure du temps de vol** est une méthode de mesure de distance basée sur l'évaluation du temps de propagation aller et retour d'un signal émis par une

---

<sup>2</sup>Lorsque l'énergie frappe une cible, elle peut être projetée dans toutes les directions. La partie de l'énergie qui est retournée directement dans la direction d'où elle vient est dite rétrodiffusée.

source et rétrodiffusée par une cible (cf. figure 1.1). Sous certaines hypothèses (faisceaux d'émission et de réception colinéaires ou quasi-parallèles, illumination de la cible de façon quasi-ponctuelle, etc. [51]), une fraction du signal émis est rétrodiffusée<sup>3</sup> dans une direction voisine de celle du signal incident. Si  $\tau$  est le temps de parcours aller et retour du signal,  $v$  est la vitesse de propagation du signal dans le milieu considéré, la distance  $D$  à mesurer est donnée par la relation :  $D = \frac{1}{2}v\tau$ . Cette technique de mesure est fréquemment utilisée pour les capteurs **radars** (« *radio detection and ranging* »), **lidars** (« *light detection and ranging* ») mais reste néanmoins employée lorsque la distance entre le capteur et la cible est grande.

### 1.2.1.2 Mesure du déphasage

Lorsque les distances sont faibles, la **télémétrie par mesure du déphasage** remplace la télémétrie par mesure du temps de vol. Une onde modulée en amplitude par un signal sinusoïdal de fréquence fixe est émise vers la cible (cf. figure 1.1). L'onde réfléchiée par l'objet est également une onde modulée en amplitude mais déphasée par rapport à l'onde émise. Le déphasage  $\phi$  entre l'onde émise et l'onde réfléchiée est proportionnel à la distance  $D$  entre le capteur et la cible, à la vitesse de propagation de l'onde  $v$  et à la fréquence de modulation  $F_m$ . Ce déphasage est donné par la relation suivante :  $\phi = 4\pi F_m \frac{D}{v} \bmod (2\pi)$ .

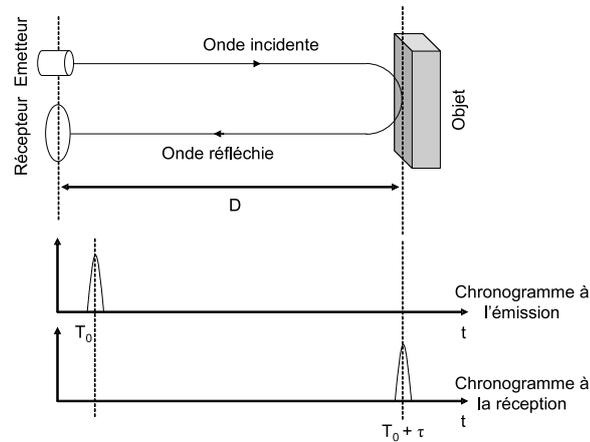
### 1.2.1.3 Mesure de la fréquence

La **télémétrie par mesure de la fréquence** utilise l'émission d'une onde modulée sinusoïdalement en amplitude par un signal lui-même modulé linéairement en fréquence (cf. figure 1.1). La différence de fréquence entre l'onde émise  $F_e(t)$  et l'onde reçue  $F_r(t)$  est proportionnelle à la distance  $D$  entre le capteur et la cible :  $F_e(t) - F_r(t) = \Delta F \frac{2D}{vT}$ , où  $\Delta F$  est l'amplitude de fréquences entre le signal émis et celui reçu et  $T$  est la période des signaux.

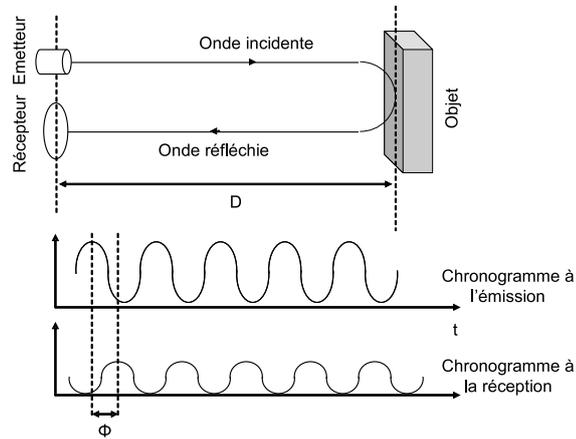
## 1.2.2 Triangulation active en lumière structurée

Dans sa forme la plus simple, la **lumière structurée** est un unique rayon émis par une source et faisant apparaître sur l'objet un point de surbrillance (cf. figure 1.2a).

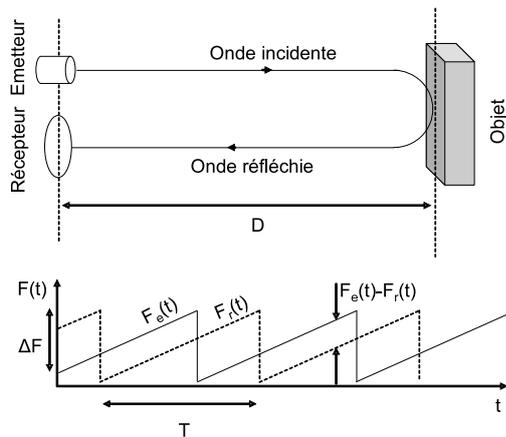
<sup>3</sup>Un signal rétrodiffusé représente la portion d'un signal acoustique ou électromagnétique que la cible redirige dans une direction très voisine de celle du faisceau incident [51].



(a) Télémétrie par mesure du temps de vol



(b) Télémétrie par mesure du déphasage



(c) Télémétrie par mesure de la fréquence

FIG. 1.1 – Technologie de capteurs par télémétrie [82].

Lorsque la direction du rayon est connue et le capteur calibré (c'est-à-dire qu'à tout point 2D correspond une ligne de vue dans l'espace 3D), il est possible de déterminer la position du point 3D définie comme l'intersection de la ligne de vue 3D et du rayon lumineux. La lumière structurée peut également être un plan de lumière (on parle alors de profilométrie) (cf. figure 1.2b). Ce plan de lumière est généré par une source lumineuse (généralement laser) devant laquelle on place une lentille cylindrique. Ce plan, d'équation connue, fait apparaître sur l'objet un segment de surbrillance observé par le capteur d'images. Aux pixels éclairés décrivant le profil sont associées des directions dans l'espace 3D (à condition que le capteur d'images soit étalonné). Il s'agit donc de déterminer les points 3D d'intersection du plan de lumière et des lignes de vue 3D associées aux pixels. Enfin, la lumière structurée peut être un motif bidimensionnel que l'on projette sur l'objet à numériser (cf. figure 1.2c). Il peut s'agir d'une grille (matrice de points) ou d'un faisceau de lignes. L'estimation de la distance entre le capteur et l'objet peut se faire par une étude de la déformation du patron projeté ou par une mise en correspondance entre la matrice de points projetée et les données observées par le capteur.

### 1.2.3 Synthèse

Nous avons vu dans les paragraphes précédents deux technologies de capteurs permettant l'acquisition de données 3D : la première est dite par télémétrie et la seconde est dite active par projection de lumière structurée.

La première technologie de capteurs utilise l'émission et la réception d'ondes radiofréquences (radars), lumineuses (lidars) ou ultra-sonores pour mesurer la distance entre le capteur et la cible visée. Mis à part le télémètre à ultrason plutôt destiné à l'acquisition de données 3D en environnements intérieurs, les radars ou les lidars sont principalement utilisés pour des applications nécessitant des portées importantes : la vitesse de propagation des ondes ainsi que les longueurs d'onde des signaux émis ne permettent pas la numérisation précise d'objets proches [82]. Dans sa version classique, la télémétrie souffre également d'une résolution azimutale faible. La synthèse d'ouverture, qui consiste à augmenter artificiellement la résolution azimutale du capteur, permet de pallier cet inconvénient. Cependant, ce processus nécessitant des déplacements successifs de la source d'émission, requiert des post-traitements conséquents sur les données reçues et ne peut donc être compatible avec des traitements temps-réel. Un autre inconvénient de cette technologie tient aux surfaces de certains matériaux, qui dans certains cas possèdent un coefficient d'absorption élevé ne permettant pas de rétro-diffuser l'onde émise, ou qui dans d'autres

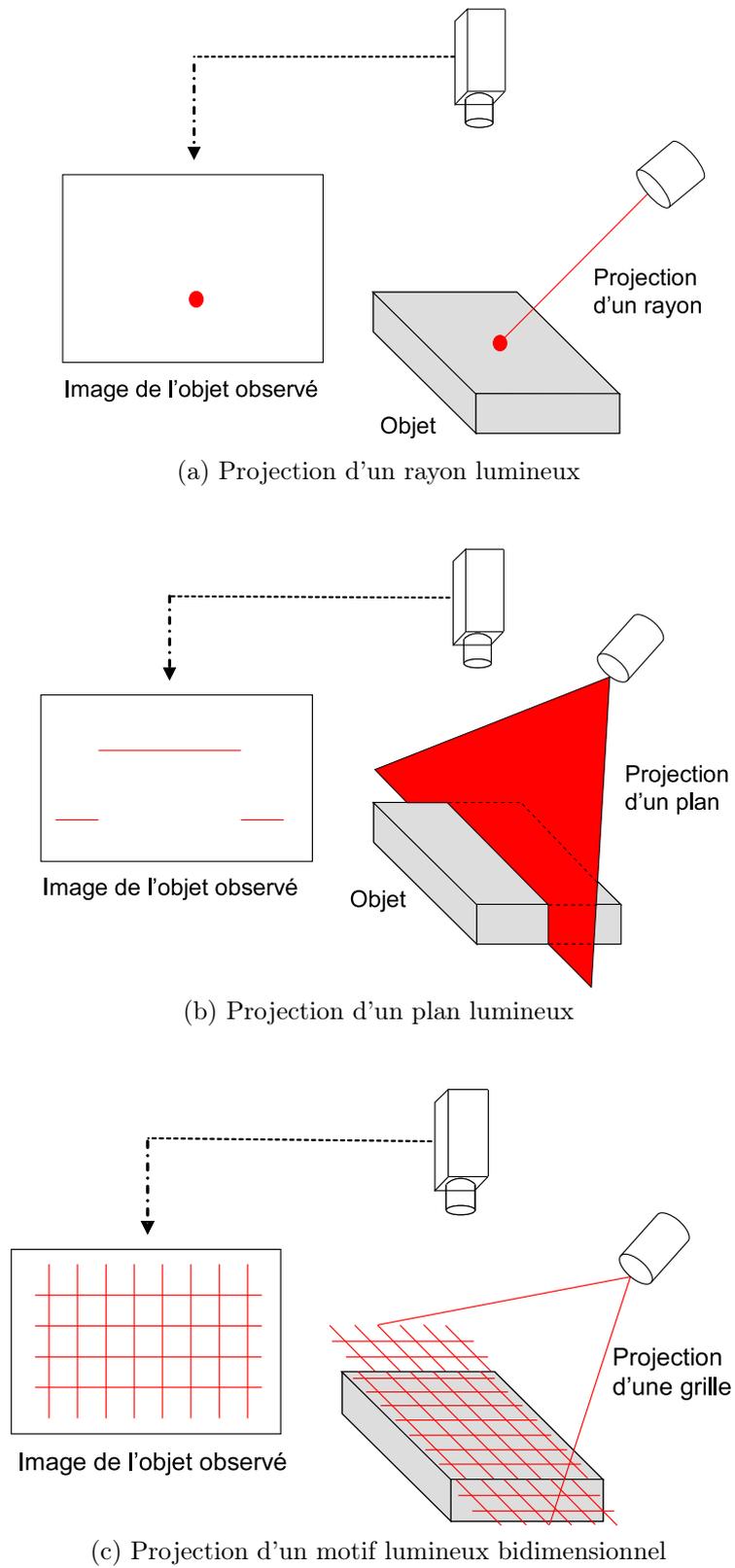


FIG. 1.2 – Technologie de capteurs par triangulation active en lumière structurée [48].

se comportent comme des miroirs réfléchissant l'onde incidente dans une direction opposée à celle émise [82].

La seconde technologie est dite active par projection de lumière structurée. Elle utilise une source lumineuse et l'observation par un capteur vidéo du motif lumineux projeté. Cette technologie tire profit des avantages liés à l'utilisation de capteurs vidéos, c'est-à-dire l'acquisition d'une quantité d'informations importante instantanément, et n'est pas soumis aux inconvénients des systèmes passifs traditionnels [44] (éclairage, surface uniforme, etc.). Lorsqu'un point ou un plan est projeté, l'estimation de distance est assez simple : il s'agit de déterminer l'intersection du ou des rayons lumineux avec les lignes de vue 3D (sous réserve que le capteur soit calibré). L'inconvénient principal apparaît lorsqu'il faut numériser l'intégralité d'un objet ou d'une scène 3D. La source de lumière doit être pilotée horizontalement et verticalement afin de permettre l'acquisition de la totalité des données 3D. Cette solution est donc difficilement compatible avec des applications nécessitant des traitements temps-réel. Une solution consiste en la projection d'un motif bidimensionnel. Aucun système mécanique de balayage n'est alors nécessaire ce qui permet d'accélérer significativement la phase d'acquisition de données. Néanmoins, l'estimation de la distance est plus complexe. En effet, la mise en correspondance entre la matrice de points projetée et les données observées par le capteur n'est pas triviale puisque rien ne permet de distinguer dans l'image un noeud d'un autre noeud [48]. Des contraintes permettent de simplifier le problème, mais des ambiguïtés subsistent. Une solution repose sur l'utilisation de patrons de lumière structurée et codée [115]. Cette technique de codage rend alors le problème de mise en correspondance univoque.

Enfin, un inconvénient commun aux deux technologies est le caractère actif de la mesure, c'est-à-dire que l'évaluation de la distance entre le capteur et la cible utilise l'émission d'un signal. Lorsque des impératifs de discrétion sont requis pour les applications, l'émission d'un signal pour acquérir les données 3D est peu envisageable.

### 1.3 Perception 3D par capteurs passifs : la vision

Les systèmes de vision sont bien adaptés à la reconstruction 3D d'environnements puisqu'ils fournissent une quantité d'informations importante en un temps très court. Cette quantité d'informations est fonction de ce que voit le capteur. Si l'on considère une sphère d'observation et un capteur vidéo placé en son centre (cf. figure 1.3) ; le capteur est dit **standard** ou **directionnel** lorsqu'il observe une portion limitée, strictement inférieure à un hémisphère, de la sphère d'observation. Lorsque son champ d'observation couvre l'intégralité de la sphère d'observation, il est qualifié d'**omnidirectionnel** [118], [100]. Enfin, il est dit **panoramique** lorsque son champ d'observation couvre  $360^\circ$  autour d'un axe. En pratique, les termes panoramique et omnidirectionnel ne sont pas différenciés. Tous deux renvoient à une même notion : celle d'un capteur d'images dont le champ d'observation couvre  $360^\circ$  en azimut [106]. Enfin, les systèmes permettant d'observer une zone plus importante qu'un capteur directionnel sont qualifiés de capteur **large champ**.

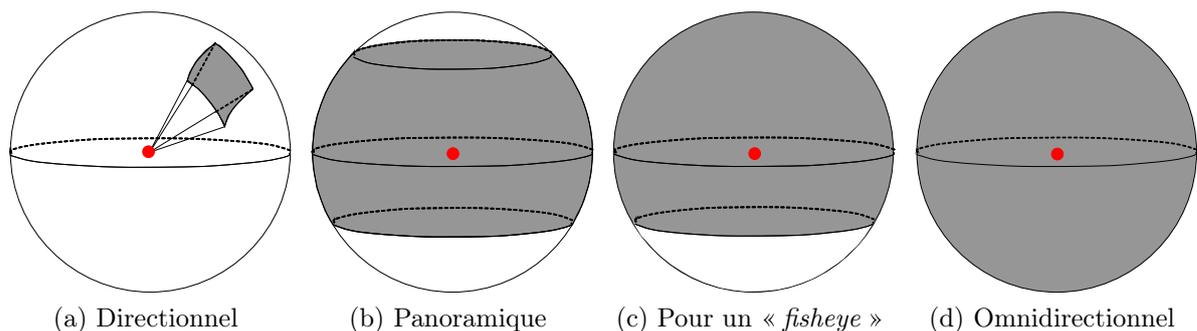


FIG. 1.3 – Différents champs d'observation des capteurs de vision [118], [100], [106]. Le point rouge matérialise le capteur vidéo placé au centre de la sphère d'observation. a) champ d'observation directionnel : une portion limitée de la sphère d'observation est vue. b) champ d'observation panoramique d'un capteur catadioptrique : le capteur visualise une bande de vue. c) champ d'observation panoramique d'un objectif « *fisheye* ». d) champ d'observation omnidirectionnel : l'intégralité de la sphère d'observation est vue.

La vision omnidirectionnelle a connu un essor important dans les années 1990 pour des domaines d'applications variés, notamment pour la robotique mobile autonome. Elle y a supplanté les capteurs classiques directionnels parce qu'elle permet une augmentation significative du champ d'observation « sans rupture géométrique et temporelle » [14]. Elle trouve aujourd'hui des utilisations pour d'autres champs d'applications : Huang & Trivedi [71] et Layerle & al. [84] ont proposé l'utilisation d'un seul capteur catadioptrique positionné dans un véhicule pour capturer simultanément le visage d'un conducteur et la scène de conduite en vue panoramique. D'un point de vue théorique, l'intérêt est majeur

puisque la vision omnidirectionnelle « pose la question de la généralisation des méthodes utilisées en vision perspective classique » [14].

Cette section est structurée de la manière suivante : un premier paragraphe s'intéresse aux capteurs de vision directionnelle. Le modèle sténopé est détaillé et les différentes aberrations apparaissant lors du processus de formation d'images sont listées. Une deuxième partie s'intéresse aux capteurs de vision omnidirectionnelle. Les techniques permettant une augmentation significative du champ d'observation sont exposées, et une attention particulière est portée aux systèmes combinant un miroir et une caméra. Finalement, une synthèse est proposée permettant de statuer sur la pertinence des technologies inventoriées en regard des applications visées par le programme *OmniViss*.

### 1.3.1 Capteurs de vision directionnelle

#### 1.3.1.1 Le modèle de projection sténopé (cf. Annexe A)

Le **modèle de projection sténopé** (trou d'épingle ou « *pinhole* ») est le modèle de projection le plus couramment utilisé en **Vision par Ordinateur**. Il permet de modéliser de manière simple une projection perspective entre l'espace 3D et le plan image d'un capteur vidéo (cf. figure 1.4).

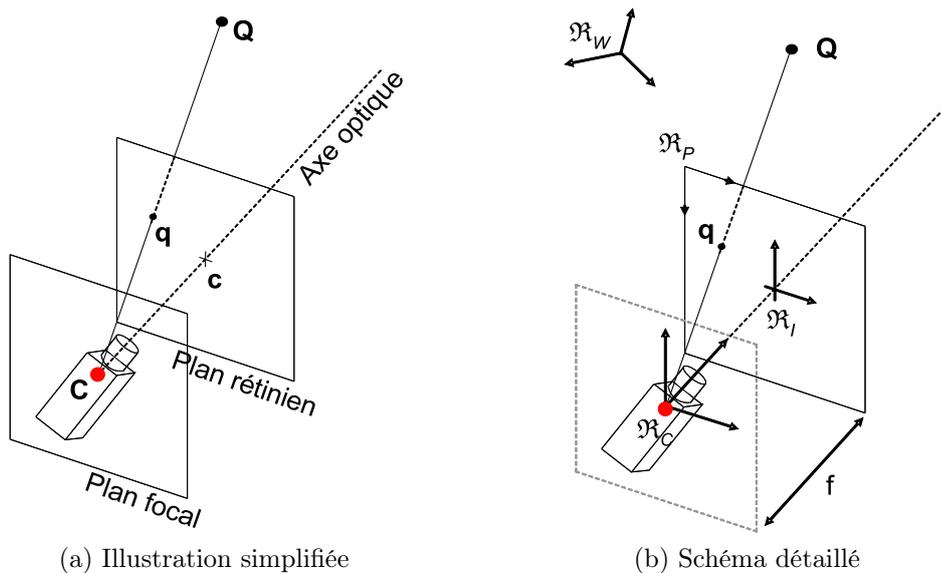


FIG. 1.4 – Le modèle sténopé [149].

La simplicité du modèle sténopé réside dans sa construction géométrique : un point 3D  $\mathbf{Q}$  se projette sur le plan image en un point 2D  $\mathbf{q}$  par une projection centrale de centre  $\mathbf{C}$ .  $\mathbf{q}$  est l'intersection de la demi-droite d'origine  $\mathbf{C}$ , passant par  $\mathbf{Q}$ , avec le plan image. Le modèle est également simple d'un point de vue mathématique : en coordonnées homogènes, la projection de  $\mathbf{Q} = (X, Y, Z, 1)^\top$  en  $\mathbf{q} = (u, v, 1)^\top$  est une application linéaire (cf. équation 1.1). La matrice  $M$  est appelée **matrice de projection perspective**, le plan image est appelé **plan rétinien**. Le plan contenant le centre de projection  $\mathbf{C}$  est appelé **plan focal**. L'**axe optique** est défini comme l'axe perpendiculaire au plan image passant par  $\mathbf{C}$ . Le **point principal**  $\mathbf{c}$  est l'intersection de l'axe optique et du plan image.

$$\mathbf{q} = M_{3 \times 4} \mathbf{Q} \quad (1.1)$$

La matrice de projection  $M$  ne fournit pas d'indications sur les grandeurs physiques mises en jeu durant le processus de prises d'images [149]. Il est donc judicieux de décomposer  $M$  sous forme d'un produit matriciel faisant clairement apparaître les différents paramètres intervenant durant le processus d'acquisition. Il s'agit des paramètres **intrinsèques** qui décrivent les caractéristiques propres au système d'acquisition et les paramètres **extrinsèques** qui décrivent la position relative du capteur vis à vis d'un repère monde  $\mathfrak{R}_W$ .

L'équation de projection d'un point de l'espace 3D en coordonnées homogènes  $\mathbf{Q} = (X, Y, Z, 1)^\top$  défini dans un référentiel monde  $\mathfrak{R}_W$  et d'un point 2D en coordonnées homogènes  $\mathbf{q} = (u, v, 1)^\top$  défini dans un référentiel pixel  $\mathfrak{R}_P$  est donnée par l'équation 1.2. Cette équation fait apparaître :

- une matrice homogène de changement de base  $S = (\mathbf{R} \mid \mathbf{t})$  modélisant les paramètres extrinsèques.  $\mathbf{R}$  est une matrice de rotation et  $\mathbf{t}$  un vecteur de translation ;
- la matrice homogène de projection perspective  $P_I$  ;
- la matrice homogène des paramètres intrinsèques  $K_P$  pour une projection perspective dépendant de la focale  $f$ , du nombre de pixels par unité de longueur sur chacun des axes  $(k_u, k_v)$ , des coordonnées du centre de l'image  $\mathbf{c} = (u_0, v_0)^\top$  dans  $\mathfrak{R}_P$ .

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \overbrace{\begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}^{K_P} \overbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}}^{P_I} \overbrace{\begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}_{3 \times 1}^\top & 1 \end{pmatrix}}^S \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (1.2)$$

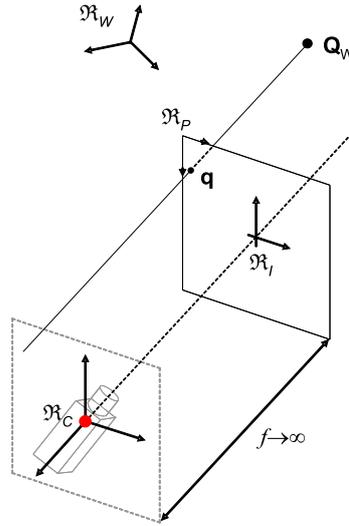


FIG. 1.5 – Le modèle de projection orthographique [61]

### 1.3.1.2 Le modèle de projection orthographique (cf. Annexe B)

Le modèle de projection orthographique est un cas limite de la projection perspective pour lequel la focale  $f$  tend vers l'infini (cf. figure 1.5). Par conséquent, le rapport  $f/z$  tend vers 1. Les rayons lumineux sont alors parallèles entre eux et parallèles à l'axe optique. Le modèle de projection orthographique établissant la relation entre un point 3D  $\mathbf{Q}_C = (X_C, Y_C, Z_C, 1)^\top$  défini dans un référentiel caméra  $\mathfrak{R}_C$  et sa projection  $\mathbf{q} = (u, v, 1)^\top$  exprimée dans un référentiel pixel  $\mathfrak{R}_P$  s'écrit alors comme suit :

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \overbrace{\begin{pmatrix} k_u & 0 & 0 & u_0 \\ 0 & k_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}^{K_O} \overbrace{\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}^{O_I} \begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix} \quad (1.3)$$

$K_O$  est définie comme la matrice homogène des paramètres intrinsèques pour une projection orthographique et  $O_I$  est la matrice homogène de projection orthographique.

### 1.3.1.3 Les distorsions

Les modèles présentés supposent une projection perspective ou orthographique parfaite. Ils ne tiennent pas compte d'un certain nombre de phénomènes qui conduisent à des écarts

entre la projection réelle d'un point de l'espace 3D et celle estimée par les modèles. Ces phénomènes, appelés **aberrations**, apparaissent généralement lorsque la distance entre le point 3D et l'axe optique est grande et que les inclinaisons des rayons lumineux sont faibles [119]. Des approximations dans le positionnement des lentilles des objectifs peuvent également conduire à des écarts entre les projections réelle et estimée. Dans le cas d'une projection orthographique qui nécessite l'utilisation de lentilles télécentriques, le centre de projection n'est pas idéalement situé à l'infini. De ce fait, les rayons lumineux se sont plus parfaitement parallèles entre eux et parallèles à l'axe optique, ce qui introduit des aberrations dans l'image.

Ces phénomènes peuvent être minimisés par des procédés mécaniques de type usinage et assemblage haute précision. Ils peuvent être également pris en compte dans l'écriture du modèle par une transcription des phénomènes sous forme mathématique [96]. Nous proposons ci-dessous un inventaire non-exhaustif des aberrations de capteurs vidéos [61] :

→ les **distorsions radiales** et **tangentiels** : les distorsions radiales, de type coussinet ou barillet, sont dues aux courbures des lentilles. Les distorsions tangentiels apparaissent lorsque les lentilles ne sont pas idéalement positionnées perpendiculairement à l'axe optique. Des approches mathématiques existent afin d'interpréter au mieux ces phénomènes. La modélisation la plus couramment utilisée consiste à approximer les distorsions par des formes polynomiales [24], [96], [83].

Soit  $\mathbf{Q}$  un point 3D,  $\mathbf{q}_s$  la projection estimée de  $\mathbf{Q}$  par le modèle sténopé et  $\mathbf{q}_d$  la projection réelle. La relation entre  $\mathbf{q}_s$  et  $\mathbf{q}_d$  est donnée par :

$$\mathbf{q}_d = \mathbf{q}_s + \Delta \quad (1.4)$$

Où  $\Delta$  est le vecteur d'aberration ou de distorsion. Ce vecteur est la somme des distorsions radiales  $\Delta_R$  et des distorsions tangentiels  $\Delta_T$ , les distorsions tangentiels étant l'association des **distorsions de décentrage** et des **distorsions prismatiques** introduites par Weng & al. [164], [97]. Le vecteur d'aberration  $\Delta$  est alors défini par :

$$\Delta = \Delta_R + \Delta_T \quad (1.5)$$

Avec :

- $\Delta_R = (\delta_{R_x}, \delta_{R_y})^\top$  : les distorsions radiales sont données par l'équation 1.6 :

$$\begin{aligned} \delta_{R_x} &= (u - u_0) \sum_{i=1}^{\infty} k_i \rho^{2i} \\ \delta_{R_y} &= (v - v_0) \sum_{i=1}^{\infty} k_i \rho^{2i} \end{aligned} \quad (1.6)$$

- $\Delta_T = (\delta_{T_x}, \delta_{T_y})^\top$  : les distorsions tangentiels sont données par l'équation 1.7 :

$$\begin{aligned}\delta_{T_x} &= (p_1 (\rho^2 + 2(u - u_0)^2) + 2p_2 (u - u_0)(v - v_0)) \left(1 + \sum_{j=1}^{\infty} p_{2+j} \rho^{2j}\right) \\ \delta_{T_y} &= (p_2 (\rho^2 + 2(v - v_0)^2) + 2p_1 (u - u_0)(v - v_0)) \left(1 + \sum_{j=1}^{\infty} p_{2+j} \rho^{2j}\right)\end{aligned}\quad (1.7)$$

Où :

- $k_i$  : les coefficients des distorsions radiales. Pour l'étalonnage du capteur, l'estimation des deux premiers paramètres est suffisante  $(k_1, k_2)$  [97],
  - $p_j$  : les coefficients des distorsions tangentielles. Pour l'étalonnage du capteur, l'estimation des paramètres  $(p_1, p_2)$  est suffisante [97],
  - $\rho^2 = (u - u_0)^2 + (v - v_0)^2$  : distance radiale depuis le centre de l'image  $\mathbf{c} = (u_0, v_0)^\top$  ;
- les **aberrations chromatiques** sont dues à une variation de l'indice de réfraction du matériau composant les lentilles en fonction de la longueur d'onde de la lumière qui les traverse. Il en résulte une variation de la distance focale, de sorte que la mise au point ne peut être effectuée simultanément pour toutes les couleurs du spectre. L'utilisation d'un achromat, c'est-à-dire l'association de deux lentilles composées de verres différents et de courbures différentes, l'une convergente et l'autre divergente, permet de réduire ces effets ;
- les **aberrations sphériques** sont dues à une plus grande convergence des rayons optiques en périphérie de la lentille que ceux proches de l'axe optique, ce qui conduit à l'obtention d'une tâche floue sur le plan image. Ce défaut est lié à la forme sphérique de la majorité des lentilles utilisées. Afin d'éviter ces effets, l'utilisation de lentilles asphériques de forme paraboloidale est préconisée ;
- les **aberrations de coma** tendent à faire apparaître une traînée vers le centre de l'image. Ceci est dû au fait que des rayons lumineux parallèles mais inclinés par rapport à l'axe optique ne convergent pas tous en un même point ;
- les **aberrations d'astigmatisme** résultent d'une variation de la distance focale en fonction de la direction du rayon lumineux ;
- les **aberrations de courbure de champ** sont dues au fait qu'un objet plan de grande dimension ne se projette pas sur une surface plane mais paraboloidale ;
- les **aberrations de vignettage** apparaissent lorsque la distance focale est importante et le diaphragme fermé. Les bords de l'image sont alors assombris.

## 1.3.2 Capteurs de vision omnidirectionnelle

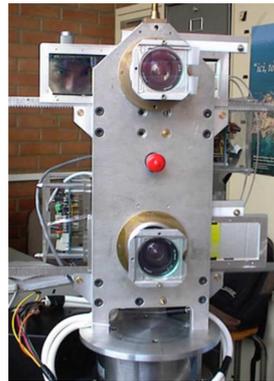
### 1.3.2.1 Comment accroître le champ de vision ?

Il existe différentes manières d'accroître le champ d'observation. Les choix technologiques dépendent étroitement des applications visées. On peut par exemple envisager un capteur directionnel en rotation autour d'axes motorisés. On parle alors de **capteurs cylindriques** ou **sphériques**. On peut également penser à multiplier les systèmes d'acquisition d'images. Les caméras forment alors un **réseau de caméras synchrone** ou **asynchrone**. On peut également utiliser des **objectifs large champ** de type « *fisheye* » par exemple. Enfin, on peut se servir conjointement d'un capteur directionnel et d'un miroir convexe. Il s'agit alors d'un capteur **catadioptrique**. Les recueils d'articles des ouvrages édités par Benosman & Kang [16], puis par Daniilidis & Klette [37] ainsi que le numéro spécial sur la vision omnidirectionnelle publié par la revue Traitement du Signal [107], proposent un inventaire détaillé et dense des avancées récentes en vision omnidirectionnelle.

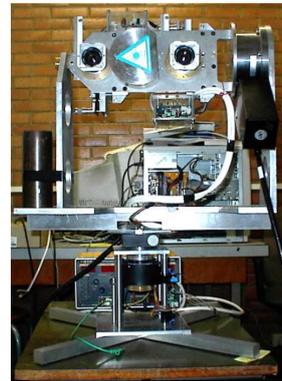
### 1.3.2.2 Capteurs cylindriques ou sphériques

Les systèmes cylindriques et-ou sphériques sont constitués de capteurs directionnels en rotation autour d'axes verticaux [133], [145], [146] et-ou de deux axes, l'un vertical et l'autre horizontal [42] (cf. figure 1.6). Les appellations cylindrique et sphérique qualifient les champs d'observation couverts par le capteur : pour un capteur cylindrique, le champ d'observation décrit un cylindre de révolution. Dans le cas d'un système sphérique, le champ d'observation couvre l'intégralité de la sphère d'observation.

Dans sa version la plus simple, le système est constitué d'un capteur CCD (ou CMOS) linéaire fixé sur un axe, lui-même placé sur un mécanisme de positionnement angulaire. L'image panoramique résultante est obtenue en concaténant les images prises pour les différentes positions du capteur. Lorsque les positions angulaires du capteur ne sont pas contraintes, l'image panoramique est le résultat d'un mosaïcage des images acquises aux différentes positions du capteur. Lorsqu'un capteur matriciel est utilisé, l'image panoramique est également le résultat d'une mosaïque des images acquises par le capteur. Pour cette solution technique, il est souhaitable que l'axe de rotation du système soit confondu avec le centre optique du capteur, ceci afin d'éviter l'apparition d'une disparité dans l'image panoramique finale [61].



(a) Capteur cylindrique



(b) Capteur sphérique



(c) Image obtenue par un capteur cylindrique

FIG. 1.6 – Exemples de capteurs cylindrique et sphérique développés conjointement par le Laboratoire des Instruments et Systèmes d’Île de France (LISIF, nouvellement ISIR) et l’Institut Supérieur d’Electronique de Paris (ISEP) [133], [42].

### 1.3.2.3 Réseau de caméras

Une autre manière d’agrandir le champ d’observation d’un capteur consiste en l’augmentation du nombre de capteurs le constituant. Chacun d’eux observe une zone de l’espace 3D et l’image panoramique est alors le résultat d’un mosaïcage des images fournies par les différents capteurs. L’association de ces différents systèmes d’acquisition forme un réseau de caméras qui peut être synchrone, lorsque les acquisitions sont faites au même instant, ou asynchrone. Les figures suivantes (cf. figure 1.7) présentent quelques exemples de réseaux de caméras : le capteur Ladybug<sup>4</sup> développé par Point Grey Research, le capteur RingCam<sup>5</sup> proposé par Microsoft Research et la série de capteurs Astro Sensor Series<sup>6</sup> développée par la société ViewPlus.

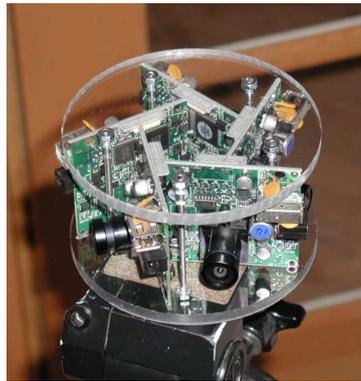
<sup>4</sup><http://www.ptgrey.com/products/ladybug2/index.asp>

<sup>5</sup><http://research.microsoft.com/~rcutler/ringcam/ringcam.htm>

<sup>6</sup><http://www.viewplus.co.jp/products/sos/astro-e.html>



(a) Ladybug2



(b) RingCam



(c) Astro serie



(d) Image par le capteur Ladybug2

FIG. 1.7 – Quelques exemples de capteurs à réseau de caméras.

#### 1.3.2.4 Objectifs grand angle

Les objectifs grand angle font souvent référence aux objectifs « *fisheye* » (oeil de poisson) qui se fixent sur les caméras standards à la place des objectifs classiques (cf. figure 1.8<sup>7</sup>). Ils permettent également l'acquisition d'images omnidirectionnelles.

#### 1.3.2.5 Capteurs catadioptriques

Un capteur catadioptrique est l'association d'une caméra et d'un miroir de courbure convexe ou concave, la caméra observant le miroir (cf. figure 1.9). L'appellation catadioptrique est la concaténation de deux termes : le premier, **catoptrique**, fait référence à la réflexion de la lumière et désigne donc le miroir. Le second, **dioptrique**, fait référence à la réfraction de la lumière et désigne donc la lentille.

Le premier brevet industriel d'un système catadioptrique a été proposé par Rees en 1970 [131]. Ce capteur, dédié à des applications militaires, associe un capteur vidéo classique à

<sup>7</sup>[http://upload.wikimedia.org/wikipedia/commons/thumb/b/b4/Fisheye\\_photo.jpg/227px-Fisheye\\_photo.jpg](http://upload.wikimedia.org/wikipedia/commons/thumb/b/b4/Fisheye_photo.jpg/227px-Fisheye_photo.jpg)



(a) Objectif fisheye Nikon-FC-E8

(b) Image par un objectif « *fi-sheye* »

FIG. 1.8 – Exemple de capteurs à objectif grand angle.

des miroirs de révolution de forme hyperboloïde et ellipsoïde. C'est dans les années 1990 que la vision omnidirectionnelle a connu son véritable essor. Le développement des activités de robotique autonome ont poussé les scientifiques à s'intéresser à cette technologie de capteurs [168], [169], [26], [48], [1], [23], [18], [97]. En effet, l'acquisition immédiate d'une image à 360° d'une scène 3D, permet à ces systèmes de répondre favorablement à des problématiques temps-réel. Les fondements théoriques, notamment la définition formelle de la **contrainte de point de vue unique**, sont posés par Nayar [111] et repris dans l'article de Baker & Nayar [6].



(a) Capteur catadioptrique



(b) Image par un capteur catadioptrique

FIG. 1.9 – Exemple de capteurs catadioptriques.

**Contrainte de point de vue unique :** La **contrainte de point de vue unique** correspond au fait qu'un point 2D dans l'image ne mesure l'intensité que d'un unique rayon lumineux dans l'espace 3D, l'ensemble des rayons s'intersectant en un point unique : le centre de projection. Les capteurs respectant cette contrainte sont dits à **projection centrale** ou **capteurs centraux**. Par opposition, ceux possédant plusieurs centres de projection sont dits à **projection non-centrale** ou **capteurs non-centraux**. L'exemple classique d'un capteur central est le système d'acquisition perspectif dont la transformation entre l'espace 3D et le plan image est décrite par le modèle « *pinhole* » (cf. figure 1.10a). Dans le cas des capteurs catadioptriques, la transformation entre l'espace 3D et le plan image ne peut être décrite uniquement avec le modèle « *pinhole* », les distorsions introduites par le miroir étant importantes. Néanmoins, pour certaines configurations miroir-caméra, le capteur catadioptrique peut être considéré comme central. La contrainte de point de vue unique est alors définie comme la condition que l'ensemble des rayons provenant du point focal (« *effective pinhole* »), intersectant le miroir, semble converger en un point : le point de vue unique (« *effective viewpoint* ») également le foyer du miroir [6] (cf. figure 1.10b). L'intérêt de la contrainte de point de vue unique est essentiel parce qu'elle permet l'utilisation des résultats établis en vision classique, ce qui conduit à une simplification des modèles de projection. L'image acquise peut être rétro-projetée sur tout type de surfaces (projection cylindrique, sphérique, etc.) et la géométrie épipolaire, qui permet la recherche de points homologues pour des systèmes stéréoscopiques, est facilitée.

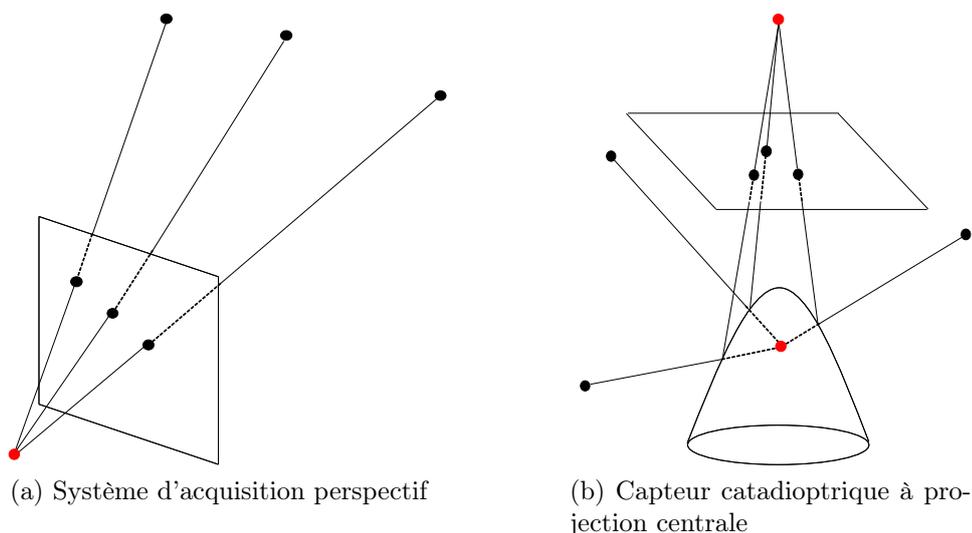


FIG. 1.10 – Deux illustrations de la contrainte de point de vue unique [136].

**Capteurs centraux :**

- le **miroir plan et une caméra perspective** (cf. figure 1.11a) : le point de vue unique est situé sur la bissectrice perpendiculaire à la droite reliant le point focal et son image virtuelle. Son utilisation reste marginale puisqu'il ne permet pas d'augmenter le champ d'observation. On peut néanmoins citer les travaux de Nalwa [110] qui utilise quatre miroirs plans et quatre caméras. Les miroirs forment une pyramide à base carrée dont les quatre centres de projection sont confondus en un centre de projection unique ;
- le **miroir ellipsoïde et une caméra perspective** (cf. figure 1.11b) : le point de vue unique et le point focal sont localisés aux deux foyers de l'ellipsoïde de révolution. Le miroir étant de courbure concave, le champ d'observation n'est que peu augmenté. Ce miroir n'offre pas d'intérêt particulier et, il n'existe pas, à notre connaissance, de travaux utilisant cette configuration matérielle ;
- le **miroir sphérique et une caméra perspective** (cf. figure 1.11c) : le point de vue unique est confondu avec le point focal et situé au centre de la sphère. Ce miroir n'a que peu d'intérêt dans cette configuration car la caméra s'observe elle-même. Néanmoins, certains s'affranchissent de cette contrainte et utilisent le miroir sphérique pour l'observation à 360° d'une scène [166], [79] ;
- le **miroir conique et une caméra perspective** (cf. figure 1.11d) : le point de vue unique est également confondu avec le point focal et est situé au sommet du cône. A l'instar du miroir sphérique, le miroir conique n'offre que peu d'intérêt dans cette configuration. Il reste néanmoins un miroir très utilisé pour les applications de robotique autonome puisque les droites verticales dans l'espace 3D sont projetées en droites radiales sur le plan image du capteur. La détection d'amers<sup>8</sup> pour des applications en navigation autonome est donc grandement simplifiée [22], [25], [147] ;
- le **miroir paraboloidé et un objectif télécentrique** (cf. figure 1.12b) : le point de vue unique est le foyer d'un paraboloidé de révolution. Le paraboloidé induisant une projection orthographique des rayons lumineux, le point focal est rejeté à l'infini. Ce type de miroir est très utilisé puisque la projection orthographique ne contraint pas de positionnement précis du miroir vis à vis de la caméra [61]. En d'autres termes, la formation géométrique des images est invariante en translation [111]. Afin de modéliser la projection orthographique, des lentilles télécentriques sont utilisées. En réalité, les lentilles télécentriques fixées sur les caméras perspectives possèdent un centre de projection à une distance lointaine mais finie du miroir. Les rayons lumineux ne sont donc pas idéalement parallèles à l'axe optique et le capteur n'est pas idéalement cen-

---

<sup>8</sup> « Un amer, terme issu du vocabulaire maritime, est défini comme un objet, bâtiment fixe et visible situé sur une côte et servant de point de repère pour la navigation, d'où le mot plus explicite de « *landmark* » en anglais » [121].

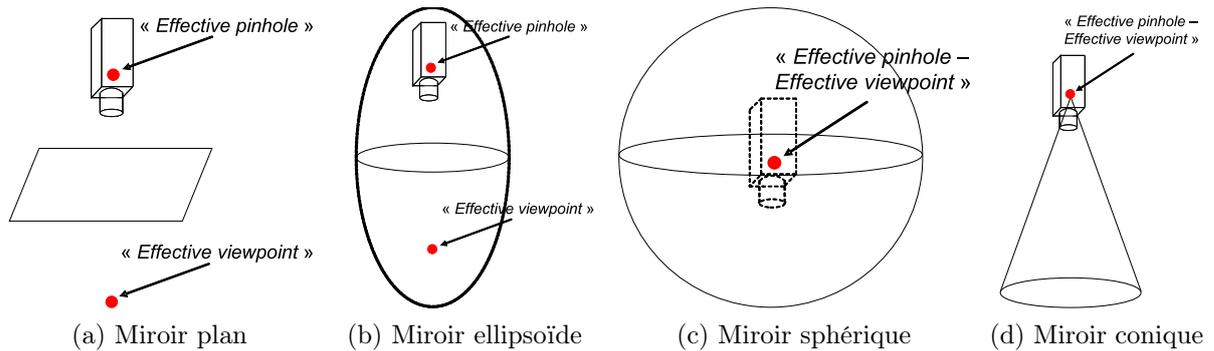


FIG. 1.11 – Exemples de configurations matérielles pour des capteurs catadioptriques centraux. Les deux configurations en figures a) et b) n’augmentent pas le champ d’observation et sont donc peu utilisées. Les deux configurations en figures c) et d) sont dégénérées puisque le point focal (« *effective pinhole* ») et le point de vue unique sont confondus [111], [6].

tral [97]. Néanmoins, l’approximation par un modèle central se vérifie parfaitement pour la majorité des applications où le capteur est employé ;

→ le **miroir hyperboloïde et une caméra perspective** (cf. figure 1.12a) : il s’agit d’un capteur hypercatadioptrique pour lequel le point de vue unique et le point focal sont situés aux deux foyers d’un hyperboloïde à deux nappes de révolution. L’une des nappes est le miroir. L’autre n’est pas utilisée et correspond à l’emplacement de la caméra. Contrairement au miroir parabololoïde utilisant un objectif télécentrique, l’hyperboloïde se couple avec une caméra perspective classique, ce qui rend le capteur compact et financièrement plus accessible. Les travaux utilisant ce type de miroirs sont donc nombreux, particulièrement ceux dédiés à la robotique autonome [154], [32], [99].

**Capteurs non-centraux :** La contrainte de point de vue unique est d’un intérêt essentiel, d’une part parce qu’elle facilite l’étalonnage du capteur en proposant un modèle de projection simplifié et d’autre part parce qu’elle facilite la recherche de points homologues pour des capteurs stéréoscopiques. Néanmoins sa mise en oeuvre reste souvent délicate et ce, pour deux raisons essentielles :

→ la contrainte de point de vue unique nécessite un positionnement idéal entre le miroir et la caméra. L’axe optique du capteur doit être parfaitement confondu avec l’axe de révolution du miroir et le capteur doit être positionné à une distance très précise du miroir (cas du capteur hypercatadioptrique). Ces contraintes de positionnement et d’ajustement sont donc difficiles à réaliser mécaniquement, ce qui ne permet pas de pouvoir considérer le capteur comme parfaitement central [45] ;

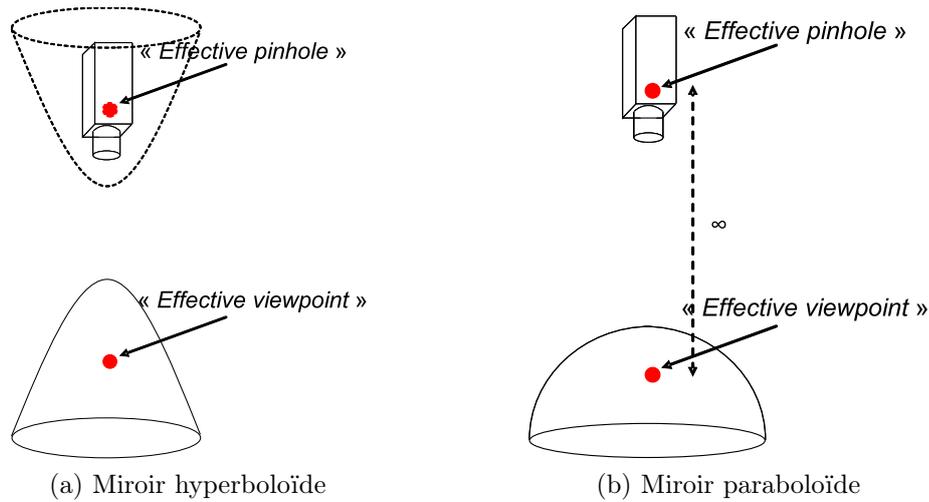


FIG. 1.12 – Autres exemples de configurations matérielles pour des capteurs catadioptriques centraux. a) un capteur hypercatadioptrique : une caméra perspective observant un miroir hyperboloïde à deux nappes de révolution. b) Un capteur paracatadioptrique : une caméra avec un objectif télécentrique observant un miroir parabololoïde [111], [6].

→ certaines applications nécessitent des propriétés spécifiques en terme de résolution qui sont, par définition, incompatibles avec des miroirs à projection centrale [68], [85]. Cette non-centralité conduit nécessairement à devoir prendre en compte les **courbes** (ou **surfaces**) **caustiques** (cf. figure 1.13). Il s'agit de l'enveloppe des rayons lumineux subissant une réflexion (**catacaustique**) ou une réfraction (**diacaustique**). Le cas des capteurs centraux est en fait un cas particulier pour lequel les courbes caustiques sont réduites en un point unique [72]. La détermination de la caustique est donc d'un intérêt important puisque cette surface décrit de manière complète le miroir. Ainsi, pour des problématiques d'étalonnage par exemple, déterminer la courbe caustique revient à étalonner le capteur puisqu'elle permet d'associer un point de l'espace 3D à un point 2D et un point 2D à une ligne de vue dans l'espace 3D.

### 1.3.3 Synthèse

Nous avons présenté dans les paragraphes précédents la technologie d'acquisition de données 3D par capteurs passifs. Il s'agit de technologies d'acquisition par vision qui s'avèrent pertinentes pour la numérisation 3D d'environnements parce qu'elles fournissent une quantité d'informations importante instantanément. Nous avons tout d'abord présenté les capteurs de vision directionnelle qui permettent l'observation d'une zone limitée de l'espace 3D, et nous avons introduit ensuite les capteurs de vision omnidirectionnelle dont le

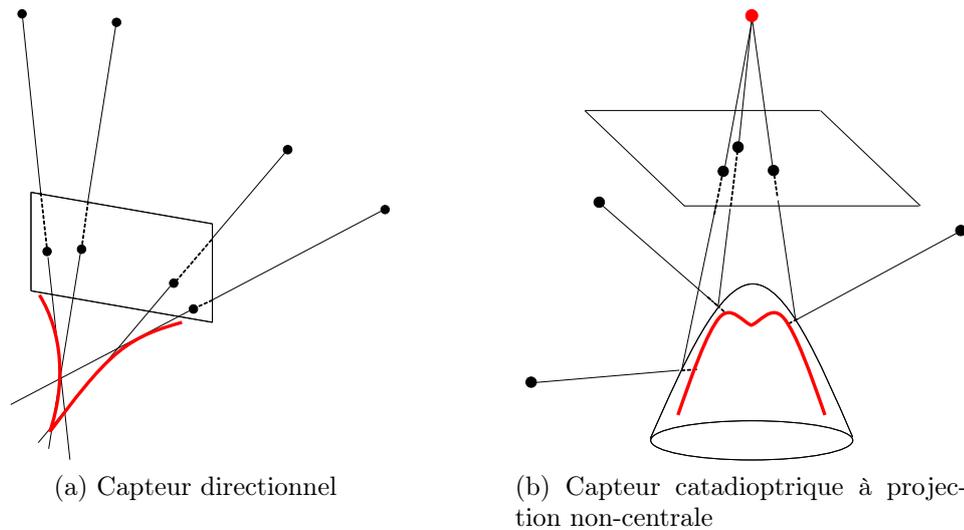


FIG. 1.13 – Deux illustrations de courbes caustiques pour des capteurs non-centraux [136].

champ de vue est augmenté significativement pour atteindre  $360^\circ$  en azimuth. Différentes technologies permettant d'accroître le champ d'observation ont alors été distinguées : les capteurs cylindriques ou sphériques, les réseaux de caméras, les objectifs grand angle et les capteurs catadioptriques :

- la première catégorie, à savoir les capteurs cylindriques ou sphériques utilisent une (ou deux) caméra linéaire ou matricielle directionnelle fixée sur un axe en rotation. Cette technologie autorise l'obtention d'images très haute résolution (ex. Benosman & Devars [15] proposent un capteur dont l'image panoramique finale est de dimensions  $1024 \times 3600$  pixels). Ces technologies nécessitent cependant une mécanique et une algorithmique (dans le cas du mosaïcage) qui sont difficilement compatibles avec des traitements temps-réel. Ces systèmes sont donc très généralement utilisés pour des applications en environnements statiques [61] ;
- la seconde catégorie concerne les réseaux de caméras qui sont l'association de plusieurs caméras positionnées de sorte à couvrir les  $360^\circ$  en azimuth. A l'instar de la solution précédente, cette technologie permet l'obtention d'images panoramiques haute résolution. Néanmoins cette solution est rarement utilisée pour des applications temps-réel : le flux de données est important et rend difficiles les traitements. Un positionnement mécanique précis des centres optiques des caméras est également nécessaire afin d'éviter l'apparition d'une disparité dans l'image panoramique finale. L'utilisation de plusieurs caméras soulève également le délicat problème de la synchronisation des prises d'images, qui s'avère d'une importance essentielle pour des applications en environne-

ment dynamique. Lorsqu'un rendu 3D est l'objectif de l'application, l'étalonnage d'un tel capteur, étape clé qui permet de faire le lien entre la scène 3D et les plans images des capteurs est délicat : les caméras ne possédant aucun champ d'observation en commun, l'estimation des poses relatives des caméras est difficile. Enfin, la multiplication des capteurs rend le prix de ces systèmes prohibitif ;

- la troisième catégorie se rapporte aux objectifs grand angle de type « *fish-eye* ». Leur principal avantage est de pouvoir acquérir des images omnidirectionnelles temps-réel sans aucune occultation [97]. Néanmoins, ils présentent certains inconvénients qui rendent leur utilisation plus complexe. D'une part, les images sont très distordues et la résolution dans la zone d'intérêt est faible (résolution faible en périphérie et bonne au centre) [61]. D'autre part, ces objectifs ne possèdent pas de centre de projection unique ce qui rend leur modélisation et leur utilisation difficiles ;
- la dernière catégorie se réfère aux capteurs catadioptriques, association d'un miroir et d'une caméra. Cette technologie est particulièrement appréciée parce qu'elle permet l'acquisition d'images omnidirectionnelles en temps-réel. L'image obtenue est fortement distordue, mais contrairement aux objectifs grand angle, certaines configurations miroir-caméra possèdent un centre de projection unique ce qui facilite l'écriture des modèles et permet d'utiliser (de réadapter) les résultats de la Vision par Ordinateur décrite pour les systèmes directionnels (géométrie épipolaire, etc.). L'inconvénient principal de ce type de capteurs est la résolution des images acquises : elle est faible au centre et bonne en périphérie [5], [98], [19].

Les systèmes de vision et notamment les systèmes de vision omnidirectionnelle, dont le champ de vue est obtenu par observation d'un miroir, semblent être particulièrement adaptés pour la numérisation de scènes 3D en temps-réel. Néanmoins, et c'est là le principal inconvénient des capteurs de vision, le rendu 3D est étroitement lié aux caractéristiques des surfaces observées. Ainsi, l'obtention d'un rendu 3D pour des surfaces non-texturées est difficile [44].

## 1.4 Capteurs de stéréovision omnidirectionnelle

Du fait de l'aspect projectif du processus de formation d'images par un capteur vidéo, l'utilisation d'une seule image ne permet pas de déterminer la géométrie 3D d'une scène observée. La **stéréovision**, c'est-à-dire la perception du relief d'une scène grâce aux différences entre les images d'objets formées sur les plans images des capteurs permet de reconstruire la structure 3D de la scène (cf. figure 1.14). La mesure par stéréovision est également appelée **photogrammétrie** ou **vidéogrammétrie** selon qu'elle privilégie la justesse des mesures (photogrammétrie) ou qu'elle favorise la vitesse d'acquisition et des traitements des données (vidéogrammétrie). Une autre définition suggère que la photogrammétrie est utilisée lorsque les traitements sont effectués en temps différé et la vidéogrammétrie, lorsqu'ils sont effectués en temps réel [171]. Néanmoins, depuis l'apparition des appareils photo numériques pouvant faire l'acquisition de 25 images par seconde, la distinction entre photogrammétrie et vidéogrammétrie est difficile.

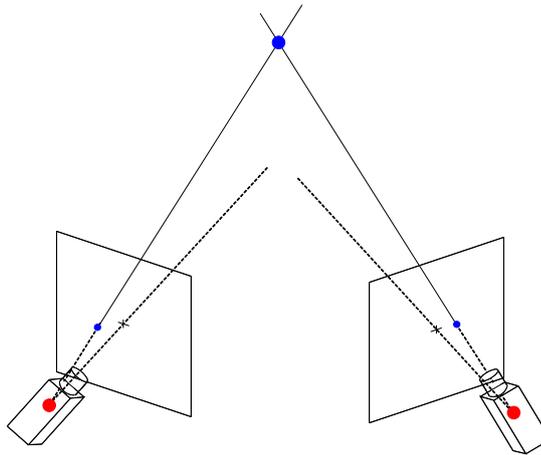


FIG. 1.14 – Détermination de la position d'un point 3D par stéréovision.

La stéréovision, c'est-à-dire les techniques mises en jeu pour acquérir les images de points de vue différents, peut être réalisée, soit par l'utilisation d'un seul capteur en déplacement, soit par l'utilisation de plusieurs capteurs. L'utilisation d'un capteur en mouvement se réfère à une problématique ayant fait l'objet de nombreux travaux et connue sous le nom de « *structure from motion* » [38]. Le « *structure from motion* » consiste à estimer conjointement les mouvements du capteur et la géométrie de la scène 3D. Pour résoudre ce type de problème, des hypothèses liées à la rigidité de la scène et à l'invariance des paramètres intrinsèques du capteur sont généralement formulées. Néanmoins, des travaux récents ont porté sur la résolution de ce type de problème pour des scènes déformables [159]. Cette

technique de reconstruction 3D par stéréoscopie monoculaire est un cas particulier et traditionnellement, lorsque l'on parle d'un système stéréoscopique, deux capteurs sont utilisés. Ayache [3] propose néanmoins un système de stéréovision trinoculaire pour lequel il soulève la question du nombre optimal de capteurs. Intuitivement, une augmentation du nombre de capteurs augmenterait la précision de la reconstruction 3D. Néanmoins, Ayache démontre qu'un nombre de capteurs trop important, au-delà de trois, diminue les performances du système stéréoscopique. Deux raisons principales peuvent être évoquées : plus le nombre de capteurs augmente, plus le champ de vision commun aux capteurs diminue, ceci restreignant la zone 3D à reconstruire. De plus, dans le cas d'un système trinoculaire, la contrainte épipolaire est réduite à un point. Des capteurs supplémentaires ne font donc que fournir une information redondante à celle existante. Comme le souligne Ayache, une configuration matérielle à trois capteurs semble être la plus pertinente. L'algorithme trinoculaire est plus simple que l'algorithme binoculaire. Sa mise en oeuvre est donc moins coûteuse et le temps d'exécution est plus court. L'algorithme trinoculaire est également plus fiable puisque la contrainte épipolaire rend l'appariement univoque. Il est aussi plus précis puisque le point 3D est à l'intersection de trois droites. Néanmoins, le champ d'observation commun est moindre que pour un système binoculaire, ce qui restreint la zone 3D à reconstruire. De plus, il faut trois capteurs et donc trois pré-traitements pour extraire une représentation de chacune des images. Le coût matériel et le temps de traitement sont donc significativement augmentés.

Traditionnellement, les techniques de reconstruction 3D partent des données 2D pour déterminer la géométrie 3D de la scène. Elles utilisent principalement l'appariement, c'est-à-dire la mise en correspondance, de primitives stéréo-correspondantes dans les différentes images acquises par le capteur, puis une triangulation des lignes de vue 3D associées aux primitives appariées pour obtenir le rendu 3D. La stéréovision peut-être dense ou éparse, selon que l'appariement s'effectue pour tous les couples stéréo-correspondants (stéréovision dense) ou uniquement pour un échantillon de primitives (stéréovision éparse). Les primitives utilisées sont généralement des points mais peuvent être des droites ou des régions [69]. Les appariements potentiels sont généralement obtenus à partir de contraintes géométriques (épipolaire ou d'orientation). Les faux appariements sont ensuite rejetés en utilisant des contraintes d'ordre, d'unicité, etc. En pratique, seules quelques contraintes sont implémentées afin d'éviter des temps de calcul prohibitifs.

Les paragraphes suivants présentent quelques travaux applicatifs ou théoriques pour la stéréovision de capteurs directionnels ou omnidirectionnels. Un rappel de la géométrie épipolaire d'un capteur stéréoscopique binoculaire standard est proposé [3], [69], [67]. Il

est suivi d'une description de la géométrie épipolaire pour deux capteurs catadioptriques en configuration générale ou coaxiale [16], [155]. Enfin, la solution matérielle retenue pour le programme *OmniViss* est présentée.

### 1.4.1 Quelques exemples de capteurs

La stéréovision de capteurs standard est un aspect très abouti de la Vision par Ordinateur, tant par les diversités des applications (aide à la conduite, construction de carte, navigation pour la robotique autonome, modélisation d'environnements etc. [163], [17], [80], [109]) que par les aspects théoriques [3], [167], [93], [174], [67].

En vision omnidirectionnelle, des capteurs stéréoscopiques catadioptriques ont également été développés. Lhuillier [88] utilise une séquence d'images acquises par un capteur monoculaire catadioptrique en mouvement pour des applications en reconstruction 3D d'environnements. Drocourt & al. [40] proposent une approche par stéréovision monoculaire catadioptrique pour la localisation de robots mobiles. Jang & al. [75] présentent un système de stéréovision monoculaire catadioptrique qui utilise deux miroirs. L'image omnidirectionnelle résultante contient en périphérie l'image d'un des miroirs et au centre l'image du second. El Mouaddib & al. [108] définissent une solution de stéréovision constituée d'une seule caméra et de plusieurs miroirs. Gluckman & al. [59], Gonzalez Barbosa [61] proposent des systèmes de stéréovision binoculaires catadioptriques pour la reconstruction 3D de scènes. Micusik [100] s'intéresse également à la vision omnidirectionnelle à partir de deux vues pour des problématiques de reconstruction de scènes 3D et de « *structure from motion* ». Ghandi & Trivedi [50] fixent un système de stéréovision omnidirectionnelle sur le toit d'un véhicule pour des applications en détection de piétons.

## 1.4.2 Géométrie épipolaire

### 1.4.2.1 Cas d'un système stéréoscopique binoculaire directionnel

On considère le système binoculaire stéréoscopique présenté sur la figure suivante (cf. figure 1.15). Soit  $\mathbf{Q}$  un point de l'espace 3D. Notons  $\mathbf{q}_1$  et  $\mathbf{q}_2$  les projetés de  $\mathbf{Q}$  sur les plans images des deux capteurs.  $\mathbf{q}_1$  et  $\mathbf{q}_2$  sont définis dans les référentiels caméras  $\mathcal{R}_{1C}$  et  $\mathcal{R}_{2C}$  d'origines respectives  $\mathbf{C}_1$  et  $\mathbf{C}_2$ . Étant donné un point  $\mathbf{q}_1$  sur le plan image  $\pi_1$ , tous les points physiques qui peuvent avoir produit  $\mathbf{q}_1$  sont sur la droite  $\mathcal{L}_1$  passant par

les points  $\mathbf{q}_1$  et  $\mathbf{Q}$ . Par conséquent, tous les homologues possibles de  $\mathbf{q}_1$  sur le plan image  $\pi_2$  sont situés sur la projection de la droite  $\mathcal{L}_1$  sur le plan image  $\pi_2$ . Cette projection est une droite infinie  $l_2$  passant par  $\mathbf{e}_2$ , qui est l'intersection de la droite passant par les foyers des caméras  $\mathbf{C}_1$  et  $\mathbf{C}_2$  avec le plan  $\pi_2$ .  $\mathbf{e}_2$  est appelé l'**épipôle** du deuxième capteur par rapport au premier, et  $l_2$  est appelée la **droite épipolaire** du point  $\mathbf{q}_1$  dans le plan image  $\pi_2$  du second capteur. L'épipôle est le point d'intersection de la droite joignant les centres des caméras avec le plan image. Un **plan épipolaire**  $\Pi$  est un plan contenant la droite reliant les centres des caméras et un point de l'espace 3D. Une ligne épipolaire est l'intersection du plan épipolaire avec le plan image. Toutes les droites épipolaires se coupent à l'épipôle.

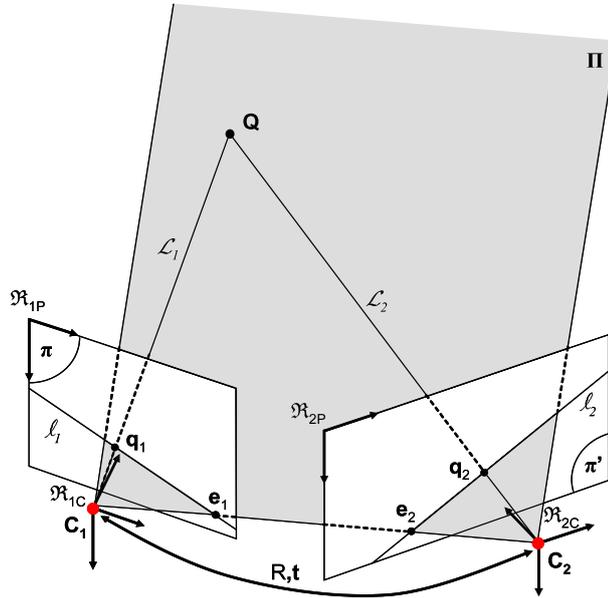


FIG. 1.15 – Géométrie épipolaire pour des capteurs perspectifs [21].

Les points  $\mathbf{Q}$ ,  $\mathbf{C}_1$ ,  $\mathbf{C}_2$ ,  $\mathbf{q}_1$  et  $\mathbf{q}_2$  sont coplanaires. Soit  $\mathbf{t}$  le vecteur de translation entre les points  $\mathbf{C}_1$  et  $\mathbf{C}_2$ . La normale du plan épipolaire peut être exprimée dans le repère associé au premier capteur par :  $\mathbf{n}_1 = \mathbf{t} \wedge \mathbf{q}_1$ . Cette normale peut aussi être exprimée dans le repère associé au second capteur par  $\mathbf{n}_2 = \mathbf{R}\mathbf{n}_1 = \mathbf{R}(\mathbf{t} \wedge \mathbf{q}_1)$ . Le point  $\mathbf{q}_2$  appartiendra au plan épipolaire (plan défini par la normale  $\mathbf{n}_2$  dans le second repère) si :  $\mathbf{q}_2^\top \mathbf{n}_2 = 0$ . La contrainte de coplanarité peut donc être exprimée par :

$$\mathbf{q}_2^\top \mathbf{R}(\mathbf{t} \wedge \mathbf{q}_1) = 0 \quad (1.8)$$

Cette contrainte de coplanarité peut être réécrite sous forme matricielle par :

$$\mathbf{q}_2^\top \mathbf{E} \mathbf{q}_1 = 0 \quad (1.9)$$

La matrice  $\mathbf{E}$  est appelée la **matrice essentielle**. Elle correspond au produit matriciel :  $\mathbf{E} = \mathbf{R}\mathbf{S}$ , où  $\mathbf{S}$  est une matrice antisymétrique représentant l'opérateur  $\mathbf{t} \wedge$  :

$$\mathbf{S} = \begin{pmatrix} 0 & -t_Z & t_Y \\ t_Z & 0 & -t_X \\ -t_Y & t_X & 0 \end{pmatrix} \quad (1.10)$$

Soit  $\mathbf{q}'_1$  et  $\mathbf{q}'_2$  les points 2D, respectivement  $\mathbf{q}_1$  et  $\mathbf{q}_2$ , définis dans les référentiels pixel  $\mathfrak{R}_{1P}$  et  $\mathfrak{R}_{2P}$  de chaque capteur. Lorsque les matrices des paramètres intrinsèques  $\mathbf{K}_1$  et  $\mathbf{K}_2$  sont connues, les points  $\mathbf{q}'$  et  $\mathbf{q}$  sont liés par la relation :  $\mathbf{q} = \mathbf{K}^{-1}\mathbf{q}'$ . En remplaçant  $\mathbf{q}$  dans l'équation 1.9, on obtient :

$$\mathbf{q}'_2{}^\top \mathbf{K}_2^{-\top} \mathbf{E} \mathbf{K}_1^{-1} \mathbf{q}'_1 = 0 \quad (1.11)$$

Notons  $\mathbf{F}$  le produit matriciel  $\mathbf{F} = \mathbf{K}_2^{-\top} \mathbf{E} \mathbf{K}_1^{-1}$ . Il s'agit de la **matrice fondamentale** qui caractérise la géométrie épipolaire pour des systèmes non-calibrés [94].

#### 1.4.2.2 Cas d'un système stéréoscopique binoculaire catadioptrique

On considère le dispositif binoculaire stéréoscopique omnidirectionnel constitué de deux capteurs hypercatadioptriques et présenté sur la figure suivante (cf. figure 1.16). Soit  $\mathbf{Q}$  un point de l'espace 3D et  $\mathbf{O}_1, \mathbf{O}_2$  les foyers des deux miroirs.

Notons  $\mathbf{Q}_1$  et  $\mathbf{Q}_2$  les projections de  $\mathbf{Q}$  sur chacun des miroirs. Pour cette nouvelle configuration, la contrainte de coplanarité concerne les points  $\mathbf{Q}, \mathbf{O}_1, \mathbf{O}_2, \mathbf{Q}_1$  et  $\mathbf{Q}_2$  de sorte que :

$$\mathbf{Q}_2^\top \mathbf{R}(\mathbf{t} \wedge \mathbf{Q}_1) = 0 \quad (1.12)$$

De manière identique au cas sténopé, on peut faire apparaître la matrice essentielle  $\mathbf{E} = \mathbf{R}\mathbf{S}$ , avec  $\mathbf{S}$  la matrice antisymétrique définie préalablement :

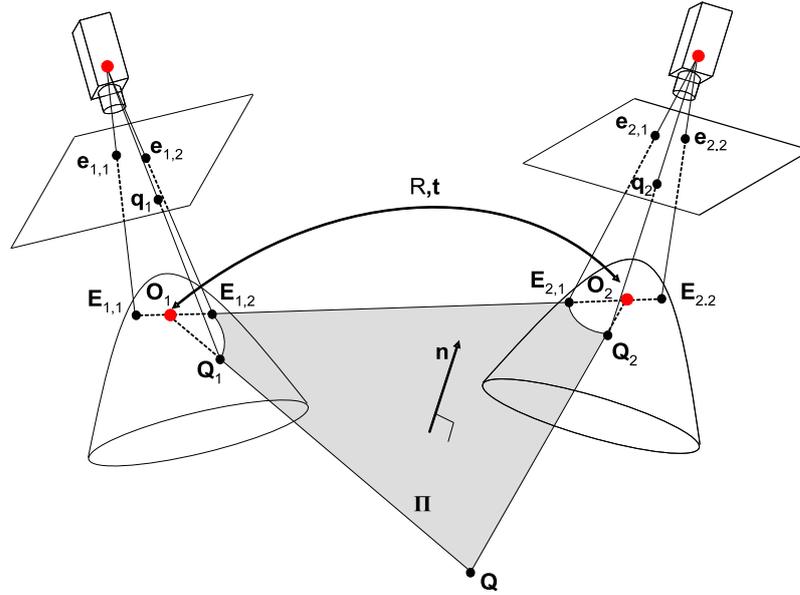


FIG. 1.16 – Géométrie épipolaire pour des capteurs catadioptriques [154].

$$\mathbf{Q}_2^\top \mathbf{E} \mathbf{Q}_1 = 0 \quad (1.13)$$

Le triplet de points  $(\mathbf{Q}\mathbf{Q}_1\mathbf{Q}_2)$  définit le plan épipolaire  $\Delta$ . Les intersections du plan et des quadriques définissent des coniques épipolaires dont les projections sur les plans images sont également des coniques. Notons que, contrairement au cas de capteurs directionnels, la droite passant par les foyers des miroirs  $\mathbf{O}_1$  et  $\mathbf{O}_2$  intersecte chaque miroir en deux points. Chaque capteur catadioptrique possède donc deux épipôles  $\mathbf{E}_{i,j}$ , l'indice  $i$  faisant référence aux capteurs catadioptriques et l'indice  $j$  aux épipôles. Par conséquent, les épipôles du premier capteur catadioptrique sont notés :  $(\mathbf{E}_{1,1}, \mathbf{E}_{1,2})$  et ceux du second sont notés :  $(\mathbf{E}_{2,1}, \mathbf{E}_{2,2})$ . Dans le cas de miroirs hyperboloïdes ou paraboïdes, les coniques peuvent être des hyperboles, des paraboles, des ellipses ou des lignes. La nature de la conique dépend principalement de la forme du miroir et de la position du point dans l'espace 3D.

Pour un point  $\mathbf{q}_1$ , la conique épipolaire associée s'écrit :

$$\mathbf{q}_2^\top \mathbf{A}_2(\mathbf{E}, \mathbf{q}_1) \mathbf{q}_2 = 0 \quad (1.14)$$

Où la matrice  $\mathbf{A}_2(\mathbf{E}, \mathbf{q}_1)$  est une fonction non linéaire de la matrice essentielle  $\mathbf{E}$ , du point  $\mathbf{q}_1$ , et des paramètres d'étalonnage [154], [61].

Il existe une configuration qui simplifie la géométrie épipolaire du banc stéréoscopique, celle pour laquelle les axes optiques des capteurs (également les axes de révolution des miroirs) sont confondus (configuration coaxiale). Les coniques épipolaires sont alors des droites radiales dans les images omnidirectionnelles (cf. figure 1.17).

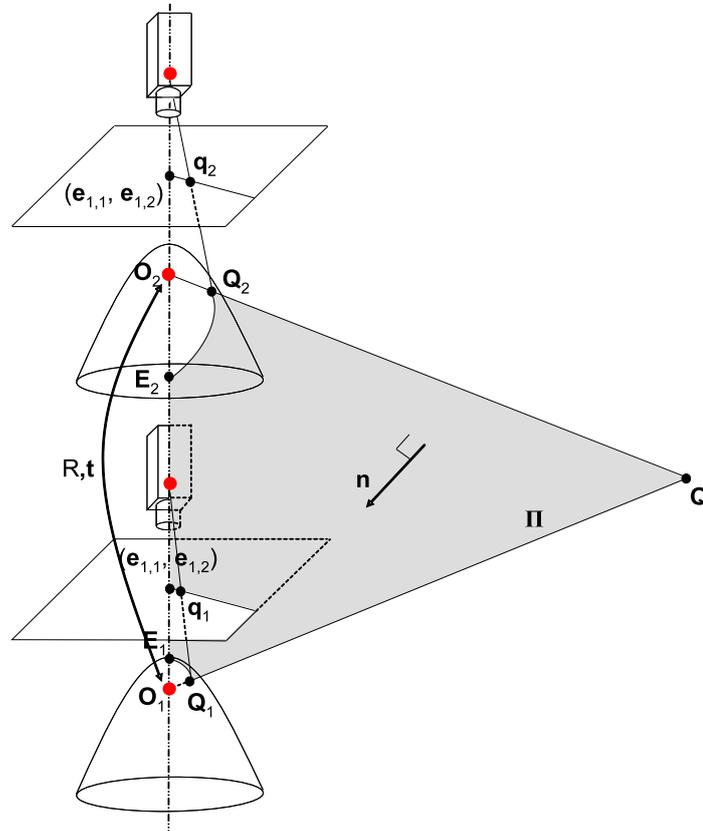


FIG. 1.17 – Géométrie épipolaire pour des capteurs catadioptriques en configuration coaxiale.

### 1.4.2.3 La rectification d'images

L'opération de **rectification** est un moyen de simplifier la géométrie épipolaire d'un capteur stéréoscopique. Si l'on considère une paire d'images d'un système stéréoscopique binoculaire, la rectification consiste en l'estimation d'une transformation de chaque plan image de sorte qu'une paire de lignes épipolaires correspondantes soit colinéaire et parallèle à l'un des axes du référentiel pixel du plan image, les épipôles étant rejetés à l'infini [49]. Le principal avantage d'une rectification est de réduire l'espace de recherche des stéréo-correspondants potentiels à une seule dimension. Ce processus est utilisé très fréquemment en stéréovision standard [4], [91], [158]. Pour un système stéréoscopique ca-

tadioptrique, l'opération de rectification existe également [58], [56], [62] et est grandement simplifiée lorsque les capteurs sont positionnés verticalement. Les droites épipolaires initialement définies suivant des rayons, sont alors des colonnes dans l'image rétro-projetée sur un cylindre de révolution<sup>9</sup>. La recherche des stéréo-correspondants est donc simplifiée puisqu'elle ne s'effectue que dans une dimension. Ce procédé n'est possible que parce que les capteurs catadioptriques sont à projection centrale, ce qui permet de rétro-projeter l'image sur tout type de surface et pour tout type de projection [61] (cf. figure 1.18). La plus couramment utilisée est la projection cylindrique. Elle permet une vision panoramique de la scène et est très facilement interprétable par l'oeil humain [74]. La désanamorphose de l'image omnidirectionnelle peut alors être envisagée de deux manières différentes :

→ une première consiste à discrétiser par une grille régulière une surface de reprojec-tion. Les points sont projetés sur l'image à l'aide des correspondances établies lors de l'étalonnage du capteur. Finalement, la surface est « dépliée » de sorte à obtenir une image panoramique de la scène. Cette méthode requiert généralement des processus d'interpolation afin de déterminer les couleurs associées aux points de la grille.

- pour la projection cylindrique (cf. figure 1.18b), il s'agit d'un cylindre de révolu-tion dont la discrétisation en hauteur et en azimuth est régulière. Les dimensions de l'image dépliée résultante sont alors fonction du nombre de points de discrétisa-tion : le nombre de points de discrétisation en hauteur définit le nombre de lignes de l'image dépliée et le nombre de points de discrétisation en azimuth définit le nombre de colonnes. Chaque point 3D du cylindre est ensuite projeté dans l'image catadiop-trique. La couleur des pixels de l'image dépliée peut être vue comme une recopie de la couleur des pixels de l'image source ou peut être obtenue par interpolation par rapport aux pixels voisins ;
- pour la projection sphérique (cf. figure 1.18c), le principe est très similaire. Une sphère est discrétisée de sorte que la résolution en azimuth et en élévation soit ré-gulière. Le nombre de points en élévation détermine le nombre de lignes de l'image dépliée et le nombre de points en azimuth définit le nombre de colonnes. Chaque point 3D de la sphère est projeté dans l'image catadioptrique et les couleurs des pixels de l'image dépliée sont, soit une recopie de celles de l'image source soit une interpolation par rapport au voisinage pixelique.

→ une seconde méthode utilise le principe « inverse ». A partir des pixels de l'image source, auxquels sont associées les lignes de vue dans l'espace 3D, on détermine l'in-tersection d'un objet 3D d'équation connue avec l'ensemble des lignes de vue. Comme la méthode précédente, l'objet 3D est ensuite « déplié ». Les dimensions de l'image

---

<sup>9</sup>Ce procédé est également connu sous le nom d'**anamorphose** ou de **désanamorphose**

---

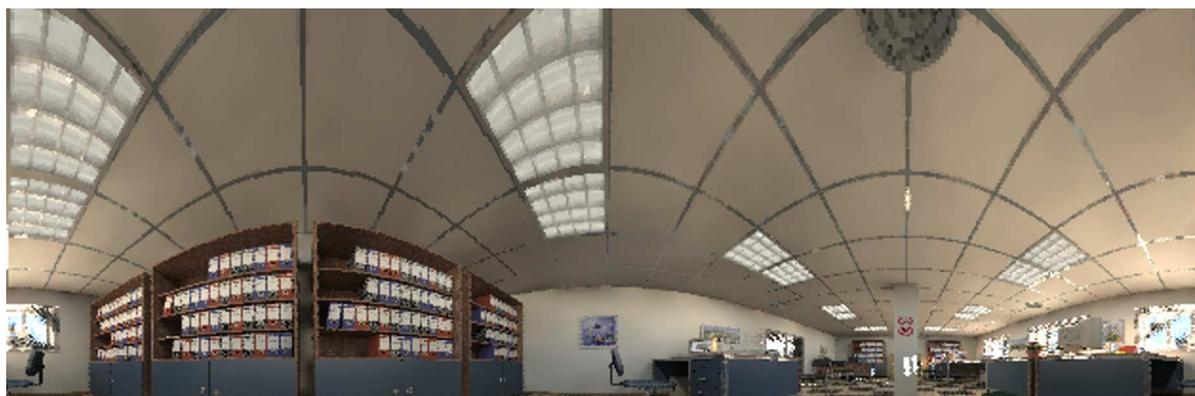
dépliée sont fixées par les paramètres de l'image source : le nombre de lignes est égal au nombre de pixels constituant le rayon maximal de l'image omnidirectionnelle et le nombre de colonnes est égal au nombre de pixels constituant son périmètre.

Cette méthode de désanamorphose ne permet pas l'obtention d'une grille de points 3D régulière. En effet, la résolution d'une image omnidirectionnelle n'est pas constante et peut être considérée comme une fonction non linéaire du rayon des pixels. La rétro-projection sur l'objet 3D de pixels appartenant à un même rayon conduit à l'apparition de « trous » dont les dimensions ne sont pas constantes.

L'intérêt de la première méthode réside principalement dans le fait que le format de reprojection est paramétrable. De plus, les mécanismes d'interpolation sont directement inclus dans l'algorithmique, ce qui évite de les faire apparaître explicitement lorsque l'image dépliée est trouée (cas de la deuxième méthode).



(a) Image omnidirectionnelle



(b) Projection cylindrique



(c) Projection sphérique

FIG. 1.18 – Projections cylindrique et sphérique d'une image omnidirectionnelle acquise par un capteur catadioptrique central. L'image omnidirectionnelle est une image de synthèse réalisée sous le logiciel de synthèse d'images **POV-Ray**.

## 1.5 Conclusions

Au travers de ce **CHAPITRE 1**, nous avons proposé une revue des systèmes permettant l'acquisition de données 3D. Cet état de l'art s'est limité aux capteurs dont les technologies mises en jeu permettent de répondre favorablement aux exigences des applications du programme *OmniViss*.

Les capteurs actifs par télémétrie et par projection d'une lumière structurée ont tout d'abord été présentés. Après un descriptif des caractéristiques techniques de ces deux technologies de capteurs, une synthèse de leurs avantages et inconvénients a permis de statuer quant à leur pertinence pour les champs applicatifs visés : le caractère actif de la mesure rend ces technologies difficilement compatibles avec des impératifs de discrétion. De plus, la mesure souvent ponctuelle nécessite une mécanique qui ne permet pas des traitements en temps-réel.

Les systèmes d'acquisition 3D par vision ont donc logiquement été introduits. Le caractère passif de la mesure répond aux exigences de discrétion que doit remplir le système de numérisation. Néanmoins, du fait d'une ouverture angulaire qui limite le champ d'observation du capteur, les systèmes de vision dans leur version simplifiée ne permettent pas la numérisation complète d'une scène 3D en temps-réel. Les systèmes de vision catadioptrique, association d'une caméra et d'un miroir permettent l'observation à 360° en azimut de l'environnement et offrent donc une solution pertinente au problème de reconstruction de scènes 3D temps-réel. Les systèmes catadioptriques sont particulièrement étudiés et utilisés pour des applications diverses car, certaines configurations miroir-caméra, appelées configurations centrales, ne possèdent qu'un centre de projection unique. Cette particularité est d'un intérêt essentiel car les principaux résultats obtenus en Vision par Ordinateur peuvent être ré-utilisés, ce qui permet de simplifier sensiblement la complexité des algorithmes de rendu 3D.

Nous avons également vu que, de part le caractère projectif du processus de formation d'images, une image ne permet pas de recréer la structure géométrique 3D de la scène observée. Deux images d'un même objet prises sous des angles de vue différents sont un minimum, formant alors un banc de stéréovision. Ce système stéréoscopique peut être monoculaire : la caméra est alors en mouvement dans la scène. Dans sa version classique, le banc stéréoscopique est binoculaire. Suite à un bref recensement des capteurs de stéréovision omnidirectionnelle que propose la littérature, nous re-détaillons la géométrie d'un capteur stéréoscopique binoculaire utilisant des caméras directionnelles. La géométrie d'un système de stéréovision omnidirectionnelle utilisant des capteurs catadioptriques

---

à projection centrale est ensuite introduite. Deux cas sont alors présentés : le premier propose une configuration générale des deux capteurs catadioptriques, c'est-à-dire que leur positionnement dans l'espace 3D n'est pas contraint. Le second détaille une configuration particulière pour laquelle les capteurs sont positionnés verticalement de sorte à avoir leurs axes optiques confondus. Cette configuration s'avère particulièrement intéressante parce qu'elle simplifie la géométrie épipolaire du banc stéréoscopique : les lignes épipolaires sont des rayons dans les images omnidirectionnelles. Nous introduisons également le processus de rectification d'images stéréoscopiques qui consiste, dans le cas standard, en une transformation de lignes épipolaires correspondantes pour qu'elles soient, après transformation, colinéaires et parallèles à l'un des axes du référentiel pixel du plan image. Dans le cas particulier d'un banc stéréoscopique coaxial, la rectification consiste à projeter l'image omnidirectionnelle sur un cylindre et de le déplier : les lignes épipolaires sont alors des colonnes dans l'image panoramique, ce qui simplifie sensiblement le processus d'appariement.

La stéréovision omnidirectionnelle accomplie par deux capteurs catadioptriques observant une même scène 3D sous des points de vue différents semble être une solution pertinente aux problématiques de reconstruction de scènes 3D. Elle répond favorablement aux exigences techniques résultant des spécificités des deux types d'applications visées par le programme *OmniViss* : un système d'acquisition de données 3D non-intrusif qui soit compatible avec une solution industrialisable et qui permette une observation large de l'environnement pour favoriser des traitements temps-réel.

# Chapitre 2

## Solution matérielle et étalonnage paramétrique

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>45</b>
<b>2.2</b>	<b>Proposition d'une solution matérielle</b>	<b>47</b>
2.2.1	Caractéristiques matérielles et techniques	47
2.2.2	Réglage du point de vue unique	49
<b>2.3</b>	<b>Étalonnage de capteurs catadioptriques</b>	<b>52</b>
2.3.1	Modèles de capteurs centraux et non-centraux	53
2.3.2	Méthodes d'étalonnage	61
2.3.3	Mesures de la précision d'un étalonnage	64
<b>2.4</b>	<b>Étalonnage paramétrique par un modèle ad hoc</b>	<b>66</b>
2.4.1	Modèles de projection et rétro-projection	67
2.4.2	Estimation des paramètres	68
<b>2.5</b>	<b>La mire d'étalonnage</b>	<b>71</b>
2.5.1	Caractéristiques techniques	71
2.5.2	Extraction des points d'intérêt	72
2.5.3	Dispositif d'aide à l'étalonnage	74
<b>2.6</b>	<b>Résultats</b>	<b>78</b>
2.6.1	Environnement de simulation et protocole d'étalonnage	78
2.6.2	Estimation des paramètres	79
<b>2.7</b>	<b>Discussion</b>	<b>82</b>
<b>2.8</b>	<b>Conclusions</b>	<b>84</b>

---

## 2.1 Introduction

La stéréovision binoculaire omnidirectionnelle utilisant deux capteurs catadioptriques est la solution matérielle retenue pour répondre aux applications du programme *OmniViss*. Lorsque l'objectif attendu est l'obtention d'un modèle 3D à partir d'un capteur de vision, l'étalonnage se révèle être une « étape incontournable [...] dès lors qu'il faut établir une relation entre l'image et le monde 3D » [83]. La littérature distingue deux types d'étalonnage : l'un concerne l'**étalonnage géométrique** d'un système de vision qui consiste à associer un point de l'espace 3D à un point 2D dans l'image (c'est-à-dire la projection), et un point 2D à une ligne de vue dans l'espace 3D (c'est-à-dire la rétro-projection) (cf. figure 2.1)<sup>10</sup>. L'autre se rapporte à l'**étalonnage radiométrique** d'un système de vision et est une problématique fondamentalement différente puisqu'elle consiste à corriger les différences photométriques des pixels [83], [157]. L'objectif principal de l'étude étant l'acquisition de données pour la reconstruction de scènes 3D, le terme étalonnage fera référence à une même et seule idée : l'étalonnage géométrique d'un système de vision.

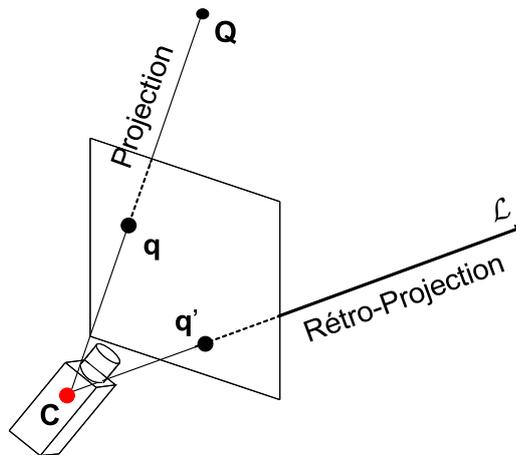


FIG. 2.1 – Principe de l'étalonnage d'un capteur de vision : projection d'un point 3D en un point 2D dans l'image et rétro-projection d'un point 2D en une ligne de vue dans l'espace 3D.

Le **CHAPITRE 2** est organisé de la manière suivante : une première partie présente et détaille la solution matérielle envisagée pour répondre aux applications du programme *OmniViss*. Une deuxième partie dresse une revue des modèles et des méthodes d'étalonnage de capteurs catadioptriques proposés par la littérature. Une troisième partie décrit

<sup>10</sup>Dans la suite du document, il ne sera pas fait de distinction entre les termes projection et rétro-projection sauf mention explicite.

---

une procédure classique d'étalonnage paramétrique utilisant une mire 3D. La mire, à savoir sa forme et la nature des points d'intérêt 3D, est présentée dans une quatrième partie. Une cinquième partie présente les résultats obtenus pour l'étalonnage paramétrique du capteur qui est suivie d'une discussion formulant une analyse critique. Finalement, une conclusion reprend les points clés soulevés et clôture ce chapitre.

## 2.2 Proposition d'une solution matérielle

Cette section détaille l'architecture matérielle retenue pour répondre aux applications du programme *OmniViss*. Une première sous-section détaille les caractéristiques techniques du capteur. Une seconde expose une procédure proposée par Zivkovic & Booij [175] pour que le système caméra-miroir respecte la contrainte de point de vue unique.

### 2.2.1 Caractéristiques matérielles et techniques

La solution matérielle retenue pour répondre aux applications du programme *OmniViss* est un banc stéréoscopique omnidirectionnel constitué de deux capteurs hypercatadioptriques positionnés verticalement (cf. figure 2.3). Chaque capteur catadioptrique est composé d'un miroir standard de forme hyperboloïdale<sup>11</sup> associé à une caméra perspective Marlin F-145C<sup>12</sup>. Le choix d'une solution matérielle de stéréovision omnidirectionnelle coaxiale s'explique principalement par le fait que la recherche de points homologues est grandement simplifiée. Cette configuration facilite également la détection de droites verticales puisque ces primitives se projettent suivant des rayons dans l'image omnidirectionnelle. Le choix d'un capteur catadioptrique à miroir de révolution hyperboloïde a été retenu parce qu'il s'agit d'un capteur à projection centrale (donc plus simple à modéliser) et que ce capteur est moins coûteux et moins encombrant qu'un capteur à miroir paraboïde utilisant une lentille télécentrique. Les différents éléments le constituant sont :

- le miroir assure un champ d'observation de 360° en azimut et de 106° théorique en élévation (cf. figure 2.3). En réalité, l'angle d'élévation est strictement inférieur puisque une portion de la caméra obstrue le champ d'observation. Dans notre cas et compte tenu des dimensions des caméras et des objectifs utilisés, l'angle d'élévation couvert par le capteur est proche de 90°. Le miroir possède un diamètre de 60mm et l'équation de l'hyperboloïde,  $\frac{z^2}{789.3274} - \frac{x^2+y^2}{548.1440} = 1$ , permet de déterminer l'écartement théorique entre le foyer du miroir et le point principal de la caméra (lieu de convergence des rayons optiques) afin de respecter la contrainte du point de vue unique : 73.143mm ;
- le capteur d'images utilisé est une caméra numérique couleur mono CCD pour laquelle le transfert des données est réalisé par une liaison firewire IEEE1394. Le capteur CCD est de dimension 1/2 pouce (largueur : 6.4mm, hauteur : 4.8mm) et les objectifs, de la marque Tamron<sup>13</sup>, sont à distance focale fixe. Afin d'assurer d'une part l'unicité du

<sup>11</sup><http://www.neovision.cz>

<sup>12</sup><http://www.imasys.fr>

<sup>13</sup><http://www.tamron.fr/>

centre de projection et de permettre à la projection du miroir de couvrir l'intégralité de l'image d'autre part, la focale doit être choisie de manière adéquate. Les figures 2.2a et 2.2b et l'équation 2.1 précisent son estimation par rapport aux caractéristiques dimensionnelles du capteur et à celles du CCD utilisé.

$$f = \frac{\delta \Delta}{\phi} \quad (2.1)$$

Où,  $\delta = 4.8\text{mm}$  désigne la hauteur du capteur CCD,  $\phi = 60\text{mm}$  désigne la largeur de l'objet observé (le diamètre du miroir) et  $\Delta = 73.143 + 9.094$  représente la distance à laquelle l'objet est situé par rapport au point principal de la caméra

La valeur de la focale obtenue est donc de  $f = 6.579\text{mm}$ . Il ne s'agit pas d'une valeur de focale standard. Le choix s'est donc orienté vers la focale de dimension standard la plus proche de la valeur théorique, soit :  $f = 6.5\text{mm}$  ;

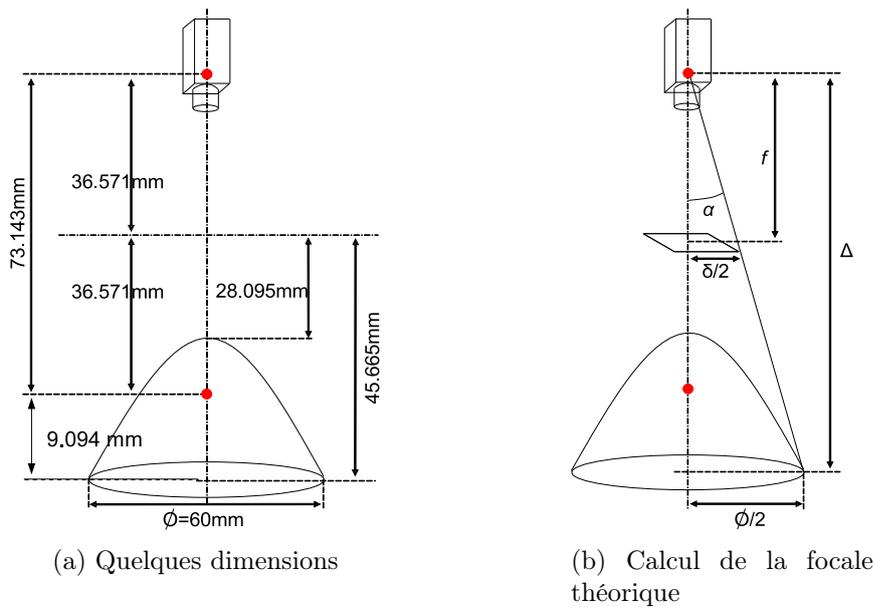
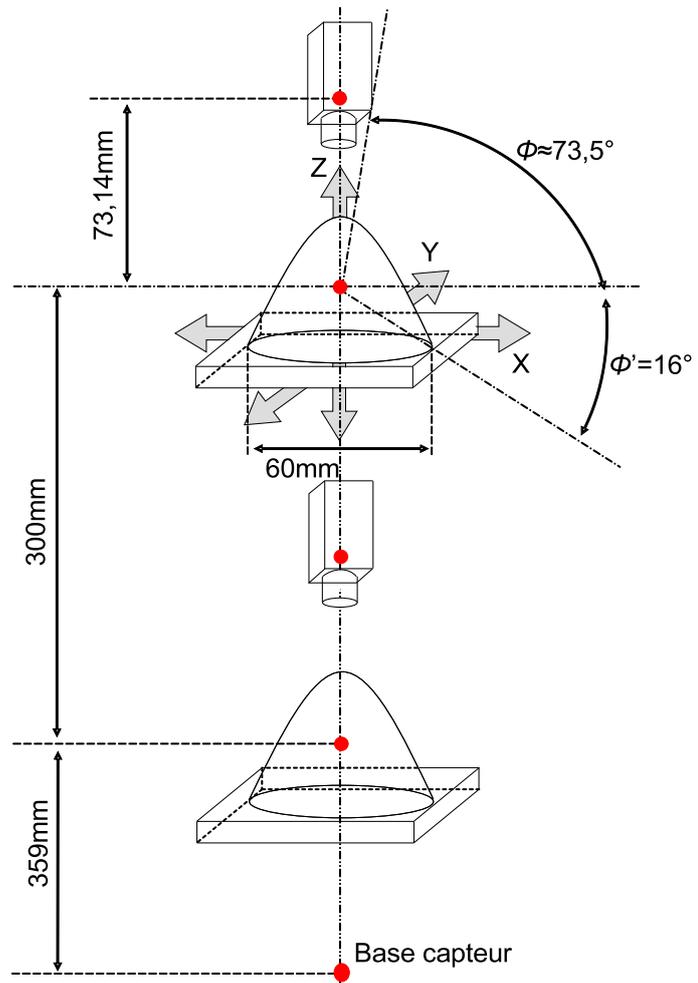


FIG. 2.2 – Détails techniques du capteur hypercatadioptrique utilisé

→ les deux capteurs sont fixés verticalement sur une potence en aluminium avec une distance inter-caméras de 300mm (cf. « *baseline* »). Le point principal de la première caméra est positionné verticalement à une distance de 359mm de la base du capteur. Les caméras sont fixées sur la potence et les miroirs sont montés sur des socles mobiles réglables suivant les trois axes à l'aide de vis micrométriques. Cette mécanique permet de positionner le miroir afin que le système (c'est-à-dire l'ensemble miroir-caméra) respecte la contrainte de point de vue unique.



(a) Photo du dispositif de stéréovision omnidirectionnelle



(b) Schéma détaillé du dispositif (les proportions ne sont pas respectées)

FIG. 2.3 – Architecture matérielle retenue pour répondre aux applications du programme *OmniViss*.

### 2.2.2 Réglage du point de vue unique

Le réglage du point de vue unique est une procédure algorithmique et mécanique qui permet de s'assurer que le capteur catadioptrique central ne possède qu'un centre de projection unique. Pour un capteur hypercatadioptrique, c'est-à-dire un capteur utilisant les propriétés d'une hyperboloïde à deux nappes de révolution, cette singularité est vérifiée pour une distance entre le foyer du miroir et le centre virtuel de la caméra (lieu de convergence des rayons lumineux et également foyer de la seconde nappe de la quadrique)

égale à  $2e$  ( $e$  est définie comme l'excentricité de la quadrique cf. paragraphe 2.3.1.1).

Lorsque ce positionnement est respecté, que la focale de l'objectif est choisie en adéquation avec le système (cf. paragraphe 2.2.1) et que les axes optique de la caméra et de révolution du miroir sont confondus, la projection du miroir dans l'image omnidirectionnelle est alors un disque qui s'inscrit dans la totalité de la hauteur du CCD et dont le centre du disque est le point principal de l'image.



FIG. 2.4 – Illustration du processus de réglage du point de vue unique pour un capteur hypercatadioptrique, d'après la méthode de Zivkovic & Booij. Le cercle en pointillé rouge correspond à la projection théorique de la bordure extérieure du miroir lorsque l'assemblage caméra-miroir respecte la contrainte de point de vue unique. L'objectif est alors de régler mécaniquement le miroir de sorte à faire coïncider la bordure extérieure du miroir réellement observée avec celle pré-estimée.

Mathématiquement ces deux points (foyer du miroir et centre virtuel de la caméra) sont parfaitement déterminés. Physiquement, il est beaucoup plus difficile de les identifier. Zivkovic & Booij [175] proposent une méthode permettant pour un système hypercatadioptrique de vérifier « au mieux » la contrainte de point de vue unique. Cette procédure permet d'optimiser la projection du miroir sur le plan image. Il s'agit à notre connaissance de la seule méthode permettant le positionnement des éléments miroir-caméra pour un tel capteur.

La caméra est tout d'abord étalonnée indépendamment du miroir. Les connaissances des paramètres intrinsèques de la caméra, de la dimension du diamètre extérieur du miroir

---

et du positionnement relatif du miroir par rapport à la caméra permettent d'estimer la projection de la bordure extérieure du miroir dans le plan image du capteur (cf. figure 2.4). La dernière étape consiste en un réglage mécanique du miroir de sorte que sa bordure extérieure coïncide avec la projection pré-estimée.

Nous supposerons pour la suite des traitements que le miroir hyperboloïde et la caméra sont positionnés de sorte à satisfaire à la contrainte de point de vue unique.

## 2.3 Etalonnage de capteurs catadioptriques

La littérature distingue trois principaux types d'étalonnage géométrique : l'étalonnage **fort**, l'étalonnage **faible** et l'**auto-étalonnage** :

- les méthodes d'étalonnage fort utilisent une mire constituée de points d'intérêt 3D dont les positions sont supposées parfaitement connues dans un référentiel attaché à la mire. Le problème consiste donc, à partir des correspondances entre les points d'intérêt 3D et leurs projections, à estimer les paramètres de la fonction de projection pour que les projections estimées correspondent au mieux aux projections réelles ;
- l'étalonnage faible consiste à étalonner deux caméras sans aucune connaissance a priori des paramètres des caméras ou sur la scène observée. Il s'agit d'estimer la structure d'un système stéréoscopique, c'est-à-dire la géométrie épipolaire du capteur. Lorsque les paramètres intrinsèques de chacun des capteurs sont connus, la géométrie épipolaire est décrite par la matrice essentielle. Lorsqu'ils ne sont pas connus, l'estimation porte sur la matrice fondamentale ;
- l'auto-étalonnage sous-entend un étalonnage automatisé des capteurs pour lequel aucune intervention manuelle n'est nécessaire ni l'utilisation d'objets spécifiques [142]. Il s'agit d'estimer les paramètres du capteur de vision en n'utilisant que l'information contenue dans les images.

De nombreux travaux ont porté sur la transposition des méthodes de étalonnage fort, faible et d'auto-étalonnage de capteurs de vision standard aux cas des capteurs catadioptriques. La littérature propose ainsi deux approches distinctes. Une première porte sur l'estimation des paramètres d'un modèle mathématique traduisant fidèlement le phénomène de projection. L'étalonnage géométrique est alors dit **paramétrique**. Une autre méthode envisage le système sous forme d'une boîte noire pour lequel aucun modèle n'est utilisé. La caméra est alors vue comme un assemblage de pixels et chaque pixel capture l'intensité lumineuse provenant d'une unique ligne de vue dans l'espace 3D [152]. L'étalonnage géométrique est alors qualifié de **non-paramétrique** :

- pour l'étalonnage paramétrique, le problème se résume à la minimisation d'une fonction coût par des méthodes d'optimisation linéaires ou itératives. La fonctionnelle à minimiser dépend des paramètres à estimer :
  - pour un étalonnage fort ou un auto-étalonnage, la fonctionnelle est choisie comme l'erreur de reprojection, très souvent définie par la somme des distances au carré entre les primitives extraites de l'image et celles estimées par le modèle [150]. Les paramètres intervenant dans le modèle se déclinent alors en deux catégories : les paramètres extrinsèques qui caractérisent la position et l'orientation de la caméra

par rapport à la scène et les paramètres intrinsèques qui caractérisent la caméra elle-même,

- pour un étalonnage faible d'un système binoculaire, le problème se résume à l'estimation de la pose relative des capteurs. Lorsque les paramètres intrinsèques sont connus, l'estimation porte sur la matrice essentielle. Un nombre de cinq appariements suffit pour effectuer l'estimation [112]. Néanmoins l'algorithme classiquement implémenté utilise huit correspondances. L'équation peut s'écrire sous la forme d'un système homogène, dont une solution est obtenue par une décomposition SVD (« *Singular Value Decomposition* »). L'algorithme des huit points est cependant très sensible aux mauvais appariements. Un algorithme de type *RanSaC*<sup>14</sup> est donc généralement implémenté afin de rendre plus robuste la procédure [67];
- l'étalonnage non-paramétrique n'utilise, par définition, aucun modèle de projection et repose sur le fait que chaque pixel du plan image mesure l'intensité lumineuse d'une unique ligne de vue dans l'espace 3D [152]. Le principe consiste donc à déterminer pour chaque pixel la ligne de vue 3D correspondante.

Cette section est un état de l'art des modèles mathématiques et des méthodes d'étalonnage pour des capteurs catadioptriques. Elle est structurée en trois sous-parties : une première propose une revue détaillée des modèles mathématiques pour des capteurs centraux et non-centraux. Une deuxième référence les méthodes d'étalonnage pour ces types de capteurs. Une troisième concerne les critères d'évaluation utilisés pour l'étalonnage de capteurs de vision.

## 2.3.1 Modèles de capteurs centraux et non-centraux

### 2.3.1.1 Modèles ad hoc pour des capteurs centraux

Les premiers modèles de transformation entre la scène 3D et le plan image ont été définis pour des configurations miroir-caméra spécifiques, assurant la centralité de la projection. Ces modèles de transformation dits **ad hoc**, utilisent les équations des miroirs et celles décrivant le processus de projection. On trouve par exemple des modèles de projection pour des capteurs hypercatadioptriques, association d'une caméra perspective et d'un miroir hyperboloïde (cf. figure 2.5a). Ce modèle s'appuie sur l'équation du miroir définie en son foyer (cf. équation 2.4) et sur la matrice de projection perspective définie pour le modèle sténopé [154] (cf. paragraphe 2.4.1.1 et annexe A).

<sup>14</sup>Random Sample Consensus : algorithme non déterministe permettant d'estimer les paramètres d'un modèle en présence de données aberrantes.

$$\frac{(z - e)^2}{\beta^2} - \frac{x^2 + y^2}{\gamma^2} = 1 \quad (2.2)$$

$\beta$  et  $\gamma$  sont les demi-axes focaux et non-focaux et  $e = \sqrt{\beta^2 + \gamma^2}$  est l'excentricité de la quadrique de révolution. Le modèle associe donc un point de l'espace 3D  $\mathbf{Q}_W = (X_W, Y_W, Z_W)^\top$  à son projeté en coordonnées homogènes  $\mathbf{q} = (u, v, 1)^\top$  dans l'image omnidirectionnelle tel que :

$$\left\{ \begin{array}{l} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K_P \frac{1}{\delta} R_C \left( \lambda R_M \left( \begin{pmatrix} X_W \\ Y_W \\ Z_W \end{pmatrix} - \mathbf{t}_M \right) - \mathbf{t}_C \right) \\ K_P = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \\ \delta = R_{C_3}^\top (\lambda R_M (\mathbf{Q}_W - \mathbf{t}_M) - \mathbf{t}_C) \end{array} \right. \quad (2.3)$$

$(R_M | -\mathbf{t}_M)$  et  $(R_C | -\mathbf{t}_C)$  définissent des changements de base,  $1/\delta$  est un coefficient de normalisation et  $K_P$  est la matrice des paramètres intrinsèques.

On trouve également des modèles de projection pour des capteurs paracatadioptriques, association d'un miroir parabolöide et d'une caméra à optique télécentrique. Ce modèle utilise l'équation d'un parabolöide centré en son foyer (cf. équation 2.4) et une matrice de projection orthographique [61] (cf. figure 2.5b et annexe B).

$$z = \frac{h^2 - (x^2 + y^2)}{2h} \quad (2.4)$$

$2h$  représente le latus rectum du parabolöide. Il s'agit d'une corde passant par le foyer de la quadrique perpendiculairement à son axe de révolution.  $h/2$  représente donc la distance entre le foyer et le sommet.

Soit  $\mathbf{Q}_M = (X_M, Y_M, Z_M)^\top$  un point de l'espace 3D défini dans un référentiel attaché au foyer du miroir et  $\mathbf{q} = (u, v, 1)^\top$  son projeté en coordonnées homogènes dans l'image omnidirectionnelle. Le modèle s'écrit alors :

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K_O O_I \begin{pmatrix} hX_M/|\mathbf{Q}_M|+Z_M \\ hX_M/|\mathbf{Q}_M|+Z_M \\ z \\ 1 \end{pmatrix} \quad (2.5)$$

$O_I$  est la matrice homogène de projection orthographique telle que définit en l'équation 1.3 et  $K_O$  est la matrice homogène des paramètres intrinsèques pour cette projection.

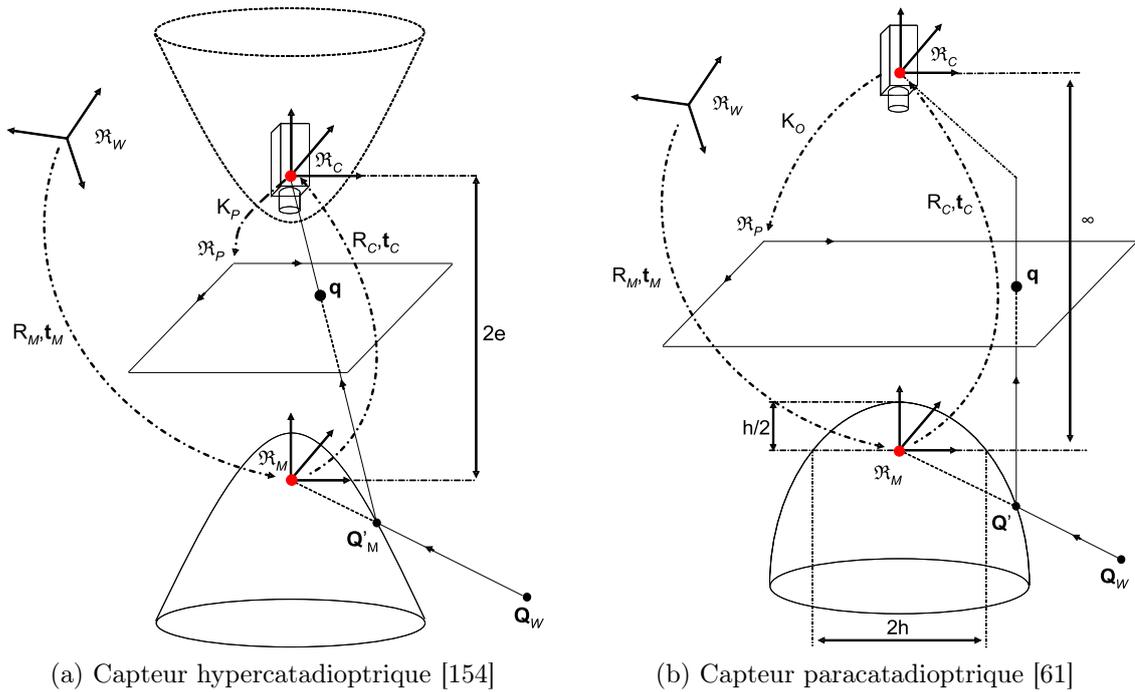


FIG. 2.5 – Détails de la projection d'un point 3D pour des modèles ad hoc de transformation entre l'espace 3D et le plan image. Les capteurs catadioptriques illustrés en figures a) et b) sont à projection centrale.

### 2.3.1.2 Modèles unifiés pour des capteurs centraux

La multiplication des configurations matérielles et la diversité des modèles ont motivé la communauté scientifique à s'intéresser à des théories unificatrices permettant de modéliser différents types de capteurs par une même approche. Ces théories unificatrices sont qualifiées de **modèles unifiés** et deux approches sont à distinguer : celles qui utilisent l'équivalence entre la projection sur une quadrique et la projection sur une sphère et celles qui approximent la projection par un polynôme.

**Sphère d'équivalence :** deux équipes ont publié la même année des travaux sur des modèles unifiés de projection [52], [9]. Ces théories reposent principalement sur l'équivalence établie par Geyer [52] entre la projection sur une surface quadrique et la projection sur une sphère. La sphère  $S$  de rayon unitaire, appelée sphère d'équivalence, permet de modéliser tous les types de capteurs à projection centrale. Ainsi, la projection d'un point 3D  $\mathbf{Q}$  défini dans un référentiel associé à la sphère  $\mathfrak{R}_S$  se décompose essentiellement en deux projections centrales :

→ une première projette centralement le point 3D  $\mathbf{Q} = (X, Y, Z)^\top$  sur la sphère  $S$  depuis le centre  $\mathbf{O}$  de la sphère en  $\mathbf{Q}' = (X', Y', Z')^\top$  ;

$$\mathbf{Q}' = \frac{\mathbf{Q}}{\|\mathbf{Q}\|} \quad (2.6)$$

→  $\mathbf{Q}'$  est exprimé dans un référentiel  $\mathfrak{R}'_S$  dont le point principal  $\mathbf{O}'$  est situé entre les deux pôles de la sphère et situé à une distance  $\xi$  (cf. tableau 2.1) du point  $\mathbf{O}$ . Cette modification du centre de projection permet de différencier les capteurs catadioptriques (hypercatadioptriques, paracatadioptriques). Ainsi, pour un capteur paracatadioptrique le centre  $\mathbf{O}'$  de la seconde projection centrale est défini au pôle de la sphère ; pour une caméra « *pinhole* » il est situé au centre ; il est localisé entre les deux pour un capteur hypercatadioptrique ;

→ une seconde projette  $\mathbf{Q}'$  sur un plan image normalisé  $\pi_N$  depuis le point principal  $\mathbf{O}'$ . Soit  $\mathbf{q}_N$  ce point 2D appartenant à  $\pi_N$ .

$$\mathbf{q}_N = \left( \frac{X'}{Z' + \xi}, \frac{Y'}{Z' + \xi}, 1 \right)^\top \quad (2.7)$$

→ finalement, la projection de  $\mathbf{q}_N$  sur le plan image de la caméra en le point  $\mathbf{q}_P$  fait intervenir la matrice des paramètres intrinsèques et une matrice dépendant des paramètres du miroir :

$$\mathbf{q}_P = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & kv & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \varphi - \xi & 0 & 0 \\ 0 & \xi - \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X'/Z'+\xi \\ Y'/Z'+\xi \\ 1 \end{pmatrix} \quad (2.8)$$

Ces modèles offrent un cadre général au problème de transformation entre l'espace 3D et le plan image d'un capteur vidéo mais souffrent de quelques imperfections : les distorsions ne sont pas prises en compte ce qui peut conduire à quelques approximations dans l'estimation du processus de projection [97]. De plus, les paramètres intrinsèques du capteur et du miroir sont dissociés ce qui amène à considérer le capteur catadioptrique comme l'association de deux entités distinctes : un miroir et une caméra. D'un point de

vue mathématique, cette décomposition peut conduire à des difficultés de convergence des algorithmes d'optimisation [160].

Récemment, les modèles initialement développés ont été repris pour être complétés et modifiés [97] (cf. figure 2.6). Tout d'abord, les distorsions radiales et tangentielles sont introduites dans le modèle sous forme d'une approximation polynomiale (cf. équation 1.6, équation 1.7). Soit  $\mathbf{q}_D$  le point 2D lorsque les distorsions sont prises en compte.  $\mathbf{q}_D$  est alors défini par :

$$\mathbf{q}_D = \mathbf{q}_N + \Delta_R(\mathbf{q}_N) + \Delta_T(\mathbf{q}_N) \quad (2.9)$$

Ensuite, les paramètres de la caméra et ceux du miroir sont exprimés en un paramètre unique : la focale généralisée  $\gamma$  (cf. tableau 2.1), ce qui permet de considérer le capteur catadioptrique comme un ensemble unique. Mathématiquement, le nombre de paramètres diminue sensiblement ce qui limite les minima locaux lors de l'étape d'estimation. Le point 2D  $\mathbf{q}_P$  sur le plan image est alors défini par :

$$\mathbf{q}_P = \begin{pmatrix} \gamma & 0 & u_0 \\ 0 & \gamma & v_0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{q}_D \quad (2.10)$$

	$\xi$	$\varphi$	$\gamma$
paraboloïde	1	$1 + h$	$-hf$
hyperboloïde	$\frac{df}{\sqrt{d^2+h^2}}$	$\frac{d+h}{\sqrt{d^2+h^2}}$	$\frac{-hf}{\sqrt{d^2+h^2}}$
ellipsoïde	$\frac{df}{\sqrt{d^2+h^2}}$	$\frac{d-h}{\sqrt{d^2+h^2}}$	$\frac{hf}{\sqrt{d^2+h^2}}$
plan	0	1	$-f$
perspectif	0	1	$f$
$d$ : distance entre les points focaux $h$ : demi latus rectum			

TAB. 2.1 – Paramètres caractéristiques pour le modèle de la sphère d'équivalence [9], [97]

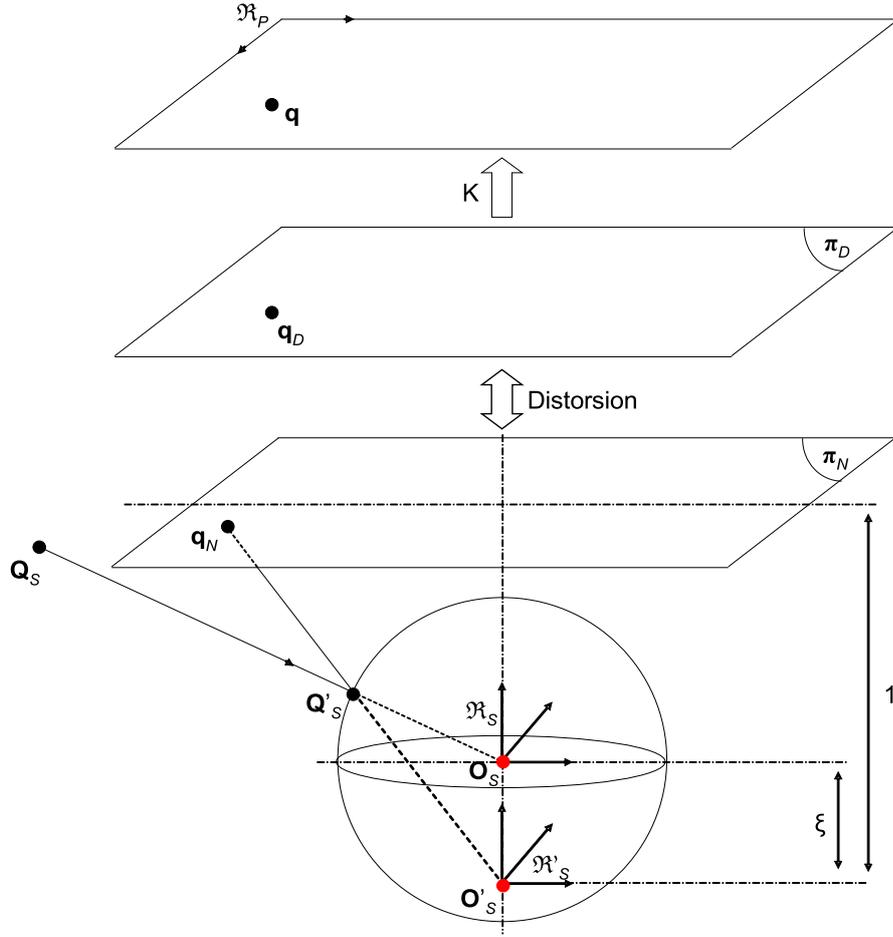


FIG. 2.6 – Détails de la projection d'un point 3D par le modèle unifié de la sphère d'équivalence [97].

**Approximation polynomiale :** Le modèle unifié proposé par Scaramuzza & al. [137] repose sur une approximation de la fonction de projection par une forme polynomiale. Il s'agit d'un prolongement des travaux initiés par Micusik & Pajdla [101]. La fonction de projection est approximée par un développement en série de Taylor au rang  $n$ . Une minimisation non linéaire permet d'estimer le degré du polynôme et ses coefficients. Soit  $\mathbf{Q} = (X, Y, Z)^\top$  un point de l'espace 3D,  $\mathbf{q}_I = (x_I, y_I)^\top$  le projeté de  $\mathbf{Q}$  défini dans un 2D dans un référentiel image  $\mathfrak{R}_I$  et  $\mathbf{q} = (u, v)^\top$  ce même point 2D dans le référentiel pixel  $\mathfrak{R}_P$ . La fonction de projection  $g$  établissant la relation entre  $\mathbf{Q}$  et  $\mathbf{q}_I$  s'écrit donc :

$$g(\|\mathbf{q}_I\|) = a_0 + a_1 \|\mathbf{q}_I\| + a_2 \|\mathbf{q}_I\|^2 + \dots + a_n \|\mathbf{q}_I\|^n \quad (2.11)$$

Les  $(a_0, a_1, a_2, \dots, a_n)^\top$  ainsi que le degré  $n$  du polynôme sont les paramètres estimés durant la procédure d'étalonnage [137]. En pratique, seuls les coefficients  $(a_0, a_1, a_2, \dots, a_n)^\top$  sont estimés. Le degré  $n$  du polynôme est généralement fixé à 4. Comme le précise l'auteur<sup>15</sup>, il s'agit d'une valeur obtenue empiriquement qui peut être augmentée ou diminuée en fonction de la pertinence des résultats.

La transformation entre  $\mathbf{q}_I$  et  $\mathbf{q}$  est décrite par :

$$\mathbf{q} = \begin{pmatrix} k_u & 0 \\ 0 & k_v \end{pmatrix} \mathbf{q}_I + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \quad (2.12)$$

En écriture homogène, il s'agit de la matrice  $K'$  introduite en l'équation A.5 qui contient les paramètres intrinsèques au processus de formation d'images  $((k_u, k_v)^\top$  : facteurs d'échelle horizontal et vertical.  $\mathbf{c} = (u_0, v_0)^\top$  : les coordonnées du centre de l'image dans  $\mathfrak{R}_P$ ).

### 2.3.1.3 Modèles pour des capteurs non-centraux

Une première approche pour modéliser les capteurs non-centraux repose sur l'estimation de la surface **caustique**. Cette surface est l'ensemble des points de convergence des rayons optiques. Elle permet d'associer un point 3D à un point 2D dans l'image et un point 2D à une ligne de vue dans l'espace 3D. Ainsi, estimer la surface caustique du miroir équivaut à étalonner le capteur. Le calcul de cette surface a donc fait l'objet de nombreux travaux pour lesquels différentes méthodologies de détermination ont été exposées. Bakstein & Pajdla [7] et Ieng & Benosman [72] proposent une revue des capteurs non-centraux, des surfaces caustiques associées et des méthodes d'estimation. Swaminathan & al. [156] proposent une méthode d'estimation basée sur la résolution d'équations différentielles, la courbe caustique étant décrite par une surface paramétrique. Ieng & Benosman [72] proposent une méthode générique reposant sur une construction géométrique. Cette méthode ne nécessite aucune connaissance de l'équation du miroir et peut être appliquée quelle que soit la surface considérée. D'autres travaux prennent en compte certaines hypothèses simplificatrices : par exemple lorsque l'axe optique de la caméra et l'axe de révolution du miroir sont confondus. Il s'agit des capteurs axiaux introduits par Ramalingam & al. [130] pour lesquels la surface caustique se réduit à une droite confondue avec l'axe du capteur.

Une seconde approche consiste en l'utilisation conjointement de modèles centraux et non-centraux [60], [88], [103]. Un modèle central est tout d'abord utilisé pour obtenir une

<sup>15</sup>[http://asl.epfl.ch/~scaramuz/research/Davide\\_Scaramuzza.htm](http://asl.epfl.ch/~scaramuz/research/Davide_Scaramuzza.htm)

première approximation du projeté d'un point 3D. Ensuite, la ligne de vue 3D correspondante est déterminée par une méthode de lancer de rayon : un premier rayon dont l'origine est le centre de la caméra et dont la direction est donnée par le point 2D intersecte le miroir au point d'incidence. La normale est ensuite calculée, permettant de connaître la direction de la ligne de vue 3D associée au point 2D. La position initiale du point 2D est alors ajustée itérativement de sorte à minimiser la distance entre le rayon de projection et la position du point 3D.

Dans le cas des capteurs non-centraux, l'estimation de la transformation entre l'espace 3D et le plan image est complexe et coûteuse en temps de calcul. Le calcul d'une surface caustique n'est pas simple et ce, malgré des hypothèses simplificatrices. L'estimation du projeté d'un point 3D par l'utilisation conjointe de modèles centraux et non-centraux est délicate et augmente sensiblement les temps de calcul.

#### 2.3.1.4 Modèles non-paramétriques

Les modèles non-paramétriques tentent de s'affranchir des contraintes inhérentes aux modèles paramétriques en considérant le capteur comme une boîte noire. Les notions de paramètres intrinsèques, extrinsèques, la focale, etc., n'existent plus, l'idée étant de pouvoir associer chaque point 3D à un point 2D dans l'image et un point 2D à une ligne de vue dans l'espace 3D. Sturm & Ramalingam [152] proposent une approche basée sur ce principe (cf. figure 2.7). Les techniques non-paramétriques ont bien souvent un caractère discret puisqu'elles mettent en correspondance, soit exhaustivement chaque pixel à une ligne de vue dans l'espace 3D, soit un sous-ensemble de pixels aux lignes de vue associées. Afin de garder le caractère continu du processus de formation d'images, c'est-à-dire que des pixels voisins aient des lignes de vue 3D voisines [151], une interpolation numérique vient souvent compléter la procédure [2].

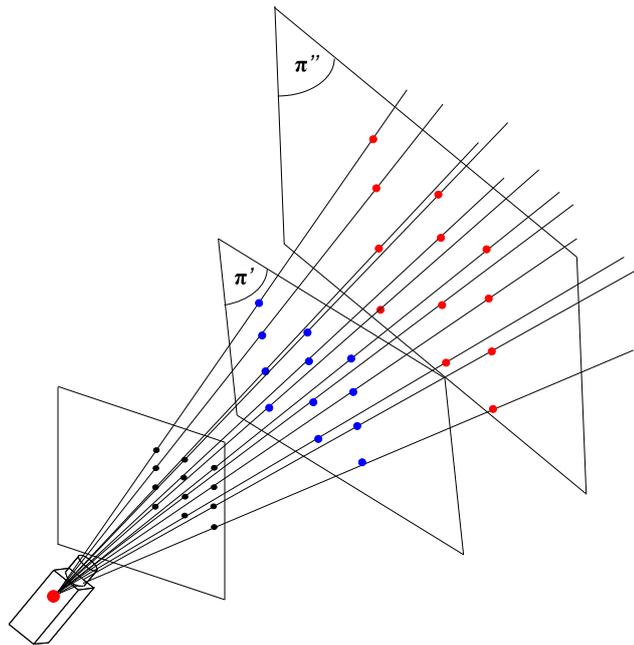


FIG. 2.7 – Illustration du modèle générique de transformation entre l'espace 3D et le plan image proposé par Sturm & al. A chaque point 2D est associée une ligne de vue dans l'espace 3D [152].

## 2.3.2 Méthodes d'étalonnage

Les paragraphes suivants présentent des méthodes d'étalonnage pour des capteurs catadioptriques. Les méthodes sont classées en trois catégories distinctes [39] suivant qu'elles utilisent une mire 3D (**étalonnage fort**), ou qu'elles estiment la géométrie d'un système de stéréovision (**étalonnage faible**) ou encore qu'elles utilisent les seules informations contenues dans la scène 3D (**auto-étalonnage**).

### 2.3.2.1 Etalonnage fort

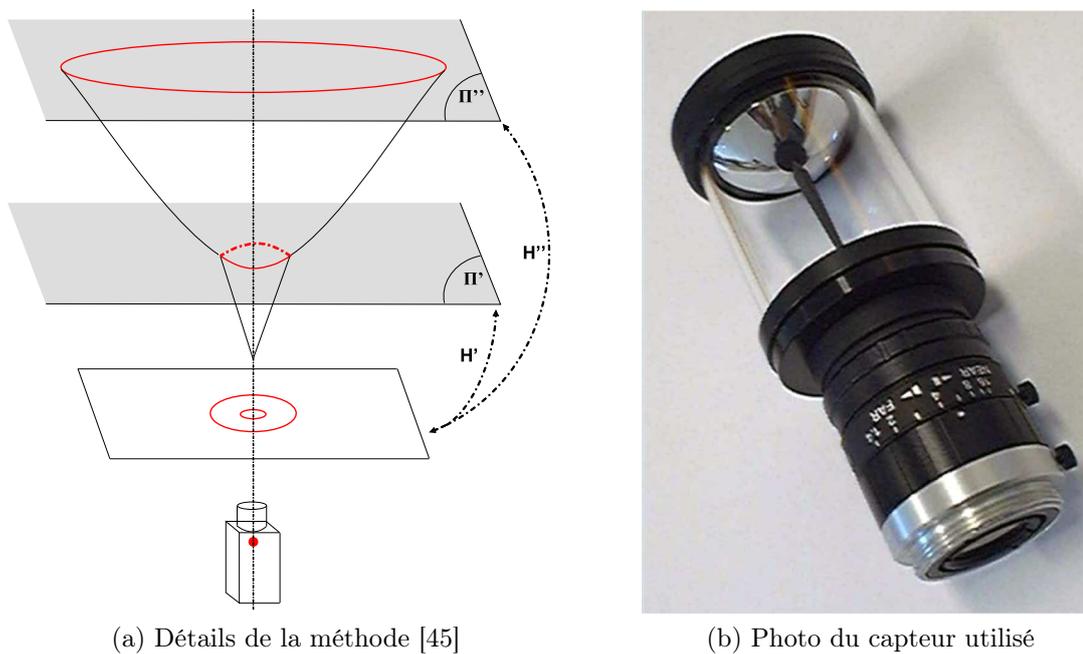
L'étalonnage fort de capteurs d'images est par essence un étalonnage supervisé qui utilise une mire 3D et s'effectue « *off line* ». La mire 3D est constituée de points d'intérêt qui peuvent être de nature différente : des points, des droites, des sphères et certains travaux utilisent l'empreinte du miroir dans l'image, considérée alors comme mire d'étalonnage. Les primitives les plus couramment employées sont des points 3D. Ils sont facilement détectables (détecteurs de Harris [65] et Shi & Tomasi [141], etc.) et la précision des motifs détectés est importante. Ils permettent également « une mesure objective de l'erreur

après étalonnage, sous forme d'une erreur de projection entre les positions observées et les positions prédites par le modèle » [160]. Les points correspondent généralement aux intersections d'un damier [26], [62], [99], d'un quadrillage [32] ou bien ils sont les barycentres de tâches lumineuses [105]. Pour un étalonnage paramétrique, l'ensemble des correspondances point 3D-point 2D est utilisé pour minimiser une fonction coût, généralement définie comme l'erreur quadratique entre les points d'intérêts extraits et des positions estimées par le modèle. Lorsque l'étalonnage est non-paramétrique, des points disposés sur une mire 3D observée dans différentes positions et couvrant l'intégralité du champ d'observation, permettent d'établir de manière exhaustive les correspondances entre les points 2D et les lignes de vue 3D associées.

Certaines méthodes utilisent des droites en 3D. Ces primitives se projettent en des portions de coniques dans l'image omnidirectionnelle fournissant trois invariants liés aux paramètres intrinsèques du capteur. L'étalonnage peut faire intervenir deux faisceaux de droites parallèles contenus dans un même plan [53], ou un échantillon de trois droites non-coplanaires [54]. Barreto & Araujo [11] et Vandeportaele & al. [161] proposent l'utilisation de droites pour l'étalonnage d'un capteur paracatadioptrique.

D'autres méthodes utilisent des sphères qui se projettent en des coniques fermées dans l'image omnidirectionnelle. Elles ne fournissent que deux informations et l'utilisation de trois sphères est donc un minimum pour étalonner le capteur [170]. Ces méthodes sont généralement plus précises que celles utilisant des droites : les droites se projetant en des portions de coniques, l'estimation peut s'avérer moins précise que dans le cas des sphères dont les coniques sont entièrement définies dans l'image omnidirectionnelle.

D'autres enfin utilisent l'empreinte du miroir dans l'image, le miroir étant lui-même considéré comme la mire d'étalonnage. Kang [76] a initialement proposé une méthode de ce type pour estimer les paramètres d'un capteur paracatadioptrique. Fabrizio & al. [45] généralisent la méthode pour tous les capteurs catadioptriques centraux à projection perspective : à la méthode initialement proposée, une technique empruntée à Gremban & al. [64] estime les transformations entre un plan de référence et deux plans d'étalonnage (le plan contenant le contour du miroir et le plan contenant le contour de l'aiguille) (cf. figure 2.8). Ces deux transformations permettent d'estimer le point principal, l'orientation du plan image ainsi que les coordonnées du centre de l'image.



(a) Détails de la méthode [45]

(b) Photo du capteur utilisé

FIG. 2.8 – Méthode d'étalonnage proposée par Fabrizio & al. utilisant l'empreinte du miroir.

### 2.3.2.2 Etalonnage faible

L'étalonnage faible concerne l'estimation de la géométrie épipolaire d'un capteur de stéréovision. Pour les capteurs centraux, la géométrie épipolaire a été décrite par Svoboda & Pajdla [155]. Mariottini & Prattichizzo [95] ont proposé une bibliothèque de fonctions intitulée « *Epipolar Geometry Toolbox* » fonctionnant sous **Matlab** pour la résolution de problèmes multi-vues<sup>16</sup>. Pajdla [117] s'intéresse à la géométrie épipolaire de certains capteurs non-centraux appelés capteurs obliques. Geyer & Daniilidis [55] décrivent les propriétés de la matrice fondamentale pour un système binoculaire paracatadioptrique et dans [57], ils définissent la géométrie épipolaire d'un capteur paracatadioptrique en mouvement. Molana & Geyer [104] généralisent la géométrie épipolaire aux capteurs non-centraux en introduisant la matrice essentielle généralisée.

<sup>16</sup>Cette bibliothèque de fonctions permet également l'élaboration de lois de commande pour l'asservissement visuel de robots autonomes lorsqu'elle est interfacée avec la bibliothèque de fonctions « *Robotics Toolbox* » développée par Corke [33].

### 2.3.2.3 Auto-étalonnage

L'auto-étalonnage de capteurs vidéo concerne toutes les méthodes de étalonnage automatique ne nécessitant aucun objet spécifique [142]. L'auto-étalonnage est par essence un étalonnage « *on line* » qui utilise les seules informations contenues dans la scène 3D. Le processus d'auto-étalonnage est mis en place lorsque le capteur est soumis à de nombreux changements ou qu'aucune mire 3D n'est disponible [149]. Il peut être monoculaire ou stéréoscopique suivant qu'il est réalisé sur un unique capteur ou sur un système de stéréovision [39]. Kang [76] propose une méthode d'auto-étalonnage pour un capteur paracatadioptrique qui utilise une séquence d'images acquises par le capteur en mouvement. La scène 3D est supposée rigide et les paramètres intrinsèques sont supposés invariants entre les prises d'images. Le problème est alors formalisé de sorte que la contrainte épipolaire pour deux images successives soit exprimée en fonction des paramètres du capteur. Une méthode d'auto-étalonnage pour un capteur paracatadioptrique est proposée par Geyer & Daniilidis [57]. Micusik & Pajdla [102] détaillent un scénario d'auto-étalonnage pour un capteur non-central qui consiste à approximer le capteur non-central par un modèle central afin d'estimer la géométrie épipolaire.

## 2.3.3 Mesures de la précision d'un étalonnage

### 2.3.3.1 Erreurs de projection et de reprojection

La mesure de la précision d'un étalonnage d'un système de vision s'effectue traditionnellement par la mesure d'une erreur géométrique dite **erreur de projection**. Il s'agit d'une mesure de distance en pixels définie comme étant la somme des carrés des distances euclidiennes entre des primitives extraites et des primitives estimées par la procédure d'étalonnage [12].

Cette mesure est parfois appelée **erreur de reprojection**. Il s'agit également d'une mesure en 2D entre des primitives extraites et estimées, les coordonnées des points 3D ayant été ré-estimées puis reprojctées dans l'image.

Considérons un système monoculaire dans lequel  $\mathbf{q}_m$  est le  $m^{\text{ème}}$  point 2D extrait de l'image et  $\hat{\mathbf{q}}_m$  la projection estimée par l'étalonnage. L'erreur de projection est alors définie par :

$$\forall m \in \{1, \dots, M\}, \quad \bar{\varepsilon} = \frac{1}{M} \sum_{m=1}^M d(\mathbf{q}_m, \hat{\mathbf{q}}_m) \quad (2.13)$$

Cette mesure d'erreur peut être facilement généralisée à un système multicaméras. Si  $\mathbf{q}_{i,j}$  est la  $j^{\text{ième}}$  primitive extraite de la  $i^{\text{ième}}$  caméra et  $\hat{\mathbf{q}}_{i,j}$  la primitive estimée correspondante l'erreur de projection  $\bar{\varepsilon}$  est alors donnée par :

$$\forall i \in \{1, \dots, I\}, \forall j \in \{1, \dots, J\}, \quad \bar{\varepsilon} = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J d(\mathbf{q}_{i,j}, \hat{\mathbf{q}}_{i,j}) \quad (2.14)$$

### 2.3.3.2 Erreurs de rétro-projection et de reconstruction

L'**erreur de rétro-projection** pour un système de vision est une mesure d'erreur en 3D. Elle est définie comme la somme des erreurs de rétro-projection de points 3D connus et de points 3D estimés, divisée par le nombre de points. Si  $\mathbf{Q}_n$  est le  $n^{\text{ème}}$  point 3D de coordonnées connues et  $\hat{\mathbf{Q}}_n$  le point 3D estimé, l'erreur de rétro-projection  $\bar{\varepsilon}$  est donnée par :

$$\forall n \in \{1, \dots, N\}, \quad \bar{\varepsilon} = \frac{1}{N} \sum_{n=1}^N d(\mathbf{Q}_n, \hat{\mathbf{Q}}_n) \quad (2.15)$$

Pour un système multicaméras, un banc de stéréovision par exemple, l'erreur de rétro-projection est très généralement définie comme l'**erreur de reconstruction** de primitives 2D appariées.

## 2.4 Etalonnage paramétrique par un modèle ad hoc

L'étalonnage de capteurs de vision concerne l'estimation des fonctions de projection  $(u, v) = g(X, Y, Z)$  et rétro-projection  $(X, Y, Z) = g^*(u, v)$  qui définissent les relations entre l'espace 3D et le plan image. Traditionnellement, ces fonctions sont décrites par des modèles paramétriques et l'étalonnage du capteur se rapporte alors à l'estimation des paramètres  $\chi$  définissant les modèles. Il s'agit donc d'un problème de minimisation pour lequel l'objectif est l'estimation d'un vecteur de paramètres optimale  $\chi^*$  minimisant une fonction coût  $J(\chi)$ .

Nous détaillons ici l'étalonnage paramétrique du banc de stéréovision binoculaire omnidirectionnelle. Les capteurs utilisés sont des capteurs hypercatadioptriques et les transformations entre l'espace 3D et le plan image sont définies par un modèle ad hoc. Les modèles utilisés sont inspirés des travaux de Svoboda [154]. Cette section est structurée de la manière suivante : les deux premières parties détaillent les modèles de projection et rétro-projection développés par Svoboda. La troisième expose la stratégie algorithmique mise en place pour étalonner le capteur.

## 2.4.1 Modèles de projection et rétro-projection

### 2.4.1.1 Modèle de projection

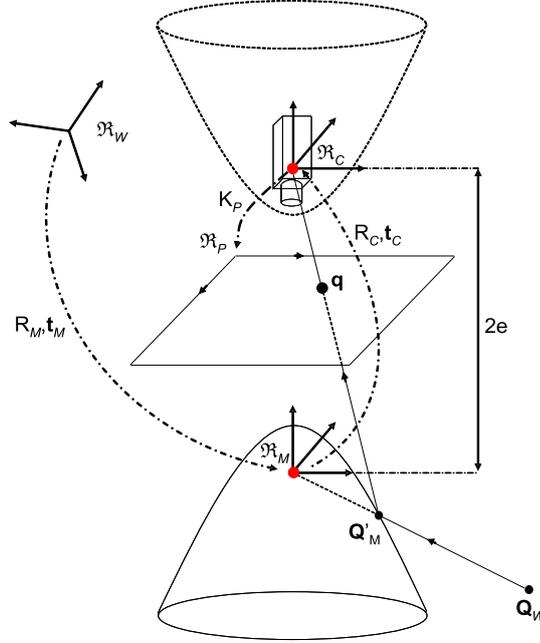


FIG. 2.9 – Détails de la projection d'un point de l'espace 3D pour un capteur hypercatadioptrique [154].

Le modèle de projection associe un point 3D à un point 2D (cf. figure 2.9). Soit  $\mathbf{Q}_W = (X_W, Y_W, Z_W)^\top$  un point 3D défini dans un référentiel monde  $\mathfrak{R}_W$  et  $\mathbf{q} = (u, v, 1)^\top$  son projeté défini en coordonnées homogènes dans un référentiel du plan image de la caméra  $\mathfrak{R}_P$ . Notons  $Q$  la quadrique et  $\mathbf{Q}_M = (X_M, Y_M, Z_M)^\top$  un point 3D défini dans un référentiel miroir  $\mathfrak{R}_M$  centré au foyer de l'hyperboloïde  $Q$ . L'équation de  $Q$  définie dans  $\mathfrak{R}_M$  est donnée par :

$$Q : \frac{(Z_M - e)^2}{\beta^2} - \frac{X_M^2 + Y_M^2}{\gamma^2} = 1 \quad (2.16)$$

$\beta$  et  $\gamma$  sont les demi-axes focaux et non-focaux et  $e = \sqrt{\beta^2 + \gamma^2}$  est l'excentricité de  $Q$ . Pour transformer le point 3D  $\mathbf{Q}_W$  en le point 2D  $\mathbf{q}$ , le point 3D est tout d'abord défini dans  $\mathfrak{R}_M$  en  $\mathbf{Q}_M$  par une matrice de changement de base  $(R_M \mid -\mathbf{t}_M)$ . Soit  $\mathcal{L} = \mathbf{O}_M + \lambda \mathbf{Q}_M = \lambda \mathbf{Q}_M$ ,  $\lambda \in \mathbb{R}_+$ , la demi-droite dans l'espace 3D issue de  $\mathbf{Q}_M$  et passant par  $\mathbf{O}_M$  l'origine du référentiel  $\mathfrak{R}_M$  et également foyer du miroir. Soit  $\mathbf{Q}'_M$  le point 3D d'intersection entre

$\mathbf{Q}$  et  $\mathcal{L}$ .  $\mathbf{Q}'_M$  est ensuite exprimé dans  $\mathfrak{R}_C$  en  $\mathbf{Q}_C$  par une matrice de changement de base  $(\mathbf{R}_C \mid -\mathbf{t}_C)$ . Le système respectant la contrainte de point de vue unique, la matrice de changement de base entre  $\mathfrak{R}_M$  et  $\mathfrak{R}_C$  est parfaitement connue :  $\mathbf{R}_C$  est égale à la matrice identité  $\mathbf{I}_{3 \times 3}$  et  $\mathbf{t}_C = (0, 0, 2e)^\top$ ,  $e$  étant l'excentricité de la quadrique  $\mathbf{Q}$ .  $\mathbf{Q}_C$  est finalement projeté dans  $\mathfrak{R}_P$  par la matrice des paramètres intrinsèques  $\mathbf{K}_P$ . Le coefficient  $1/\delta$  intervenant dans l'équation 2.17 permet l'obtention de coordonnées homogènes normalisées,  $\mathbf{R}_{C_3}^\top$  étant la troisième colonne de  $\mathbf{R}_C$ . L'équation de projection associant un point 2D  $\mathbf{q}$  à un point 3D  $\mathbf{Q}_W$  est donnée comme suit :

$$\left\{ \begin{array}{l} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{K}_P \frac{1}{\delta} \mathbf{R}_C \left( \lambda \mathbf{R}_M \left( \begin{pmatrix} X_W \\ Y_W \\ Z_W \end{pmatrix} - \mathbf{t}_M \right) - \mathbf{t}_C \right) \\ \mathbf{K}_P = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \\ \delta = \mathbf{R}_{C_3}^\top (\lambda \mathbf{R}_M (\mathbf{Q}_W - \mathbf{t}_M) - \mathbf{t}_C) \end{array} \right. \quad (2.17)$$

#### 2.4.1.2 Modèle de rétro-projection

Le modèle de rétro-projection associe un point 2D à une ligne de vue dans l'espace 3D. Soit  $\mathbf{q} = (u, v, 1)^\top$  un point 2D défini dans  $\mathfrak{R}_P$  et  $\mathcal{L}_W$  la ligne de vue dans l'espace 3D associée à  $\mathbf{q}$  et définie dans le référentiel monde  $\mathfrak{R}_W$ . A  $\mathbf{q}$  on associe la ligne de vue 3D  $\mathcal{L}_C$  définie dans  $\mathfrak{R}_C$  par la matrice des paramètres intrinsèques inverse  $\mathbf{K}^{-1}$ , telle que :  $\mathcal{L}_C = \mathbf{O}_C + \lambda \mathbf{K}^{-1} \mathbf{q} = \lambda \mathbf{K}^{-1} \mathbf{q}$ ,  $\lambda \in \mathbb{R}_+$ .  $\mathcal{L}_C$  est exprimée dans  $\mathfrak{R}_M$  en la droite  $\mathcal{L}_M$  par une matrice de changement de base  $(\mathbf{R}_C^\top \mid \mathbf{t}_C)$  :  $\mathcal{L}_M = \lambda \mathbf{R}_C^\top \mathbf{K}^{-1} \mathbf{q} + \mathbf{t}_C$ . Finalement la ligne de vue 3D  $\mathcal{L}_W$  associée à  $\mathbf{q}$  dans  $\mathfrak{R}_W$  est obtenue en utilisant la matrice de changement de base  $(\mathbf{R}_M \mid \mathbf{t}_M)$  et est donnée par :

$$\mathcal{L}_W : \quad \{ \mathcal{L}_W = \mathbf{R}_M^\top (\lambda \mathbf{R}_C^\top \mathbf{K}^{-1} \mathbf{q} + \mathbf{t}_C) + \mathbf{t}_M, \lambda \in \mathbb{R}_+ \} \quad (2.18)$$

### 2.4.2 Estimation des paramètres

L'étalonnage paramétrique d'un capteur de vision consiste en l'estimation des paramètres  $\chi$  qui définissent et caractérisent le modèle. Ainsi, le modèle présenté en l'équation 2.17 est

constitué des paramètres extrinsèques  $(\mathbf{R}_M, \mathbf{t}_M)$  et des paramètres intrinsèques contenus dans la matrice  $\mathbf{K}$  de sorte que :  $\boldsymbol{\chi} = (\mathbf{K}, \mathbf{R}_M, \mathbf{t}_M)$ . Le problème d'étalonnage paramétrique se pose donc comme un problème de minimisation multivariable pour lequel on souhaite déterminer un vecteur de paramètres optimal  $\boldsymbol{\chi}^*$  minimisant une fonction coût  $J(\boldsymbol{\chi})$ . La fonction coût est très généralement définie comme l'erreur quadratique d'un ensemble de points d'intérêts extraits dans les images et des positions estimées par le modèle. Si  $\mathbf{q}_i$  est le  $i^{\text{ème}}$  point d'intérêt extrait de l'image et  $\hat{\mathbf{q}}_i(\boldsymbol{\chi})$  la position correspondante estimée par le modèle, le problème de minimisation s'écrit alors :

$$\begin{cases} \boldsymbol{\chi}^* = \min_{\mathbf{K}, \mathbf{R}_M, \mathbf{t}_M} J(\boldsymbol{\chi}) \\ J(\boldsymbol{\chi}) = \frac{1}{2} \sum_{i=1}^I (\mathbf{q}_i - \hat{\mathbf{q}}_i(\boldsymbol{\chi}))^2 \end{cases} \quad (2.19)$$

La fonctionnelle à minimiser étant non linéaire par rapport aux paramètres à estimer, des algorithmes d'optimisation non linéaire tels que l'algorithme de **Levenberg-Marquardt** sont utilisés [87]. L'algorithme de Levenberg-Marquardt est un standard en Vision par Ordinateur pour la résolution de problèmes de minimisation non linéaire. Il s'agit d'un algorithme d'estimation itératif qui combine la **méthode de Descente de Gradient** et celle de **Gauss-Newton**. Le principe est le suivant : partant d'un point fixé arbitrairement, la méthode de descente du gradient est tout d'abord utilisée pour assurer une décroissance de l'erreur. Ensuite, la méthode de Gauss-Newton est appliquée ce qui permet une convergence rapide vers le minimum de la fonction [67], [73].

Néanmoins, l'algorithme de Levenberg-Marquardt ne se soustrait pas à l'étape essentielle d'initialisation du vecteur de paramètres. En effet, les méthodes de descente de gradient de Gauss-Newton se heurtent au problème des minima locaux. Il est donc nécessaire d'initialiser le processus itératif avec une valeur du vecteur de paramètres qui soit aussi proche que possible du vecteur optimal des estimées. Pour y parvenir, une technique classiquement utilisée conjugue des méthodes d'estimation linéaires pour déterminer le vecteur des estimées initiales  $\boldsymbol{\chi}_0$ , puis les algorithmes de minimisation itératifs sont appliqués pour estimer le vecteur de paramètres optimal  $\boldsymbol{\chi}^*$  [86].

#### 2.4.2.1 Estimation du vecteur initial des estimées $\boldsymbol{\chi}_0$

Soit  $(\mathbf{R}_{M_0}, \mathbf{t}_{M_0})$ , les valeurs initiales de la matrice de rotation  $\mathbf{R}_M$  et du vecteur de translation  $\mathbf{t}_M$ . Sans perte de généralité, on suppose les orientations des deux trièdres identiques,

ce qui conduit à :  $R_{M_0} = I_{3 \times 3}$ . Les axes de révolution du cylindre et optique du capteur sont supposés confondus, ce qui permet d'écrire :  $\mathbf{t}_{M_0} = (0, 0, t_{M_0}^Z)^\top$ ,  $t_{M_0}^Z$  pouvant être aisément connu de par la géométrie de la mire et l'architecture du capteur.

Compte tenu des hypothèses formulées, le modèle de projection s'écrit sous une forme linéaire par rapport aux paramètres à estimer  $(\alpha_u, \alpha_v, u_0, v_0)$ , c'est-à-dire les paramètres intrinsèques de la matrice  $K_P$ . Le modèle s'écrit alors :

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\lambda X_W}{\lambda(Z_W - t_{M_0}^Z) - 2e} \\ \frac{\lambda Y_W}{\lambda(Z_W - t_{M_0}^Z) - 2e} \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi X_W \\ \xi Y_W \\ 1 \end{pmatrix} \quad (2.20)$$

Pour l'ensemble des données utilisées pour étalonner le capteur, le modèle s'écrit :

$$\overbrace{\begin{pmatrix} u_1 & \cdots & u_i & \cdots & u_I \\ v_1 & \cdots & v_i & \cdots & v_I \\ 1 & \cdots & 1 & \cdots & 1 \end{pmatrix}}^{X_{3 \times I}} = \overbrace{\begin{pmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}}^{K_{P_0}} \overbrace{\begin{pmatrix} \xi X_{W_1} & \cdots & \xi X_{W_i} & \cdots & \xi X_{W_I} \\ \xi Y_{W_1} & \cdots & \xi Y_{W_i} & \cdots & \xi Y_{W_I} \\ 1 & \cdots & 1 & \cdots & 1 \end{pmatrix}}^{W_{3 \times I}} \quad (2.21)$$

Une solution est donnée par la matrice pseudo-inverse  $W^+$  de  $W$ , telle que l'estimée initiale de la matrice des paramètres intrinsèques  $K_{P_0}$  soit égale à :

$$K_{P_0} = \overbrace{(W^\top W)^{-1} W^\top X}^{W^+} \quad (2.22)$$

Le vecteur des estimées initiales  $\boldsymbol{\chi}_0$  est donc défini par :  $\boldsymbol{\chi}_0 = (K_{P_0}, R_{M_0}, \mathbf{t}_{M_0})$

#### 2.4.2.2 Estimation du vecteur optimal des paramètres $\boldsymbol{\chi}^*$

Le vecteur initial des estimées  $\boldsymbol{\chi}_0$  étant déterminé, l'algorithme itératif de Levenberg-Marquardt est appliqué à la fonction coût  $J(\boldsymbol{\chi})$  et permet de déterminer le vecteur optimal des paramètres  $\boldsymbol{\chi}^*$  (cf. annexe C). Afin de s'assurer que la rotation  $R_M$  soit une matrice orthonormale, c'est-à-dire que  $R_M^\top R_M = I$ , une décomposition en valeurs singulières (cf. annexe D) est appliquée à chaque itération [165]. Une décomposition en valeurs singulières d'une matrice  $A$  est donnée par :  $A = U \Sigma V^\top$ . La matrice orthonormale la plus proche de  $A$  est donnée par :  $A^\dagger = UV^\top$ .

## 2.5 La mire d'étalonnage

### 2.5.1 Caractéristiques techniques

L'étalonnage paramétrique proposé s'inscrit dans la catégorie des étalonnages forts. Il associe des points 3D  $\mathbf{Q}$  répartis sur une mire d'étalonnage dont les coordonnées sont supposées parfaitement connues dans le repère monde associé à la mire  $\mathfrak{R}_W$ , à leurs projetés  $\mathbf{q}$  dans le plan image.

A l'instar des travaux de Bakstein & Padjla [8] sur l'étalonnage de capteurs « *fish-eye* » ou de ceux proposés par Moldovan & Wada [105] et Comby & al. [32] sur l'étalonnage de capteurs hypercatadioptriques, la forme de la mire tire profit de la symétrie axiale du capteur. Il s'agit d'un cylindre de révolution dont les points d'intérêt 3D sont des marqueurs lumineux de type diode électroluminescente (LED).

Les caractéristiques de la mire sont les suivantes (cf. figure 2.10) :

- la forme de la mire est un cylindre de révolution dont le rayon est égal à 300mm et la hauteur est de 1400mm ;
- les points d'intérêt 3D sont des LEDs de couleur blanche. L'angle de diffusion est supérieur à  $70^\circ$  et des diffuseurs à LED sont assemblés aux LEDs afin d'accroître le phénomène de diffusion. De ce fait, les marqueurs situés en hauteur sur le cylindre d'étalonnage sont observés par le capteur ;
- le montage des LEDs est réalisé de sorte que l'ensemble LED-diffuseur soit positionné à ras de la paroi intérieure du cylindre. Un marqueur lumineux est donc vu par la caméra sous la forme d'un disque lumineux. Les coordonnées 3D du marqueur sont approximées aux coordonnées du centre du disque ;
- les points d'intérêt 3D sont répartis suivant les cercles directeurs  $\{C_1, \dots, C_m, \dots, C_M\}$  et suivant des génératrices. Un cercle  $C_m$  est défini par huit points d'intérêt 3D cocycliques tel que ces points soient équirépartis sur  $C_m$ . De plus, ces huit points d'intérêt appartiennent aux génératrices du cylindre ;
- le repère associé à la mire est noté  $\mathfrak{R}_W$ . Il est défini de sorte que son origine  $\mathbf{O}_W$  soit localisée à la base du cylindre et que l'axe  $\mathbf{Z}_W$  soit d'une part confondu avec l'axe de révolution du cylindre et d'autre part confondu avec l'axe optique du capteur de stéréovision ;

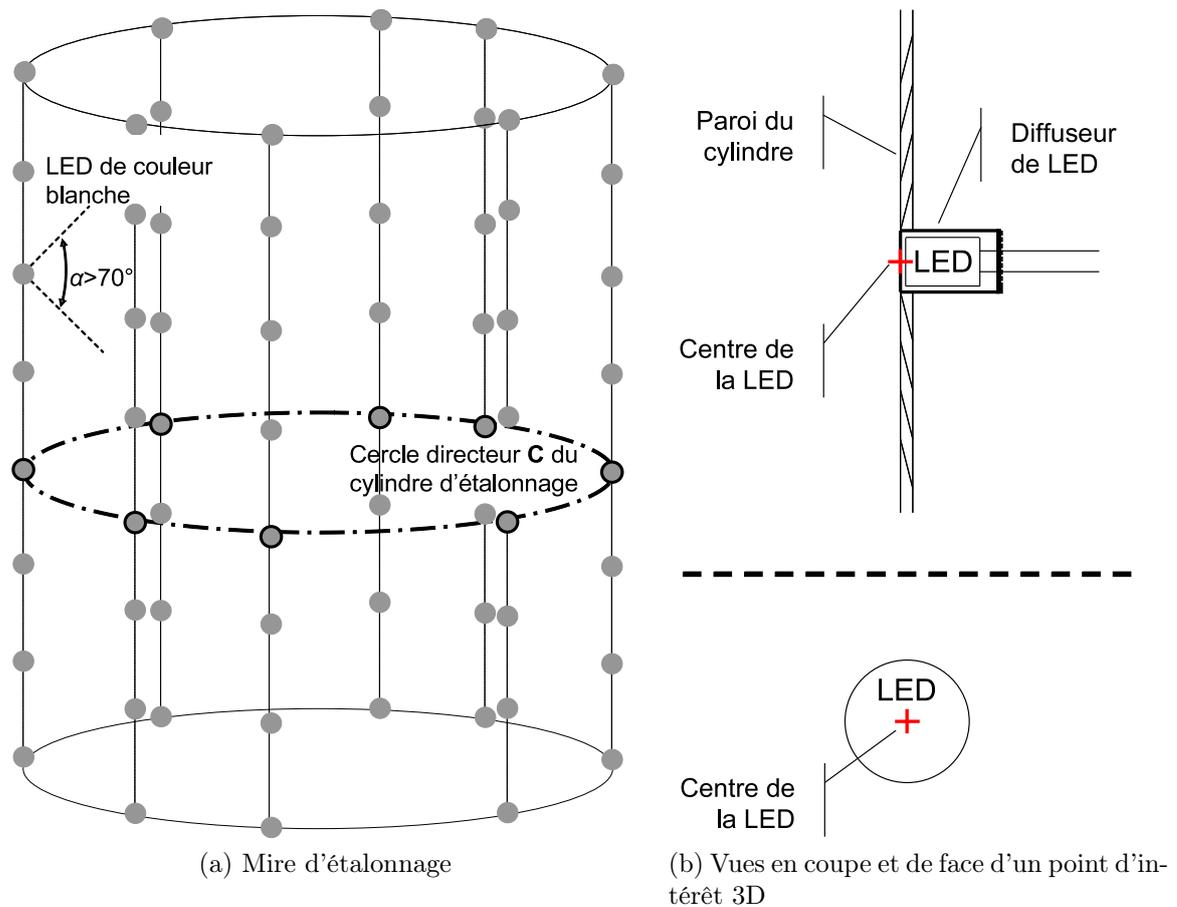


FIG. 2.10 – Représentation de la mire d'étalonnage et différentes vues d'un point d'intérêt 3D. La mire est un cylindre de révolution dont les points d'intérêt 3D sont des LEDs de couleur blanche. L'angle de diffusion des LEDs est supérieur à  $70^\circ$  et des diffuseurs sont couplés aux LEDs ce qui permet d'augmenter la diffusion. Les LEDs sont positionnées à ras de la paroi intérieure du cylindre de sorte que la caméra observe un disque lumineux en 3D.

### 2.5.2 Extraction des points d'intérêt

Les points d'intérêt 3D vus par le système d'acquisition sont des disques lumineux se projetant en des halos lumineux dans l'image omnidirectionnelle. La couleur des tâches est un dégradé de blanc et de bleu, et leurs formes s'apparentent à des ovoïdes (cf. figure 2.11).

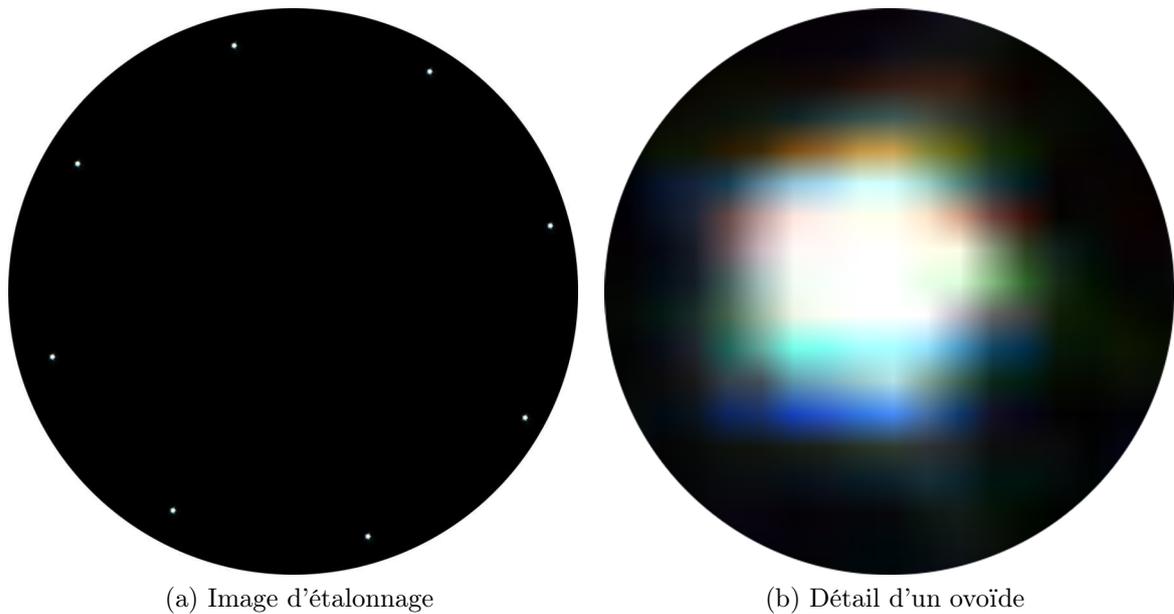


FIG. 2.11 – Exemple d'une image d'étalonnage acquise par le système (mire-capteur) réel et détail d'un ovoïde. Les huit tâches blanches sont les projections des huit marqueurs appartenant à un même cercle directeur du cylindre d'étalonnage. L'ovoïde est un halo dont la couleur est un dégradé de blanc et de bleu.

Les images d'étalonnage sont tout d'abord transformées de l'espace colorimétrique  $RVB$  en l'espace  $CIE L^*a^*b^*$ <sup>17</sup>. Cette transformation est un prérequis nécessaire au vu de l'ovoïde (cf. figure 2.11b). Afin de statuer sur l'appartenance ou non d'un pixel à l'ovoïde, la définition d'un critère sur les seuls paramètres de chrominance ne semble pas judicieuse. Il semble plus pertinent de définir un critère de similarité sur le paramètre de luminance  $L \in [0, 100]$  puisque l'information réellement utile pour caractériser la projection d'un marqueur 3D est l'intensité lumineuse émise par la LED en direction du capteur.

Les ovoïdes sont ensuite extraits de l'image par un algorithme de segmentation par régions homogènes. L'objectif de cette segmentation est de dissocier les pixels constituant le fond de l'image, de ceux qui sont susceptibles d'appartenir à l'ovoïde. L'algorithme implémenté utilise une récursivité locale qui consiste à faire croître au maximum une région  $R$  avant de s'intéresser à la suivante. L'algorithme utilise également la 8-connexité et définit une région  $R$  de l'image comme l'union d'un cardinal de  $I$  pixels  $\{\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_I\}$ . Le critère de décision utilisé pour statuer sur la connexité de deux pixels est défini comme la

<sup>17</sup>L'espace  $CIE L^*a^*b^*$  a été créé en 1976 par la *Commission Internationale de l'Eclairage*. Il possède un paramètre lié à la luminance  $L^*$  et deux paramètres liés à la chrominance  $a^*$  et  $b^*$ . Il s'agit d'un espace couleur perceptuellement uniforme, c'est-à-dire que la distance euclidienne entre deux couleurs est proportionnelle à la différence de couleur perçue par l'oeil humain.

distance euclidienne des paramètres de luminance  $\{L_1, \dots, L_i, \dots, L_I\}$  associés aux pixels  $\{\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_I\}$  [28]. Pour qu'il y ait appartenance du pixel à l'ovoïde, la distance euclidienne doit être inférieure à un seuil  $s$  préalablement fixé. L'influence des valeurs prises par le seuil  $s$  se porte principalement sur la surface de calcul de l'ovoïde, c'est-à-dire sur le nombre de pixels constituant la tâche.

$$\mathbf{R} : \bigcup_{i=1}^I \mathbf{q}_i \mid d(L_i, L_{i+1}) < s \quad (2.23)$$

Finalement, une estimation de barycentre par une moyenne pondérée permet d'approximer ces ovoïdes à des projections ponctuelles. Soit  $\bar{\mathbf{q}} = (\bar{u}, \bar{v})^\top$  le barycentre de la région  $\mathbf{R}$  défini comme la moyenne pondérée par les paramètres de luminance  $\{L_1, \dots, L_i, \dots, L_I\}$  associés aux pixels  $\{\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_I\}$ . Cette approximation suggère implicitement que la surface couverte par l'ovoïde est petite de sorte que la projection du centre du disque lumineux 3D soit définie comme le barycentre de ce dernier.

$$\forall i \in \{1, \dots, I\}, \bar{\mathbf{q}} \in \mathbf{R} : \quad \bar{u} = \frac{\sum_{i=1}^I u_i L_i}{\sum_{i=1}^I L_i}, \quad \bar{v} = \frac{\sum_{i=1}^I v_i L_i}{\sum_{i=1}^I L_i} \quad (2.24)$$

### 2.5.3 Dispositif d'aide à l'étalonnage

Le choix de points d'intérêt 3D de type diode électroluminescente nécessite, d'une part une électronique particulière pour l'alimentation des composants, et d'autre part un dispositif permettant de régler les intensités lumineuses des LEDs : une LED située dans le même plan que le miroir requiert une intensité lumineuse plus faible qu'une LED située en hauteur sur le cylindre.

Ces remarques nous ont amenés à développer un dispositif d'aide à l'étalonnage. Ce dispositif est composé de deux parties distinctes (cf. figure 2.12) : l'une « *software* »<sup>18</sup> propose une interface graphique pour l'étalonnage du capteur. La seconde « *hardware* »<sup>19</sup> est une carte électronique dédiée au pilotage des LEDs. L'inter-connexion des deux blocs est assurée par un microcontrôleur et la communication est réalisée par une liaison série RS232.

<sup>18</sup>Ce travail a fait l'objet d'un stage de Master de M. Ravi Solanki (Manipal Institute of Technology, Inde)

<sup>19</sup>Ce travail a fait l'objet d'un stage de licence de M. Miguel Angel Garcia Domene (Universidad de Valencia, Espagne)

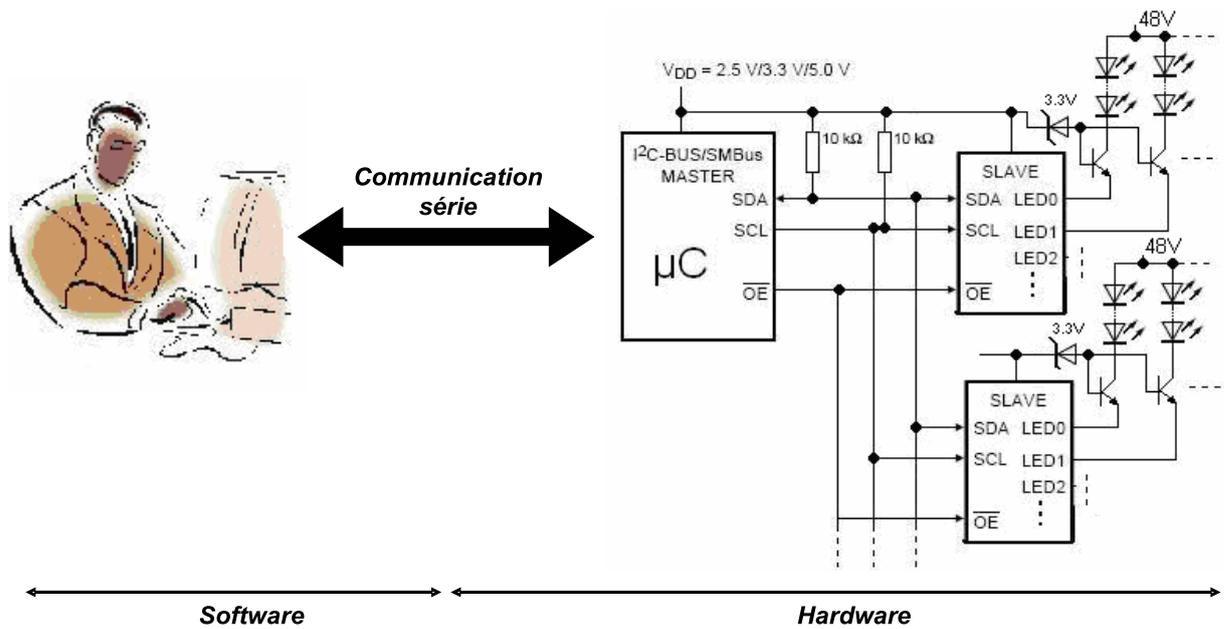


FIG. 2.12 – Schéma du dispositif d'aide à l'étalonnage.

### 2.5.3.1 Interface graphique

L'interface graphique (cf. figure 2.13), développée en langage C++, fait appel à la librairie graphique multi-plateformes « *wxWidgets* ». Les fonctionnalités implémentées doivent simplifier la procédure d'étalonnage du banc de stéréovision. Elles ont trait :

- aux réglages des paramètres des caméras ;
- à l'envoi d'un signal pour l'acquisition des images ;
- à la récupération du flux vidéo en sortie des caméras ;
- au pilotage des LEDs à l'allumage et à l'extinction ;
- au réglage des intensités lumineuses.

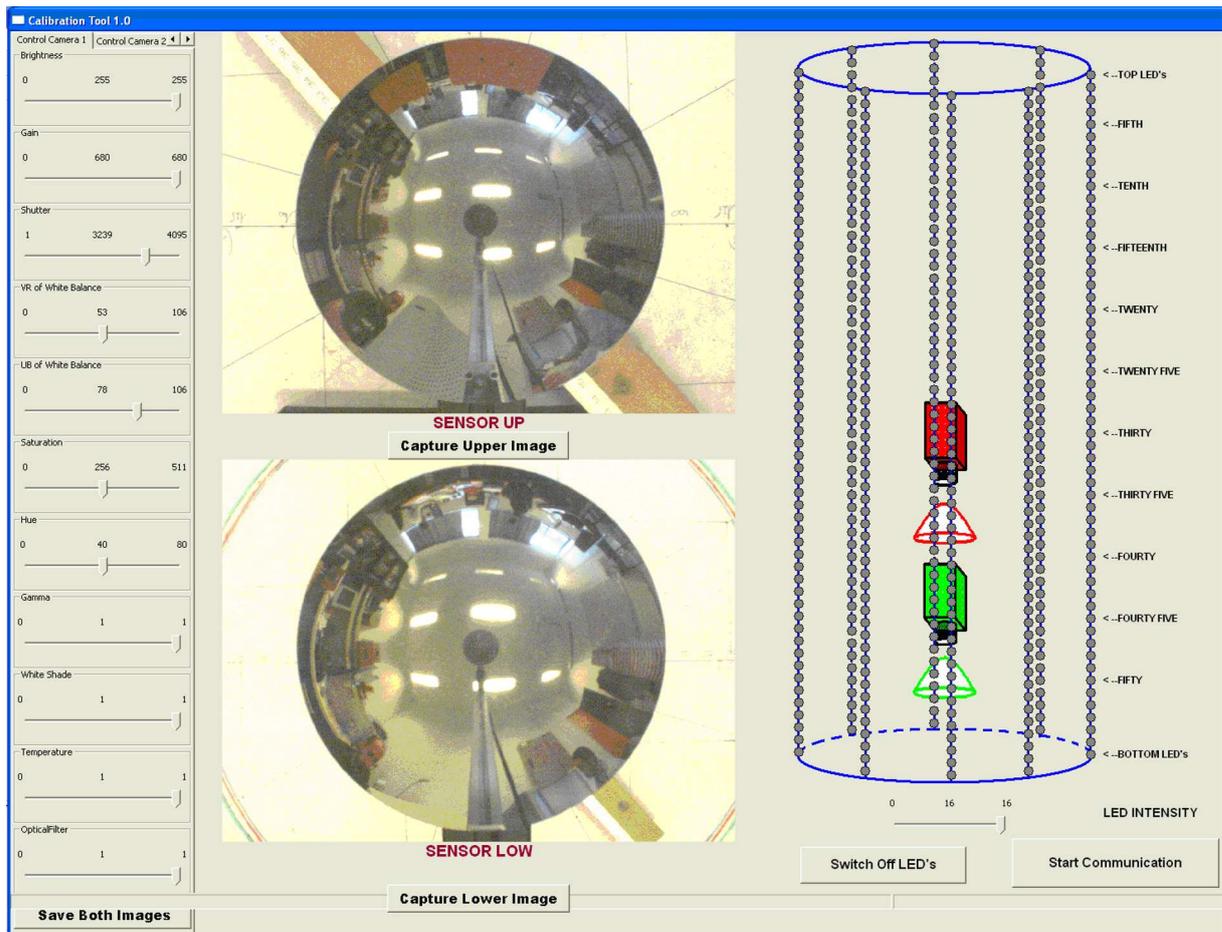
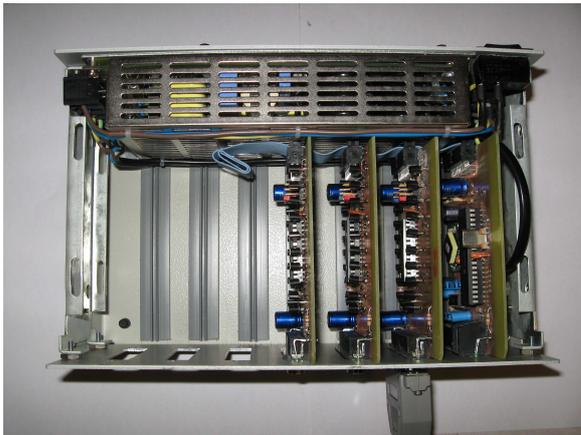


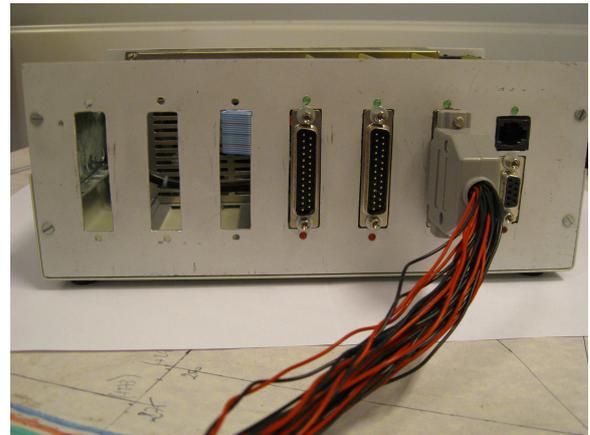
FIG. 2.13 – Interface graphique pour une aide à l'étalonnage.

### 2.5.3.2 Carte électronique

Le dispositif électronique (cf. figure 2.14) est composé d'une carte maître comprenant un microcontrôleur de type PIC et de cartes auxiliaires constituées d'un composant principal : le MAX6956. Le MAX6956 est un driver de LEDs qui permet de modifier directement le courant dans chacune des LEDs et autorise une variation de l'intensité suivant seize niveaux. Ces seize niveaux permettent une variation suffisante des intensités lumineuses des LEDs pour éclairer faiblement les LEDs proches des miroirs et fortement celles qui en sont éloignées. Le MAX6956 fonctionne en composant esclave qui envoie et reçoit des données du microcontrôleur par une interface I2C.



(a) Vue de dessus



(b) Vue de face

FIG. 2.14 – Carte électronique pour le pilotage des marqueurs lumineux 3D.



FIG. 2.15 – Photo de la mire d'étalonnage

## 2.6 Résultats

Les résultats illustrés ci-après sont des résultats obtenus en simulation à l'aide du logiciel POV-Ray. Ces résultats se rapportent à l'étalonnage paramétrique du capteur de stéréovision omnidirectionnelle. Néanmoins, même si le capteur est entièrement immergé dans l'outil d'étalonnage, la procédure implémentée ne permet l'étalonnage que d'un seul capteur à la fois. Ainsi, les résultats présentés n'ont trait qu'à un seul des deux capteurs et sauf mention contraire, il s'agit du capteur bas du système de stéréovision.

Une première partie présente donc l'environnement de simulation modélisant à la fois le capteur et la mire 3D. Le protocole d'étalonnage est également détaillé. Puis, une seconde partie décrit les résultats obtenus lors de l'estimation des paramètres : intrinsèques d'une part et intrinsèques-extrinsèques d'autre part.

### 2.6.1 Environnement de simulation et protocole d'étalonnage

L'environnement de simulation est réalisé avec le logiciel POV-Ray (cf. figure 2.16) et intègre le système de stéréovision omnidirectionnelle, la mire et les marqueurs lumineux, afin de simuler de manière réaliste la procédure d'étalonnage. A l'instar du capteur expérimental, le capteur simulé respecte la contrainte de point de vue unique et possède des caractéristiques techniques et dimensionnelles identiques à ce dernier. La mire 3D virtuelle est également définie au plus juste des caractéristiques du cylindre réel. Les LEDs sont simulées par des cylindres de couleur blanche positionnés perpendiculairement à la surface du cylindre et dont les sommets arrivent au ras de sa paroi.

La procédure d'étalonnage consiste à allumer successivement les LEDs situées sur un même cercle directeur. Pour un cercle directeur allumé, une image est prise et stockée en mémoire. La base d'images créée et le fichier texte contenant les coordonnées des points d'intérêt 3D dans le repère associé à la mire, constituent les éléments centraux de la procédure d'étalonnage. Pour l'étalonnage paramétrique proposé, les cercles directeurs sont espacés verticalement de  $\Delta z = 5\text{mm}$  et un cercle directeur est défini par huit marqueurs. Le cercle directeur le plus bas et observé par le capteur bas est défini en  $z = 275\text{mm}$  et le cercle directeur le plus haut, observé par le capteur bas est localisé en  $z = 1375\text{mm}$ . Au total 221 images d'étalonnage constituent la base de données utilisées pour étalonner le capteur et le fichier texte contient  $221 \times 8 = 1768$  coordonnées de points d'intérêt 3D exprimées dans le repère mire.

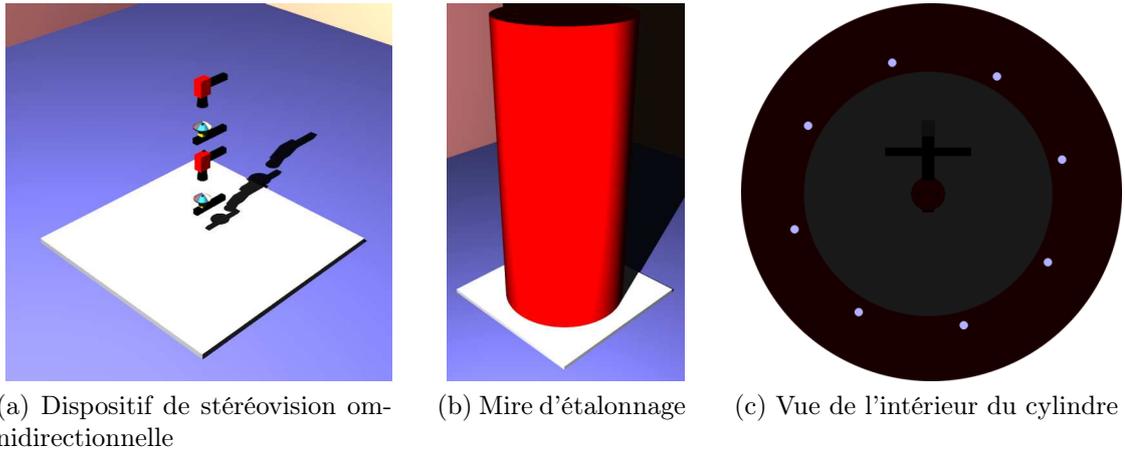


FIG. 2.16 – Environnement pour l'étalonnage du capteur réalisé avec le logiciel de synthèse d'images POV-Ray.

## 2.6.2 Estimation des paramètres

Le vecteur initial des estimées  $\boldsymbol{\chi}_0 = (\mathbf{K}_0, \mathbf{R}_{M_0}, \mathbf{t}_{M_0})$  est obtenu en posant connus les paramètres extrinsèques. Les résultats sont résumés en l'équation 2.25 et la figure 2.17 illustre les points 2D extraits et les points 2D estimés.

$$\boldsymbol{\chi}_0 = (\mathbf{K}_0, \mathbf{R}_{M_0}, \mathbf{t}_{M_0}) : \begin{cases} \mathbf{K}_0 = \begin{pmatrix} 1272.1 & 0 & 639.5 \\ 0 & 1272.1 & 479.5 \\ 0 & 0 & 1 \end{pmatrix} \\ \mathbf{R}_{M_0} = \mathbf{I}_{3 \times 3} \\ \mathbf{t}_{M_0} = (0, 0, t_M^Z = 359\text{mm})^\top \end{cases} \quad (2.25)$$

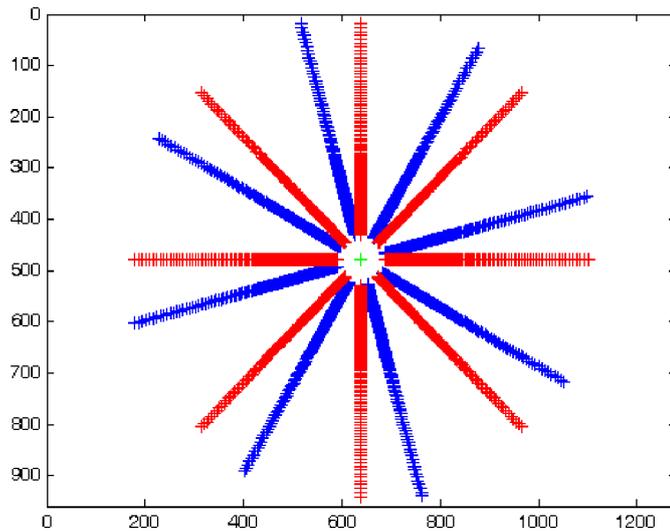


FIG. 2.17 – Initialisation du vecteur de paramètres  $\chi_0$ . Représentation des points 2D extraits des images (en bleu) et des points 2D estimés (en rouge).

Le vecteur des paramètres  $\chi^*$  est obtenu en minimisant une fonction coût  $J(\chi)$  par une méthode itérative utilisant l'algorithme de Levenberg-Marquardt. Il s'agit de ré-estimer les paramètres intrinsèques et extrinsèques du capteur. La résolution de ce problème a été implémentée en utilisant la fonction « **lsqcurvefit** » du logiciel **Matlab**. Le vecteur des estimées  $\chi_0$  est celui déterminé précédemment et le vecteur des paramètres  $\chi^* = (K, R_M, t_M)$  minimisant la fonction coût  $J(\chi)$  est donné par l'équation 2.26. La figure 2.18 illustre les points 2D extraits et les points 2D estimés pour les valeurs du vecteur des paramètres  $\chi^*$ .

Une erreur moyenne de projection a ensuite été calculée sur l'ensemble des points 2D utilisés pour étalonner le capteur. Elle est égale à  $\bar{\varepsilon} = 0.02 \pm 0.29\text{pix}$ .

$$\boldsymbol{\chi}^* = (\mathbf{K}, \mathbf{R}_M, \mathbf{t}_M) : \begin{cases} \mathbf{K} = \begin{pmatrix} 1313.1 & 0 & 639.5 \\ 0 & 1313.1 & 479.6 \\ 0 & 0 & 1 \end{pmatrix} \\ \mathbf{R}_M = \begin{pmatrix} 0.9659 & -0.2588 & -0.0003 \\ 0.2588 & 0.9659 & -0.0002 \\ 0.0004 & 0.0001 & 1.0000 \end{pmatrix} \\ \mathbf{t}_M = \begin{pmatrix} 0.0094 \\ -0.1053 \\ 359.8841 \end{pmatrix} \end{cases} \quad (2.26)$$

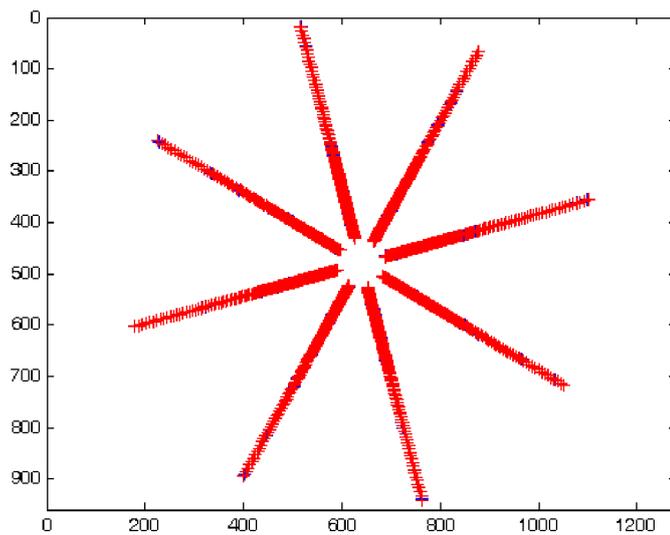


FIG. 2.18 – Estimation du vecteur optimal des paramètres  $\boldsymbol{\chi}^*$  minimisant la fonction coût  $J(\boldsymbol{\chi})$ . Représentation des points 2D extraits des images d'étalonnage (en bleu) et les points 2D estimés (en rouge).

## 2.7 Discussion

Comme illustrée sur la figure 2.17, la phase d'initialisation des paramètres consiste en une pré-estimation des paramètres intrinsèques. Un biais apparaît entre les points 2D extraits et ceux estimés, ce qui se justifie par le fait que la matrice  $R_M$  est initialisée à l'identité.

La figure 2.18 présente les points 2D extraits et ceux estimés après l'étape d'estimation globale utilisant la méthode itérative de Levenberg-Marquardt. Les points extraits et ceux estimés coïncident alors parfaitement. L'ensemble de la procédure est donc validée puisque les valeurs estimées des paramètres sont cohérentes par rapport aux valeurs théoriques (cf. tableau 2.2).

	K				$R_M$ en °			$t_M$ en mm		
	$\alpha_u$	$\alpha_v$	$u_0$	$v_0$	$\alpha_x$	$\alpha_y$	$\alpha_z$	$t_M^x$	$t_M^y$	$t_M^z$
valeurs théoriques	1315.8	1315.8	640.0	480.0	0.00	0.00	15.00	0.0	0.0	359.0
valeurs estimées	1313.1	1313.1	639.5	479.6	0.00	-0.02	14.99	0.0	-0.1	359.8

TAB. 2.2 – Valeurs théoriques et estimées des paramètres pour l'étalonnage paramétrique

La validité de l'étalonnage est confirmée par le calcul de l'erreur moyenne de projection obtenue sur l'intégralité des 1768 points 2D extraits et-ou estimés et égale à  $\bar{\varepsilon} = 0.02 \pm 0.29$ pix.

L'étalonnage étant obtenu dans des conditions idéales - respect de la contrainte de point de vue unique, coordonnées des points d'intérêt 3D idéalement connues, bruits de numérisation inexistant - l'erreur résiduelle est en fait due aux approximations lors de l'estimation des projections des centres des LEDs (cf. paragraphe 2.5.2). La projection du centre d'une LED est approximée par un barycentre calculé par une moyenne pondérée des valeurs de luminance des pixels constituant l'ovoïde. La faible valeur de l'erreur moyenne de projection permet donc de valider cette hypothèse.

L'étalonnage paramétrique d'un capteur catadioptrique se pose comme un problème de minimisation non linéaire multivariable. La stratégie de résolution adoptée dans notre cas consiste à :

- pré-estimer les paramètres intrinsèques par une méthode d'estimation linéaire ;
- ré-estimer l'intégralité des paramètres par une méthode de minimisation itérative ;

→ contraindre a posteriori l'orthonormalité des matrices de rotation (décomposition en valeurs singulières ou orthonormalisation de Gram-Schmidt).

Il s'agit de « forcer » a posteriori la matrice estimée à être une matrice de rotation. Ce problème aurait pu être envisagé différemment par des méthodes qui ne nécessitent pas une orthonormalisation a posteriori. On peut tout d'abord penser à une paramétrisation ad hoc par les angles d'Euler. Cette notation introduit néanmoins deux principaux problèmes : l'estimation de la matrice jacobienne est plus délicate et pour certaine configuration la perte d'un degré de liberté peut apparaître (cf. « *gimbal lock* »). Pour pallier en partie ce problème [162], Hamilton a introduit en 1843 les quaternions. Il s'agit d'une notation sous forme d'un quadruplet  $(x, y, z, \sigma)^\top$  qui peut être interprétée comme la rotation dans un espace 4D autour d'un axe  $(x, y, z)^\top$  avec un angle  $\sigma$  [134], [120].

## 2.8 Conclusions

Au travers de ce **CHAPITRE 2** nous avons proposé un système de perception 3D qui offre une solution pertinente pour les applications du programme *OmniViss*. Il s'agit d'un système de stéréovision binoculaire omnidirectionnelle constitué de deux capteurs hypercatadioptriques. Les deux capteurs sont positionnés verticalement et leurs axes optiques sont confondus.

La suite du **CHAPITRE 2** s'est rapportée à l'étalonnage des capteurs catadioptriques en général et au dispositif développé pour le programme *OmniViss* en particulier. Une revue des modèles mathématiques et des méthodes d'étalonnage pour des capteurs catadioptriques a tout d'abord été proposée. Puis, une procédure classique d'étalonnage paramétrique utilisant un modèle ad hoc et une mire 3D a été implémentée. Mathématiquement, cet étalonnage peut être vu comme un problème de minimisation non linéaire multivariable. La solution retenue s'appuie sur l'algorithme de Levenberg-Marquardt et sur une procédure adaptée pour l'estimation des paramètres initiaux.

Le **CHAPITRE 2** se poursuit par une description détaillée de la mire d'étalonnage réalisée. Il s'agit d'un cylindre de révolution qui, de part sa forme, tire profit de la symétrie axiale du capteur. La répartition des points d'intérêt est effectuée à la fois, suivant des cercles directeurs du cylindre et suivant des génératrices. Les points d'intérêts sont des LEDs de couleur blanche dont les projections sont des ovoïdes dans l'image omnidirectionnelle. Un dispositif d'aide à l'étalonnage constitué d'une interface homme-machine et d'une électronique particulière a été créée afin de faciliter la procédure.

Le **CHAPITRE 2** se termine par une présentation des résultats d'étalonnage obtenus par une approche classique paramétrique. Ils mettent en avant une erreur moyenne de projection faible, tout en prenant soin de repositionner ces résultats dans leur contexte d'obtention. Il s'agit de résultats de simulation qui ont permis d'une part de vérifier la validité du modèle proposé par Svoboda et d'autre part de confirmer l'estimation de la projection du centre de la LED par un calcul de moyenne pondérée par les valeurs de luminance des pixels constituant l'ovoïde.

Les modèles en général et celui-ci en particulier pré-supposent l'unicité du centre de projection, c'est-à-dire que les capteurs respectent la contrainte de point de vue unique. Comme le mentionnent Fabrizio & al. [45] et Ieng & al. [72], cette hypothèse, même si elle simplifie la modélisation des capteurs, est difficile à mettre en oeuvre en pratique.

Il semble donc judicieux de pouvoir étalonner le banc de stéréovision par une approche qui s'affranchisse des contraintes inhérentes aux modèles et qui soit transposable sur un système dont les ressources (taille mémoire, temps d'exécution) sont limitées.

# Chapitre 3

## Etalonnage non-paramétrique et conception d'une mire adaptée

### Sommaire

---

<b>3.1</b>	<b>Introduction</b>	<b>87</b>
<b>3.2</b>	<b>Etalonnage non-paramétrique par un modèle discret</b>	<b>88</b>
3.2.1	Principe	88
3.2.2	Projection	90
3.2.3	Rétro-projection	93
<b>3.3</b>	<b>Conception d'une mire adaptée</b>	<b>97</b>
3.3.1	Nombre de points d'intérêt par cercle directeur	97
3.3.2	Distribution verticale des cercles directeurs sur le cylindre	99
<b>3.4</b>	<b>Evaluation en simulation et sur le système réel</b>	<b>102</b>
3.4.1	Protocoles expérimentaux et méthodes de validation	102
3.4.2	Résultats	105
<b>3.5</b>	<b>Discussion</b>	<b>115</b>
<b>3.6</b>	<b>Conclusions</b>	<b>121</b>

---

## 3.1 Introduction

Les fonctions de projection  $(u, v) = g(X, Y, Z)$  et rétro-projection  $(X, Y, Z) = g^*(u, v)$  sont généralement décrites par des modèles paramétriques. Les modèles utilisés, lorsqu'ils existent, sont des approximations du processus de formation d'images. Le nombre de paramètres est un compromis entre la complexité et le degré de précision requis.

Nous présentons dans ce chapitre une méthode d'étalonnage s'affranchissant des contraintes inhérentes aux approches paramétriques en ce sens qu'elle n'utilise aucun modèle analytique pour établir les relations entre l'espace 3D et le plan image. L'approche d'étalonnage proposée est qualifiée de **non-paramétrique** puisque l'ensemble miroir-caméra est considéré comme une boîte noire. Elle est également qualifiée de **discrète** puisque le modèle de caméra utilisé est un modèle discret qui caractérise la relation entre les points 2D et les lignes de vue 3D par un échantillon de correspondances. La relation caractérisant la transformation entre des points 2D et des lignes de vue en 3D s'appuie sur le fait que la distance d'un point 2D au centre de l'image (c'est-à-dire le rayon du pixel dans un référentiel défini au centre de l'image) caractérise le vecteur directeur de la ligne de vue 3D associée à ce point 2D.

Ce **CHAPITRE 3** présente notre contribution au problème d'étalonnage de capteurs catadioptriques. Il s'articule en trois parties : une première définit les relations entre l'espace 3D et le plan image. La deuxième détaille les procédures algorithmiques pour projeter des points 3D dans l'image omnidirectionnelle et rétro-projeter des points 2D dans l'espace 3D. La troisième s'intéresse à la conception d'une mire d'étalonnage pour que la distribution des points d'intérêt soit la mieux adaptée au capteur. Une quatrième partie présente les résultats obtenus. Finalement, une discussion et une conclusion viennent clore ce chapitre.

## 3.2 Etalonnage non-paramétrique par un modèle discret

### 3.2.1 Principe

Pour un capteur catadioptrique dont la forme du miroir est convexe ou concave, il existe une relation entre l'angle d'élévation d'un point 3D et le rayon de sa projection dans l'image omnidirectionnelle. Par exemple, à un point 3D porté par l'axe optique correspond un rayon nul dans l'image omnidirectionnelle (le projeté se trouvant au centre de l'image). La méthode proposée s'appuie sur cette relation pour déterminer les fonctions de projection  $g$  et rétro-projection  $g^*$  décrivant le passage entre l'espace 3D et le plan image.

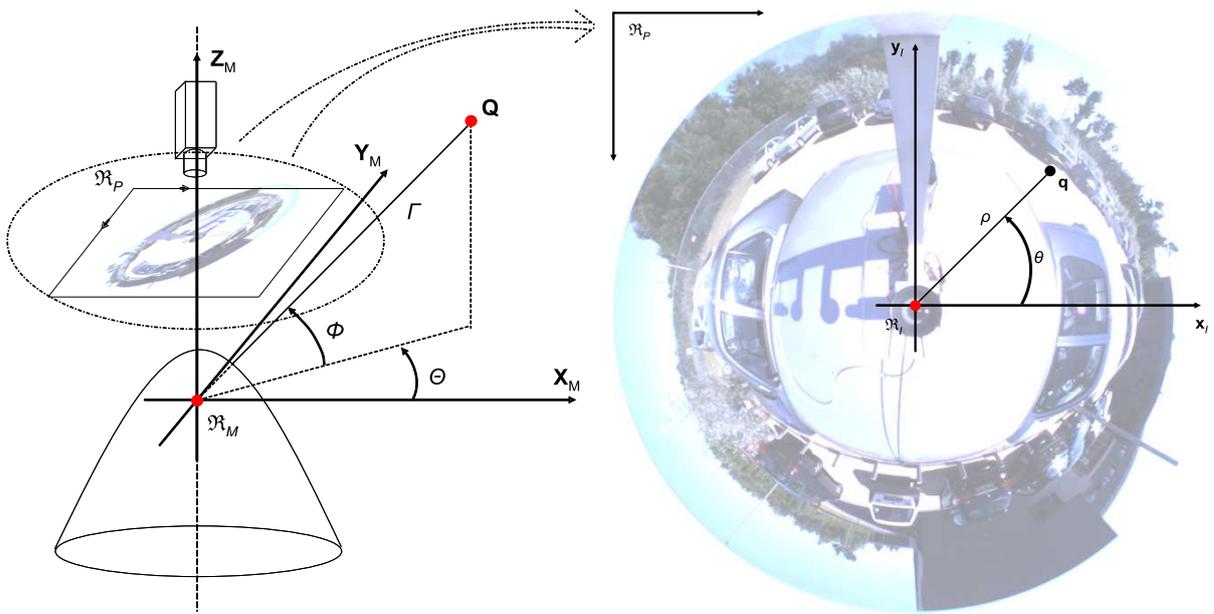


FIG. 3.1 – Relations entre l'espace 3D et le plan image pour un étalonnage non-paramétrique.

Soit  $\mathbf{Q} = (\Theta, \Phi, \Gamma)^\top$  un point 3D en coordonnées sphériques défini dans un référentiel  $\mathfrak{R}_M$  dont l'origine est le foyer du miroir. Les vecteurs de  $\mathfrak{R}_M$  forment une base orthonormée directe et l'axe  $\mathbf{Z}_M$  est colinéaire avec l'axe optique du capteur. Par convention, l'angle d'azimut du point 3D est noté  $\Theta \in [-\pi, \pi]$ , l'angle d'élévation est noté  $\Phi \in ]-\pi/2, \pi/2[$  et  $\Gamma \in \mathbb{R}_+$  est le rayon. Soit  $\mathbf{q} = (\theta, \rho)^\top$  un point 2D en coordonnées polaires défini dans un référentiel  $\mathfrak{R}_I$  dont l'origine est le centre de l'image (cf. figure 3.1). La fonction de

projection  $g$  établit donc la correspondance entre un point 3D  $\mathbf{Q}$  et un point 2D  $\mathbf{q}$ , et la fonction de rétro-projection  $g^*$  définit les relations entre un point 2D  $\mathbf{q}$  et la ligne de vue 3D  $\mathcal{L}$  correspondante. En écriture paramétrique une demi-droite  $\mathcal{L}$  est définie par son origine et un vecteur directeur, tel que de manière générale :  $\mathcal{L} = \mathbf{O} + \lambda \mathbf{d}$ ,  $\lambda \in \mathbb{R}_+$ . L'origine choisie est celle du référentiel, c'est-à-dire le foyer du miroir  $\mathbf{O}_M$ . Le vecteur directeur  $\mathbf{d}$  est fonction des angles d'azimut  $\Theta$  et d'élévation  $\Phi$ . En conséquence les fonctions  $g$  et  $g^*$  sont définies par :

$$\begin{aligned} g : \quad \mathbb{R}^3 &\rightarrow \mathbb{R}^2 \\ (\Theta, \Phi, \Gamma) &\mapsto (\theta, \rho) = g(\Theta, \Phi, \Gamma) \end{aligned} \quad (3.1)$$

$$\begin{aligned} g^* : \quad \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (\theta, \rho) &\mapsto (\Theta, \Phi) = g^*(\theta, \rho) \end{aligned} \quad (3.2)$$

- lorsque le capteur est correctement assemblé, c'est-à-dire l'axe optique de la caméra et l'axe de révolution du miroir sont confondus,  $\Theta$  et  $\theta$  décrivent le même angle, noté  $\theta$  par la suite, ce qui conduit à définir un point 3D et un point 2D par :  $\mathbf{Q} = (\theta, \Phi, \Gamma)^\top$ ,  $\mathbf{q} = (\theta, \rho)^\top$  ;
- des points 3D appartenant à une même demi-droite se projettent en un même point 2D. En conséquence, la fonction  $g$  est invariante par rapport à la variable  $\Gamma$  du point 3D ;
- lorsque le miroir est de révolution, les fonctions  $g$  et  $g^*$  sont invariantes par rapport à l'angle d'azimut  $\theta$ . Cependant, afin de pouvoir généraliser l'approche à des systèmes bas-coût, il semble judicieux de prendre en compte des défauts d'usinage et autres effets indésirables pouvant apparaître. Les fonctions  $g$  et  $g^*$  ne sont donc plus considérées comme invariantes par rapport à l'angle d'azimut  $\theta$ .

On adoptera donc pour la suite les définitions de  $g$  et  $g^*$  telles que celles proposées en l'équation 3.3 et en l'équation 3.4 :

$$\begin{aligned} g : \quad \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (\theta, \Phi) &\mapsto \rho = g(\theta, \Phi) \end{aligned} \quad (3.3)$$

$$\begin{aligned} g^* : \quad \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (\theta, \rho) &\mapsto \Phi = g^*(\theta, \rho) \end{aligned} \quad (3.4)$$

Les fonctions  $g$  et  $g^*$  sont des fonctions d'un nouvel espace de représentation paramétré par le triplet de variables  $(\theta, \Phi, \rho)$ . A chaque couple point 3D-point 2D  $(\mathbf{Q}, \mathbf{q})$  correspond un point  $\mathbf{V}$  dans ce nouvel espace :

$$(\mathbf{Q}, \mathbf{q}) \longrightarrow \mathbf{V}(\theta, \Phi, \rho) \quad (3.5)$$

La procédure d'étalonnage ne proposant qu'une mise en correspondance d'un nombre fini de points 3D et de leurs projetés, les fonctions  $g$  et  $g^*$  ne seront connues qu'en un nombre fini de points. On définit donc un **maillage** sur l'espace des variables  $(\theta, \Phi, \rho)$  formé des points notés  $\{\mathbf{V}_1, \dots, \mathbf{V}_i, \dots, \mathbf{V}_I\}$ . Au  $i^{\text{ième}}$  couple point 3D-point 2D  $(\mathbf{Q}_i, \mathbf{q}_i)$  de l'étalonnage correspond un point  $\mathbf{V}_i$  dans ce nouvel espace tel que :

$$\forall i \in \{1, \dots, I\}, \quad (\mathbf{Q}_i, \mathbf{q}_i) \longrightarrow \mathbf{V}_i(\theta_i, \Phi_i, \rho_i) \quad (3.6)$$

Chaque couple point 3D-point 2D  $(\mathbf{Q}_i, \mathbf{q}_i)$  ne fournit qu'une estimation locale des valeurs des fonctions de projection et rétro-projection. Il est donc nécessaire de disposer d'algorithmes d'**interpolation numérique** afin d'estimer des approximations de  $g$  et  $g^*$  pour tous les couples point 3D-point 2D [2].

### 3.2.2 Projection

Soit  $\widehat{\mathbf{Q}} = (\widehat{\theta}, \widehat{\Phi}, \widehat{\Gamma})^\top$  un point 3D de coordonnées sphériques connues dans  $\mathfrak{R}_M$  pour lequel on souhaite estimer une approximation de sa projection  $\widehat{\mathbf{q}} = (\widehat{\theta}, \widehat{\rho})^\top$  telle que :  $\widehat{\mathbf{q}} = g(\widehat{\mathbf{Q}})$ . L'approximation du point 2D  $\widehat{\mathbf{q}}$  se résume à l'estimation de son rayon  $\widehat{\rho}$  dans le référentiel  $\mathfrak{R}_I$ . Cette estimation repose principalement sur la connaissance de l'orientation du vecteur directeur  $\widehat{\mathbf{d}}$  de la ligne de vue passant par le point 3D paramétré par les angles  $(\widehat{\theta}, \widehat{\Phi})$ . Pour ce faire, la méthodologie proposée recherche tout d'abord les trois points du maillage 'les plus proches' du point à approximer. Une estimation  $g(\widehat{\mathbf{Q}})$  est alors obtenue par une approximation linéaire à l'aide des trois points sélectionnés.

La méthode d'estimation locale du rayon  $\widehat{\rho}$  du point 2D, connaissant le vecteur directeur  $\widehat{\mathbf{d}}$  de la ligne de vue passant le point 3D, est décrite comme suit :

→ à un point  $\widehat{\mathbf{Q}}$  correspondent les variables  $(\widehat{\theta}, \widehat{\Phi})$  du point quelconque  $\widehat{\mathbf{V}} = (\widehat{\theta}, \widehat{\Phi}, \widehat{\rho})^\top$  correspondant dans l'espace des variables  $(\theta, \Phi, \rho)$  (cf. figure 3.2) ;

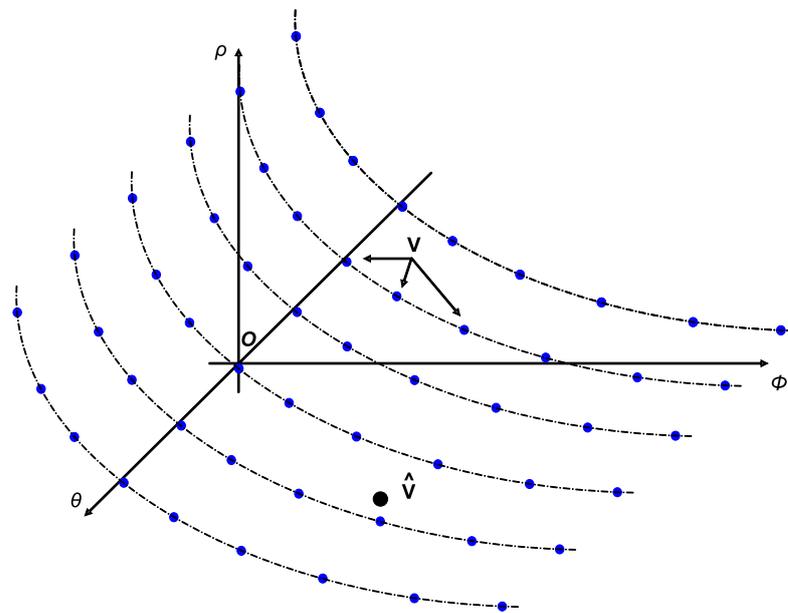


FIG. 3.2 – Représentation de l'espace des variables  $(\theta, \Phi, \rho)$ . Au nouvel espace de représentation est défini un maillage formé des points  $\mathbf{V}$  (en couleur bleue). Au point 3D  $\widehat{\mathbf{Q}}$  dont on souhaite estimer la projection  $\widehat{\mathbf{q}}$  est associé un point  $\widehat{\mathbf{V}}$  (en couleur noire) dans ce nouvel espace.

- les points  $\mathbf{V}$  sont projetés dans le plan  $(\theta\mathbf{O}\Phi)$  (cf. figure 3.3);
- on applique une triangulation de **Delaunay** au nuage de points projeté dans le plan  $(\theta\mathbf{O}\Phi)$  (cf. figure 3.3);
- on sélectionne le triplet de points  $\{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\}$  dont les projetés forment le triangle de Delaunay contenant le projeté de  $\widehat{\mathbf{V}}$ ;

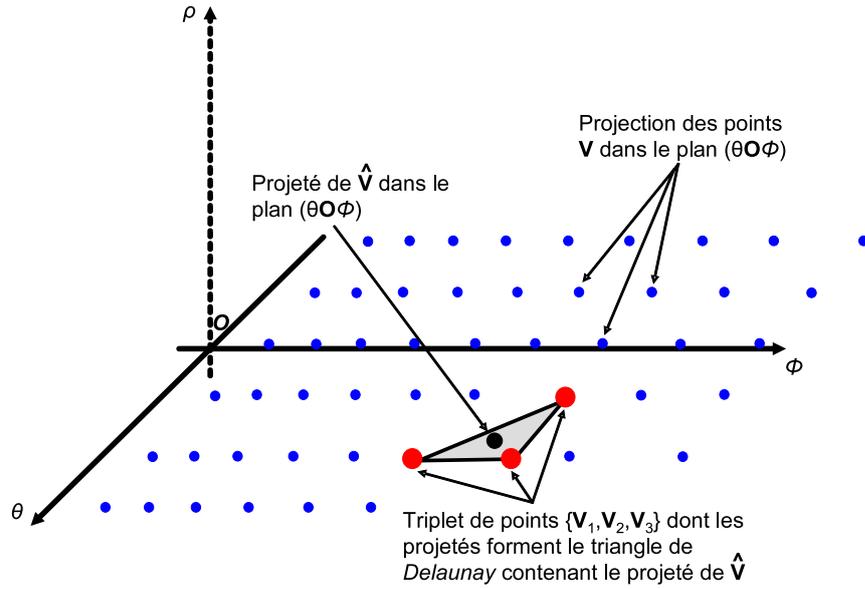


FIG. 3.3 – Projection dans le plan  $(\theta\mathbf{O}\Phi)$  des points  $\mathbf{V}$  (en bleu) et du point quelconque  $\hat{\mathbf{V}}$  (en noir). Une triangulation de Delaunay permet de sélectionner le triplet de points  $\{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\}$  dont les projetés forment le triangle de Delaunay contenant le projeté de  $\hat{\mathbf{V}}$ .

→ le triplet de points sélectionné définit un plan  $\Pi$  de paramètres  $(\eta, \mu, \varrho)^\top$  tel que (cf. figure 3.4) :

$$\forall j \in \{1, 2, 3\}, \{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\} \in \Pi \text{ ssi : } \rho_j = \eta\theta_j + \mu\Phi_j + \varrho \quad (3.7)$$

→  $\hat{\mathbf{V}}$  appartient à  $\Pi$  et vérifie donc son équation :

$$\hat{\mathbf{V}} \in \Pi \text{ ssi : } \hat{\rho} = \eta\hat{\theta} + \mu\hat{\Phi} + \varrho \quad (3.8)$$

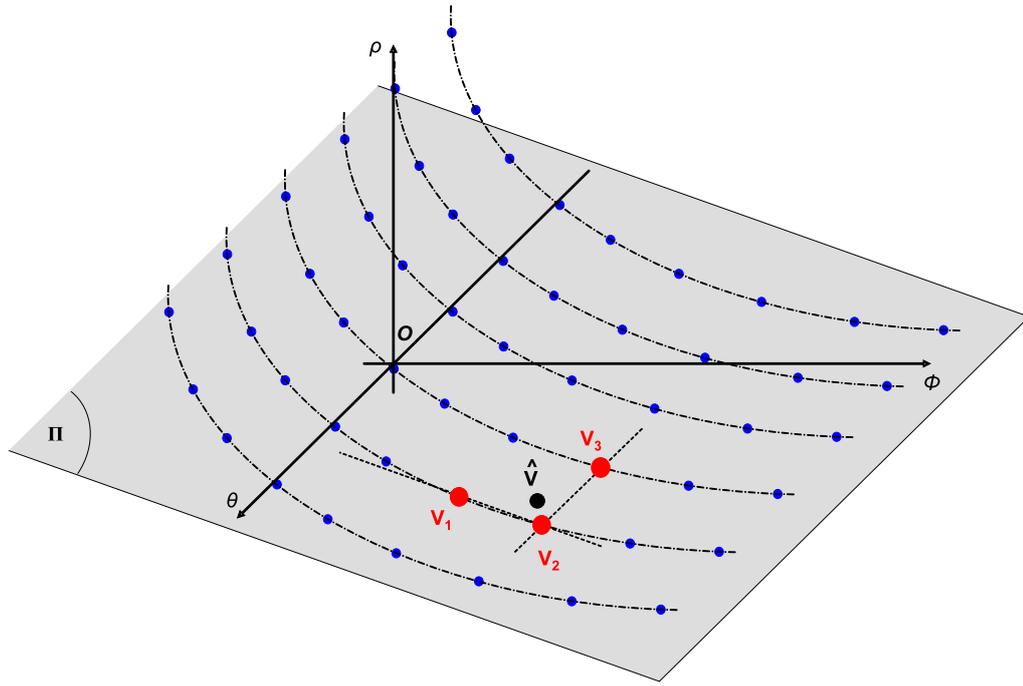


FIG. 3.4 – Dans l'espace des variables  $(\theta, \Phi, \rho)$  le triplet de points  $\{\mathbf{V}_1, \mathbf{V}_2, \mathbf{V}_3\}$  définit un plan  $\Pi$ .  $\hat{\mathbf{V}}$  appartient à  $\Pi$  et vérifie son équation, ce qui permet de déterminer la valeur du rayon  $\hat{\rho}$  du point 2D  $\hat{\mathbf{q}}$ .

→ le point 2D  $\hat{\mathbf{q}} = (\hat{\theta}, \hat{\rho})^\top$  est donc défini dans  $\mathfrak{R}_I$ . Une transformation polaire en cartésien permet de déduire les coordonnées  $(\hat{x}, \hat{y})$  dans  $\mathfrak{R}_I$ . Les coordonnées du point 2D  $\hat{\mathbf{q}} = (\hat{u}, \hat{v})^\top$  dans  $\mathfrak{R}_P$  se déduisent des coordonnées  $(\hat{x}, \hat{y})$  par un changement de variables donné par :  $\hat{u} = \hat{x} + u_0$ ,  $\hat{v} = -\hat{y} + v_0$ , où  $(u_0, v_0)^\top$  sont les coordonnées du centre  $\mathbf{c}$  de l'image dans  $\mathfrak{R}_P$ .

### 3.2.3 Rétro-projection

Soit  $\hat{\mathbf{q}} = (\hat{\theta}, \hat{\rho})^\top$  un point 2D quelconque défini dans  $\mathfrak{R}_I$ . On souhaite estimer une approximation du vecteur directeur  $\hat{\mathbf{d}}$  de la ligne de vue 3D  $\hat{\mathcal{L}}$  défini dans  $\mathfrak{R}_M$  associée à  $\hat{\mathbf{q}}$  telle que :  $\hat{\mathbf{d}} = g^*(\hat{\mathbf{q}})$ . Le vecteur directeur  $\hat{\mathbf{d}}$  est paramétré par les angles d'azimut et d'élévation  $(\hat{\theta}, \hat{\Phi})$ . Le point 2D fournissant l'angle d'azimut  $\hat{\theta}$ , le vecteur directeur  $\hat{\mathbf{d}}$  à approximer consiste en l'estimation de l'angle d'élévation  $\hat{\Phi}$ . La méthode est très similaire à celle décrite précédemment et se compose des étapes suivantes :

→ à un point  $\hat{\mathbf{q}}$  correspondent les variables  $(\hat{\theta}, \hat{\rho})$  du point quelconque  $\hat{\mathbf{V}} = (\hat{\theta}, \hat{\Phi}, \hat{\rho})^\top$  correspondant dans l'espace des variables  $(\theta, \rho, \Phi)$  (cf. figure 3.5) ;

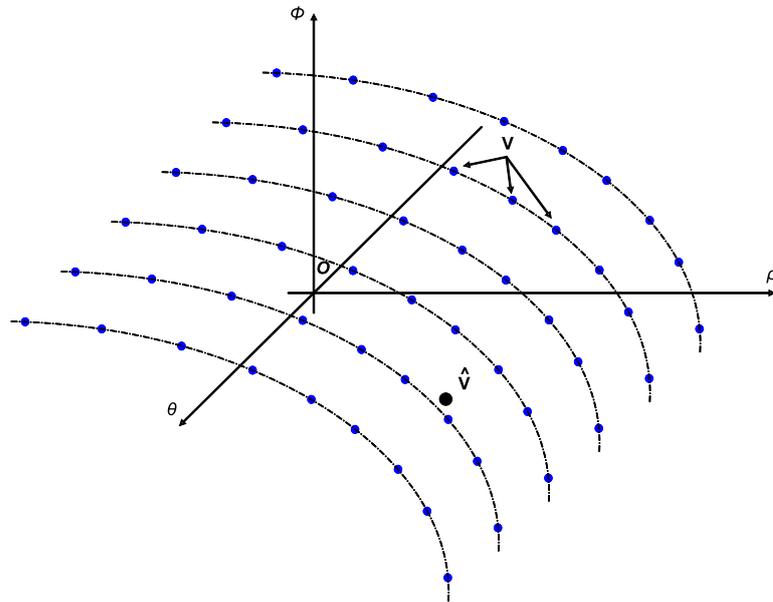


FIG. 3.5 – Représentation de l'espace des variables  $(\theta, \rho, \Phi)$ . Les points  $\mathbf{V}$  du maillage sont représentés en bleu. Le point  $\widehat{\mathbf{V}}$  dont on souhaite estimer le vecteur directeur  $\widehat{\mathbf{d}}$  de la ligne de vue 3D  $\widehat{\mathcal{L}}$  est représenté en noir.

- les points  $\mathbf{V}$  sont projetés dans le plan  $(\theta\mathbf{O}\rho)$  (cf. figure 3.6) ;
- on applique une triangulation de Delaunay au nuage de points projeté dans le plan  $(\theta\mathbf{O}\rho)$  (cf. figure 3.6) ;
- on sélectionne le triplet de points  $\{\mathbf{V}'_1, \mathbf{V}'_2, \mathbf{V}'_3\}$  dont les projetés forment le triangle de Delaunay contenant le projeté de  $\widehat{\mathbf{V}}$  ;

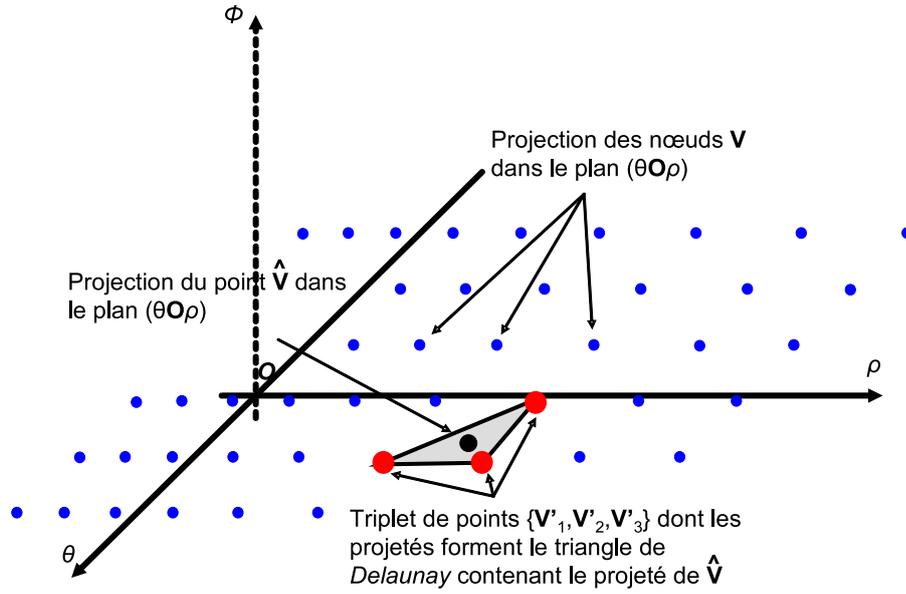


FIG. 3.6 – Projection dans le plan  $(\theta\mathbf{O}\rho)$  des points  $\mathbf{V}$  (en bleu) et du point  $\hat{\mathbf{V}}$ . Une triangulation de Delaunay permet de sélectionner le triplet de points  $\{\mathbf{V}'_1, \mathbf{V}'_2, \mathbf{V}'_3\}$  dont les projetés forment le triangle de Delaunay contenant le projeté de  $\hat{\mathbf{V}}$ .

→ Le triplet de points sélectionné définit un plan  $\Pi'$  de paramètres  $(\eta', \mu', \varrho')^\top$  tel que (cf. figure 3.7) :

$$\forall j \in \{1, 2, 3\}, \{\mathbf{V}'_1, \mathbf{V}'_2, \mathbf{V}'_3\} \in \Pi' \text{ ssi : } \rho_j = \eta'\theta_j + \mu'\Phi_j + \varrho' \quad (3.9)$$

→  $\hat{\mathbf{V}}$  appartient à  $\Pi'$  et vérifie donc l'équation :  $\hat{\Phi} = \eta'\hat{\theta} + \mu'\hat{\rho} + \varrho'$ . Les variables  $\hat{\theta}$  et  $\hat{\rho}$  sont fournies par le point 2D : un changement de repères de  $\mathfrak{R}_P$  dans  $\mathfrak{R}_I$  puis une transformation des coordonnées cartésiennes en les coordonnées polaires. Les paramètres  $(\eta', \mu', \varrho')^\top$  de  $\Pi'$  étant calculés,  $\hat{\Phi}$  est estimé ;

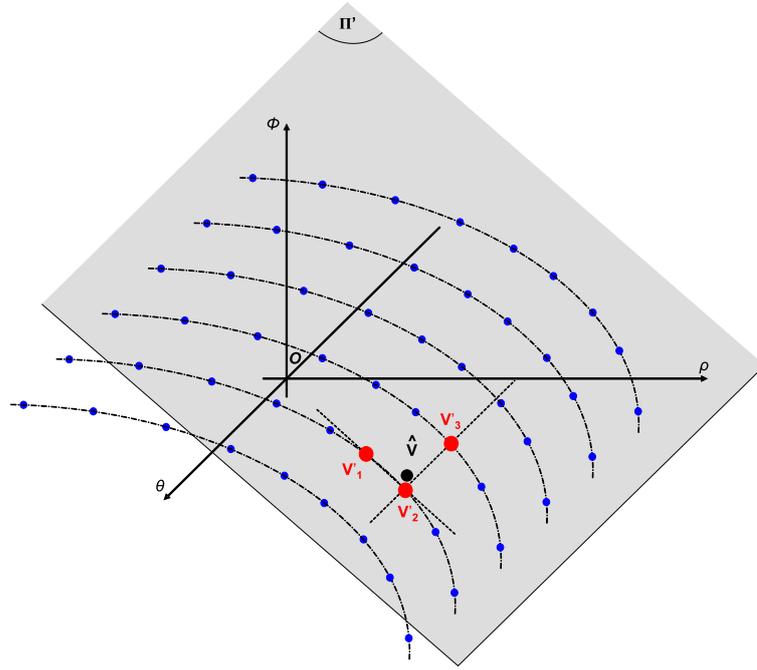


FIG. 3.7 – Dans l'espace des variables  $(\theta, \Phi, \rho)$ , le triplet de points  $\{\mathbf{V}'_1, \mathbf{V}'_2, \mathbf{V}'_3\}$  définit un plan  $\Pi'$ .  $\hat{\mathbf{V}}$  appartient à  $\Pi'$  et vérifie son équation, ce qui permet de déterminer la variable  $\hat{\Phi}$  du vecteur directeur  $\hat{\mathbf{d}}$  de la ligne de vue 3D  $\hat{\mathcal{L}}$ .

→ Connaissant  $\hat{\Phi}$  et  $\hat{\theta}$ , le vecteur directeur  $\hat{\mathbf{d}}$  de la ligne de vue associée au point 2D défini dans  $\mathfrak{R}_M$  se déduit simplement :

$$\hat{\mathbf{d}} = \begin{pmatrix} \sin \hat{\Phi} \cos \hat{\theta} \\ \sin \hat{\Phi} \sin \hat{\theta} \\ \cos \hat{\Phi} \end{pmatrix} \quad (3.10)$$

→ La ligne de vue 3D définie  $\mathfrak{R}_M$  d'origine  $\mathbf{O}_M$ , l'origine du référentiel, se déduit alors :

$$\hat{\mathcal{L}} : \left\{ \hat{\mathcal{L}} = \mathbf{O}_M + \lambda \hat{\mathbf{d}}, \lambda \in \mathbb{R}_+ \right\} \quad (3.11)$$

### 3.3 Conception d'une mire adaptée

La procédure d'étalonnage définit la transformation entre l'espace 3D et le plan image pour un nombre fini de correspondances. La précision de l'étalonnage est donc fonction du nombre de correspondances établies. Elle dépend également de la répartition des points d'intérêt sur la mire 3D. Nous proposons dans les paragraphes suivants une procédure pour le pré-dimensionnement de la mire d'étalonnage afin de :

- déterminer le nombre minimum de correspondances espace 3D-plan image nécessaires pour une précision suffisante de l'étalonnage ;
- définir une répartition des points d'intérêt 3D qui soit la plus pertinente pour le capteur.

Cette étude, réalisée au préalable, est décomposée en deux étapes successives : une première consiste en une analyse du nombre de points d'intérêt par cercle directeur. Une seconde s'intéresse à la distribution verticale des cercles sur le cylindre d'étalonnage.

#### 3.3.1 Nombre de points d'intérêt par cercle directeur

Le premier élément qui définit la précision de l'étalonnage est la densité horizontale du maillage, c'est-à-dire le nombre de points d'intérêt par cercle directeur.

Une première méthode consisterait à multiplier physiquement les points d'intérêt sur la mire afin d'atteindre une précision d'étalonnage requise. Cette solution n'est pas envisageable matériellement car elle conduirait à un outil d'étalonnage irréalisable : alimentation des cartes électroniques, câblage des LEDs, etc.

Une alternative basée sur une approche mathématique, utilise les données existantes pour re-crée des couples point 3D-point 2D supplémentaires. Cette solution permet de s'affranchir des contraintes liées à la structure de la mire et elle s'appuie principalement sur :

- l'estimation de l'équation des projections des cercles de mire dans l'image omnidirectionnelle ;
- un sur-échantillonnage iso-angulaire de ces projections ;

Soit  $C$  un cercle directeur du cylindre d'étalonnage décrit par un ensemble de points d'intérêt 3D de coordonnées supposées connues dans  $\mathfrak{R}_M$ . Lorsque le capteur est idéalement placé par rapport à la mire, c'est-à-dire que les axes optiques du capteur et de révolution du cylindre sont parfaitement confondus, la projection des points 3D décrit également un cercle  $c$  dans l'image omnidirectionnelle dont le centre est l'origine du référentiel  $\mathfrak{R}_I$  [32].

Néanmoins, l'hypothèse d'une colinéarité parfaite de l'axe optique du capteur et de l'axe de révolution du cylindre est difficilement réalisable. Le cercle directeur  $\mathbf{C}$  ne se projette alors plus en un cercle  $\mathbf{c}$  mais en une ellipse  $\mathbf{e}$  dans l'image omnidirectionnelle [32].

L'estimation de l'équation de l'ellipse peut être vue comme l'estimation de l'équation d'une conique dont la nature (parabole, hyperbole, ellipse, etc.) sera déterminée a posteriori [173], [46], [47]. Zhang [173] et Barreto & al. [10] proposent une méthode d'estimation facilement implémentable et qui requiert la connaissance d'au minimum cinq points. Néanmoins, cette méthode est fortement contrainte par le fait que les points doivent décrire « au mieux » l'intégralité de la conique pour que l'estimation réalisée soit correcte. Par conséquent les points 3D sont donc répartis sur l'intégralité du cercle de manière iso-angulaire. Afin de rendre l'estimation robuste aux points aberrants, un nombre de huit points a été choisi pour définir un cercle directeur.

Soit  $\omega$  une conique paramétrée par  $(a, b, c, d, e, f)^\top$ . Elle est définie par la projection des points 3D  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_i, \dots, \mathbf{Q}_8\}$  en les points 2D  $\{\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_8\}$  de sorte que les coordonnées du point  $\mathbf{q}_i$  soient données par :  $\mathbf{q}_i = (x_i, y_i)^\top$ . Idéalement, la conique  $\omega$  et les  $\{\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_8\}$  vérifient l'équation 3.12 :

$$\forall i \in \{1, \dots, 8\}, \{\mathbf{q}_1, \dots, \mathbf{q}_i, \dots, \mathbf{q}_8\} \in \omega : \quad (3.12)$$

$$\overbrace{\begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 & 2x_1 & 2y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_i^2 & 2x_iy_i & y_i^2 & 2x_i & 2y_i & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_8^2 & 2x_8y_8 & y_8^2 & 2x_8 & 2y_8 & 1 \end{pmatrix}}^{\mathbf{A}_{8 \times 6}} \overbrace{\begin{pmatrix} a \\ b \\ c \\ d \\ e \\ f \end{pmatrix}}^{\omega_{6 \times 1}} = \mathbf{0}_{I \times 1}$$

L'équation 3.12 retranscrit un cas idéal, celui pour lequel les données vérifient parfaitement l'équation de la conique. Ce cas est difficilement vérifiable en pratique car les données sont très souvent bruitées ce qui conduit à :  $\mathbf{A}\omega \neq 0$ . L'objectif est donc l'estimation d'un vecteur de paramètres  $\omega$  minimisant une fonction coût  $J(\omega)$  telle que la conique s'ajuste aux données. Zhang [173] et Barreto & Araujo [10] proposent une minimisation sous contrainte d'une fonction coût représentant l'erreur quadratique. Le vecteur  $\omega^*$  est alors défini comme le vecteur propre normé de  $\mathbf{A}^\top \mathbf{A}$  correspondant à la plus petite des valeurs propres.

$$\begin{cases} \boldsymbol{\omega}^* = \min_{a,b,c,d,e,f} J(\boldsymbol{\omega}) \\ J(\boldsymbol{\omega}) = \boldsymbol{\omega}^\top \mathbf{A}^\top \mathbf{A} \boldsymbol{\omega} \\ \boldsymbol{\omega}^\top \boldsymbol{\omega} = 1 \end{cases} \quad (3.13)$$

La connaissance du modèle analytique de l'ellipse permet de sur-échantillonner cette dernière iso-angulairement afin de re-cr  er des couples point 3D-point 2D suppl  mentaires. Ceci permet donc de pallier les inconv  nients li  s    la structure de la mire et plus particuli  rement au type de marqueurs choisis.

Si  $\Delta\theta$  d  finit une variation angulaire, le cercle  $\mathbf{c}$  peut   tre r  -  chantillonn   en un ensemble de points  $\{\mathbf{q}_1, \dots, \mathbf{q}_j, \dots, \mathbf{q}_J\}$  tel que les  $j^{\text{i  me}}$  et  $(j+1)^{\text{i  me}}$  points 2D appartenant     $\mathbf{c}$ ,  $\mathbf{q}_j = (\theta_j, \rho_j)^\top$  et  $\mathbf{q}_{j+1} = (\theta_{j+1}, \rho_{j+1})^\top$ , soient d  finis par la relation suivante :

$$\forall j \in \{1, \dots, J\}, (\mathbf{q}_j, \mathbf{q}_{j+1}) \in \mathbf{c} : \begin{cases} \theta_{j+1} = \theta_j + \Delta\theta \\ \rho_{j+1} = \rho_j \end{cases} \quad (3.14)$$

Aux  $\{\mathbf{q}_1, \dots, \mathbf{q}_j, \dots, \mathbf{q}_J\}$  nouvellement cr  es sont associ  s des points d'int  r  t 3D cocycliques  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_j, \dots, \mathbf{Q}_J\}$  pour le cercle  $\mathbf{C}$  tels que les coordonn  es des points d'int  r  t  $\mathbf{Q}_i$  et  $\mathbf{Q}_j$  soient d  finies par :

$$\forall j \in \{1, \dots, J\}, \forall i \in \{1, \dots, 8\}, (\mathbf{Q}_j, \mathbf{Q}_i) \in \mathbf{C} : \begin{cases} \theta_j \neq \theta_i \\ \Phi_j = \Phi_i \\ \Gamma_j = \Gamma_i \end{cases} \quad (3.15)$$

Le param  tre de sur-  chantillonnage iso-angulaire  $\Delta\theta$  permet de re-cr  er des couples point 3D-point 2D suppl  mentaires et est fix   de mani  re empirique. Cette valeur doit   tre ajust  e en fonction du degr   de pr  cision requis pour l'  talonnage.

### 3.3.2 Distribution verticale des cercles directeurs sur le cylindre

Le second   l  ment qui d  finit la pr  cision de l'  talonnage est li      la densit   verticale du maillage, c'est-  -dire au nombre de cercles directeurs et    leurs positions sur le cylindre d'  talonnage. Nous proposons une m  thode structur  e en deux   tapes successives qui

permet d'estimer le nombre de cercles nécessaires et leur distribution pour une précision requise pour l'étalonnage :

→ une première réalise un maillage tel que la répartition des cercles directeurs  $\{C_1, \dots, C_m, \dots\}$   $\{\dots, C_M\}$  soit uniforme sur la mire, c'est-à-dire que deux cercles directeurs contigus  $C_m$  et  $C_{m+1}$  possèdent un écartement vertical  $\Delta z$  constant. L'écartement  $\Delta z$  est également choisi faible afin d'obtenir un maillage dense. Cette distribution uniforme en 3D conduit, de part la résolution spécifique du capteur catadioptrique, à une répartition des projections des cercles qui est non-uniforme dans l'image omnidirectionnelle (cf. figure 3.8) ;

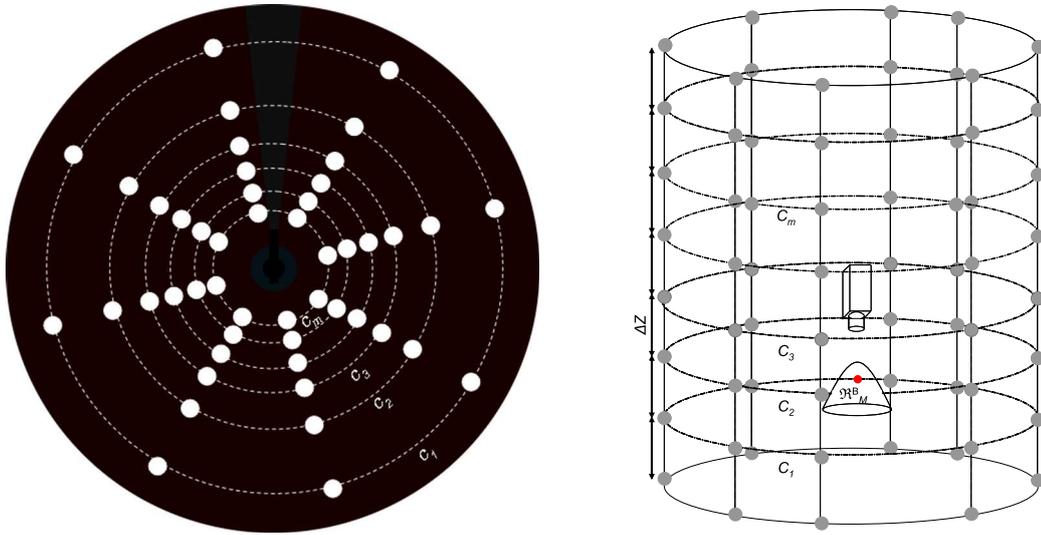


FIG. 3.8 – Répartition uniforme des cercles directeurs en 3D. Les marqueurs lumineux sont répartis uniformément suivant les cercles directeurs  $\{C_1, \dots, C_m, \dots, C_M\}$  et les génératrices du cylindre d'étalonnage. Deux cercles contigus  $C_m$  et  $C_{m+1}$  ont un écartement vertical  $\Delta z$  constant et faible. Lorsque l'axe optique du capteur et l'axe de révolution du cylindre sont confondus, les marqueurs se projettent en des cercles  $\{c_1, \dots, c_m, \dots, c_M\}$  dans l'image omnidirectionnelle. La répartition des cercles en 2D est non-uniforme (image de synthèse dont les ovoïdes sont volontairement représentés par des disques blancs de grande taille afin de faciliter la compréhension des schémas).

→ une seconde définit une nouvelle répartition des cercles directeurs 3D  $\{C_1, \dots, C_n, \dots\}$   $\{\dots, C_N\}$  de sorte que les projections de ces cercles  $\{c_1, \dots, c_n, \dots, c_N\}$  soient uniformément distribuées dans l'image omnidirectionnelle. En d'autres termes, les cercles  $\{C_1, \dots, C_n, \dots, C_N\}$  se projettent en des cercles  $\{c_1, \dots, c_n, \dots, c_N\}$  concentriques dans l'image omnidirectionnelle (axe de révolution de la mire et axe optique du capteur colinéaire) et l'écartement  $\Delta \rho$  entre deux cercles contigus  $c_n$  et  $c_{n+1}$  est constant (cf. figure 3.9).

Pour déterminer cette nouvelle distribution des cercles en 3D, un rayon  $\ell$  de l'image omnidirectionnelle est discrétisé en un échantillon de points 2D  $\{\mathbf{q}_1, \dots, \mathbf{q}_n, \dots, \mathbf{q}_N\}$  de sorte que l'écartement  $\Delta\rho$  entre deux points 2D contigus  $\mathbf{q}_n$  et  $\mathbf{q}_{n+1}$  soit constant. Ainsi deux points 2D  $\mathbf{q}_n$  et  $\mathbf{q}_{n+1}$  sont définis par :

$$\forall n \in \{1, \dots, N\}, \{\mathbf{q}_n, \mathbf{q}_{n+1}\} \in \ell : \begin{cases} \theta_{n+1} = \theta_n \\ \rho_{n+1} = \rho_n + \Delta\rho \end{cases} \quad (3.16)$$

Les vecteurs directeurs  $\{\mathbf{d}_1, \dots, \mathbf{d}_n, \dots, \mathbf{d}_N\}$  des lignes de vue 3D  $\{\mathcal{L}_1, \dots, \mathcal{L}_n, \dots, \mathcal{L}_N\}$  correspondant aux points 2D sont déduits des données du maillage précédemment réalisé (cf. distribution uniforme des cercles en 3D). Ces lignes de vue 3D sont définies dans le référentiel capteur  $\mathfrak{R}_M$  d'origine  $\mathbf{O}_M$ . Pour un angle d'azimut nul, une  $n^{\text{ième}}$  ligne de vue 3D  $\mathcal{L}_n$  est définie par :

$$\forall n \in \{1, \dots, N\}, \mathcal{L}_n : \left\{ \mathcal{L}_n = \mathbf{O}_M + \lambda \begin{pmatrix} \sin \Phi_n \\ 0 \\ \cos \Phi_n \end{pmatrix}, \lambda \in \mathbb{R}_+ \right\} \quad (3.17)$$

Ces lignes de vue 3D intersectent le cylindre en de nouvelles positions 3D  $\{\mathbf{Q}_1, \dots, \mathbf{Q}_n, \dots, \mathbf{Q}_N\}$ . A ces points 3D sont associés des cercles directeurs  $\{\mathcal{C}_1, \dots, \mathcal{C}_n, \dots, \mathcal{C}_N\}$ .

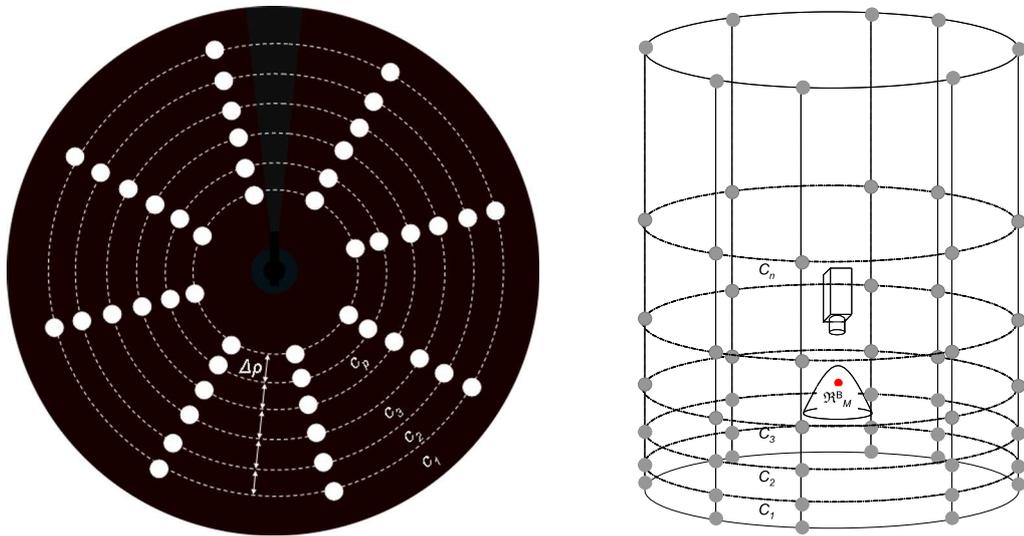


FIG. 3.9 – Distribution adaptée des cercles directeurs en 3D. Les cercles en 2D sont répartis uniformément dans l'image omnidirectionnelle : deux cercles contigus  $c_n$  et  $c_{n+1}$  ont un écartement  $\Delta\rho$  constant. Les cercles directeurs  $\{\mathcal{C}_1, \dots, \mathcal{C}_n, \dots, \mathcal{C}_N\}$  de la mire d'étalonnage sont alors répartis de manière non-uniforme en 3D (image de synthèse dont les ovoïdes sont volontairement représentés par des disques blancs de grande taille afin de faciliter la compréhension des schémas).

## 3.4 Evaluation en simulation et sur le système réel

Les résultats présentés font référence à l'étalonnage non-paramétrique d'un capteur seul et du système de stéréovision omnidirectionnelle dans sa globalité.

Une première partie présente les résultats obtenus pour une distribution régulière des points d'intérêt 3D, en simulation et pour le système réel. Une seconde partie présente des résultats de simulation se rapportant à une distribution des points d'intérêt adaptée au capteur.

### 3.4.1 Protocoles expérimentaux et méthodes de validation

#### 3.4.1.1 Environnement de simulation pour une distribution régulière des points d'intérêt

L'environnement de simulation s'apparente à celui présenté au paragraphe 2.6.1 :

- la mire est un cylindre de 600mm de diamètre ;
- la distribution des points d'intérêt en 3D est une grille régulière : les LEDs sont réparties suivant les cercles directeurs et les génératrices du cylindre. Deux cercles contigus ont un écartement vertical constant :  $\Delta z = 5\text{mm}$  et un cercle est défini par huit points 3D cocycliques ;
- le pas angulaire de sur-échantillonnage  $\Delta\theta$  est égal à :  $\Delta\theta = 1^\circ$  ;
- les cercles directeurs sont allumés successivement et pour chaque cercle allumé, l'image est acquise simultanément ;
- un fichier texte contient les coordonnées des point d'intérêt 3D définies dans le repère associé au capteur  $\mathfrak{R}_M$ .

#### 3.4.1.2 Environnement du système réel pour une distribution régulière des points d'intérêt

Pour des raisons d'intégration matérielle, l'écartement vertical entre deux cercles directeurs contigus est plus important que dans l'étude réalisée en simulation ( $\Delta z = 25\text{mm}$ ) pour une distribution régulière des points d'intérêt. Un cercle directeur est décrit par huit LEDs réparties iso-angulairement sur ce cercle et le pas de sur-échantillonnage des projections des cercles 3D est fixé à  $1^\circ$ . Le premier cercle est positionné en la coordonnée  $z = 275\text{mm}$  et le dernier en la coordonnée  $z = 1375\text{mm}$ .

### 3.4.1.3 Environnement de simulation pour une distribution adaptée des points d'intérêt

La mire possède un diamètre identique (600mm) et le pas angulaire de sur-échantillonnage  $\Delta\theta$  est égal à  $1^\circ$ . Néanmoins, pour obtenir cette répartition adaptée des points d'intérêt sur la mire, un protocole spécifique est mis en place :

- le rayon de l'image omnidirectionnelle d'angle d'azimut nul est discrétisé tel que l'écartement  $\Delta\rho$  entre deux points 2D contigus soit constant. Le nombre de points 2D échantillonnés définit le nombre de cercles sur la mire ;
- les points 2D sont rétro-projetés sur le cylindre d'étalonnage grâce aux données d'étalonnage obtenues pour une distribution régulière des points d'intérêt 3D ;
- l'intersection des lignes de vue avec le cylindre d'équation connue définit les nouvelles positions verticales des cercles sur la mire ;
- pour ces nouvelles positions des cercles directeurs, une procédure d'étalonnage non-paramétrique est mise en place.

### 3.4.1.4 Erreurs de projection

Il s'agit d'une mesure de distance euclidienne en pixel entre des primitives 2D dont les coordonnées sont estimées dans un environnement de simulation et des primitives 2D dont les projections sont estimées par les données d'étalonnage.

Le jeu d'évaluation est constitué de 29 points test 3D matérialisés par des sphères de couleur blanche qui se projettent sous forme d'ovoïdes dans l'image omnidirectionnelle. Les points test choisis, de coordonnées parfaitement connues, sont répartis de manière aléatoire dans un environnement 3D de sorte à couvrir l'intégralité du champ d'observation du capteur. Au final, les projections des points test 3D sont déterminées par une estimation de barycentres telle que présentée au paragraphe 2.5.2.

### 3.4.1.5 Erreurs de rétro-projection et de reconstruction

**Erreur de rétro-projection :** Afin d'estimer une erreur de rétro-projection par capteur, des cercles directeurs d'un cylindre sont projetés dans l'image omnidirectionnelle (cf. figure 3.10a). Chaque cercle est défini par un échantillon de points 3D répartis iso-angulairement. Une attention particulière est portée afin que ces points d'évaluation n'appartiennent pas aux données d'étalonnage. L'échantillon de points 3D projeté est ensuite

rétro-projeté sur le cylindre, ce qui définit de nouveaux points 3D. L'erreur de rétro-projection est alors définie comme la distance euclidienne entre l'échantillon de points 3D initial et les points 3D estimés. Les caractéristiques techniques du mode opératoire sont les suivantes :

- un cylindre de révolution de diamètre 600mm ;
- trois cercles arbitrairement positionnés en :  $z = 282\text{mm}$ ,  $z = 702\text{mm}$  et  $z = 1302\text{mm}$  ;
- chaque cercle est défini par 35 points 3D répartis iso-angulairement sur ce cercle.

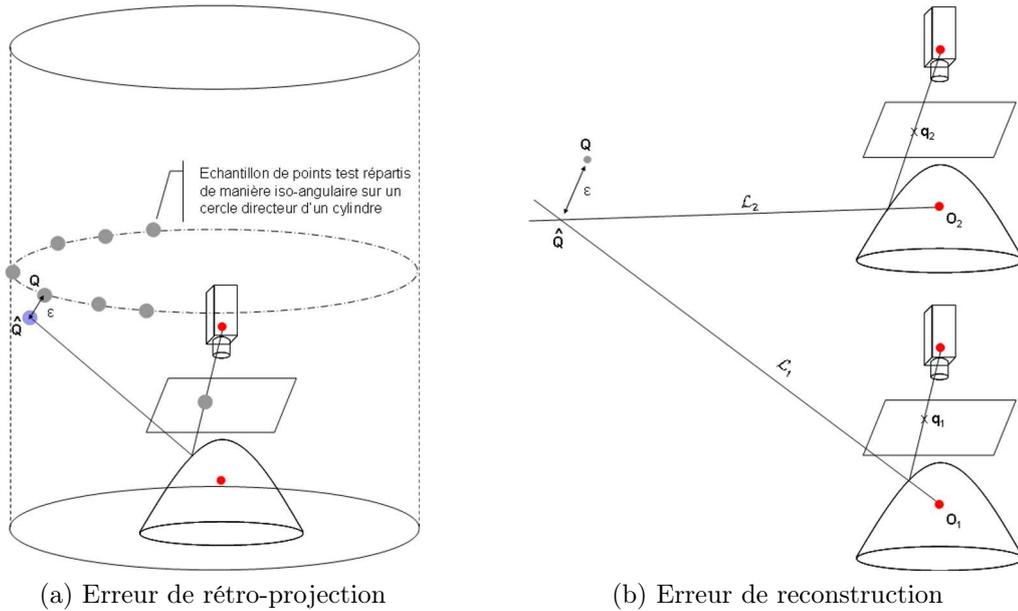


FIG. 3.10 – Schémas illustrant les erreurs de rétro-projection et de reconstruction

**Erreur de reconstruction :** Afin d'évaluer le capteur stéréoscopique dans son intégralité, une erreur de reconstruction est estimée pour les deux capteurs (cf. figure 3.10b). La procédure est réalisée en simulation à l'aide du logiciel POV-Ray. Un échantillon de 24 points de coordonnées 3D connues est réparti aléatoirement dans une scène 3D définie comme un cube de dimensions  $5000\text{mm} \times 5000\text{mm} \times 5000\text{mm}$ . Ces points test se projettent respectivement dans les images des capteurs bas et haut en des ovoïdes et chaque ovoïde est assimilé à une projection ponctuelle par une estimation de barycentre. A ces points 2D sont associés des lignes de vue dans l'espace 3D. Une triangulation des lignes de vue par une méthode du « *middle point* » permet de déterminer les coordonnées du point 3D estimé.

Soit  $\mathbf{Q}$  un point de l'espace 3D de coordonnées parfaitement connues et  $\hat{\mathbf{Q}}$  l'estimé de  $\mathbf{Q}$ .  $\mathbf{Q}$  se projette dans les plans images des capteurs bas et haut en  $\mathbf{q}_1$  et  $\mathbf{q}_2$ . Sont associées à

ces points 2D des lignes de vue dans l'espace 3D, notées respectivement  $\mathcal{L}_1$  et  $\mathcal{L}_2$ . Notons également  $\mathbf{O}_1$  et  $\mathbf{O}_2$  les foyers respectifs des miroirs bas et haut du système de stéréovision et  $\mathbf{O}_1\mathbf{O}_2$  le vecteur entre ces deux points. Le point  $\hat{\mathbf{Q}}$  est alors défini comme le barycentre du segment  $\Lambda$  minimisant l'écart entre  $\mathcal{L}_1$  et  $\mathcal{L}_2$ , tel que :

$$\hat{\mathbf{Q}} \in \Lambda \quad \Lambda = \min[\mathcal{L}_1, \mathcal{L}_2], \quad \Lambda = \mathbf{O}_1\mathbf{O}_2 \frac{\mathcal{L}_1 \wedge \mathcal{L}_2}{\|\mathcal{L}_1 \wedge \mathcal{L}_2\|} \quad (3.18)$$

## 3.4.2 Résultats

### 3.4.2.1 Distribution régulière des point d'intérêt 3D en simulation

Le maillage réalisé est défini par  $221 \times 8 = 1768$  correspondances obtenues par la mire auquel s'ajoute  $221 \times 360 = 79560$  correspondances re-crées à partir des projections des cercles directeurs 3D dans l'image omnidirectionnelle. La figure 3.11 présente une illustration du maillage obtenu pour une distribution régulière des points d'intérêt 3D.

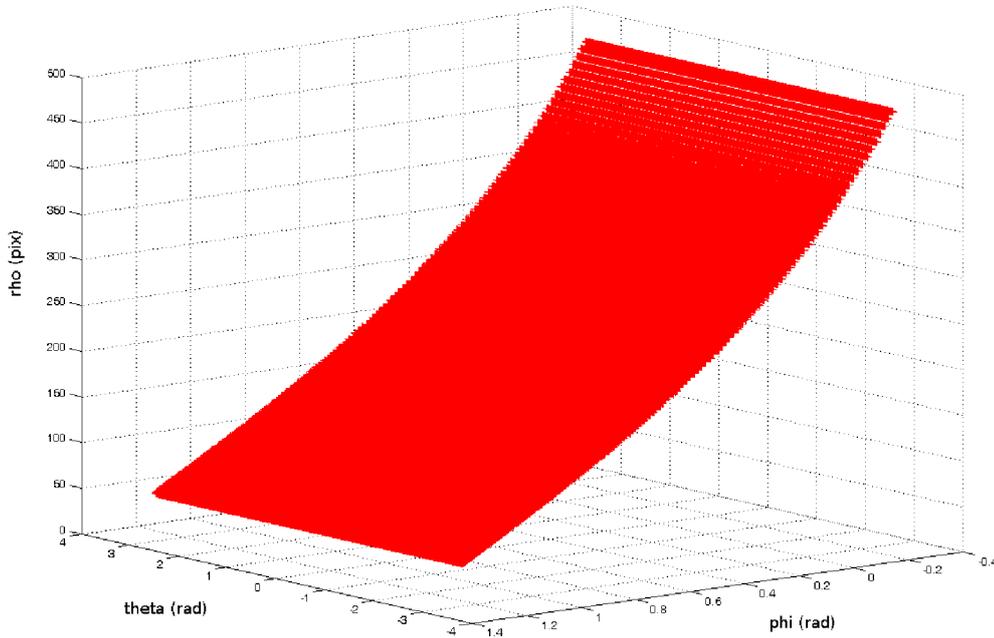


FIG. 3.11 – Maillage pour une distribution régulière des points d'intérêt 3D obtenu pour un environnement simulé. Les points rouges modélisent les points  $\mathbf{V}$  du maillage.

La figure 3.13 illustre les erreurs de rétro-projection obtenues pour chacun des cercles projetés et pour chacun des 35 points test. L'erreur moyenne globale obtenue est alors de :  $\bar{\epsilon} = 1.77 \pm 1.27\text{mm}$ . La figure 3.12 est une représentation du plan image où figurent les projections des points appartenant aux trois cercles 3D.

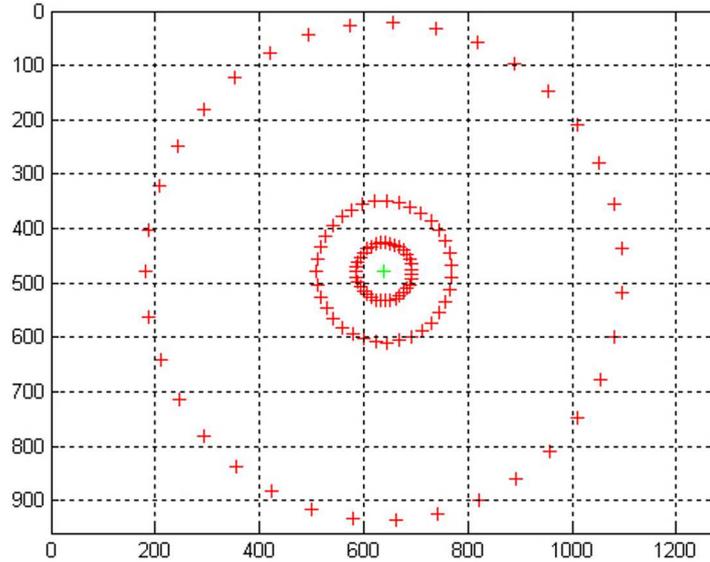


FIG. 3.12 – Projection des trois cercles 3D dans l'image omnidirectionnelle. Ces cercles appartiennent à un cylindre de révolution virtuel. Les points 2D sont ensuite rétro-projetés sur le cylindre afin de déterminer l'erreur de rétro-projection.

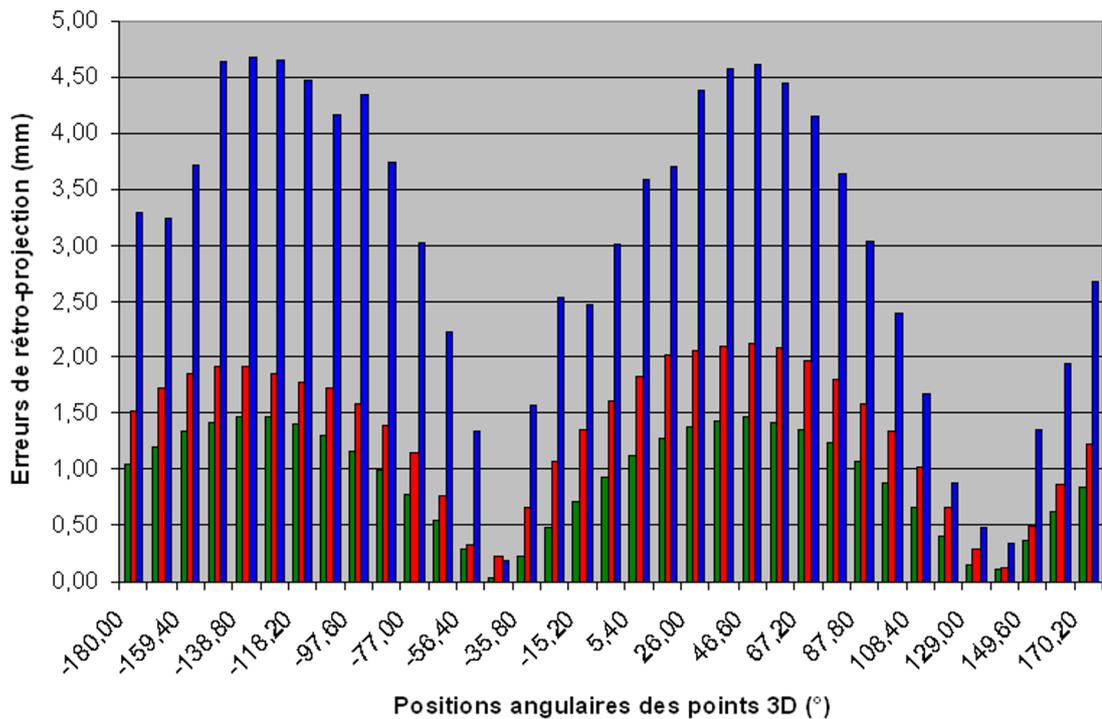


FIG. 3.13 – Erreur de rétro-projection pour une répartition uniforme des points d'intérêt 3D (données d'étalonnage obtenues en simulation). Ces graphiques illustrent les erreurs de rétro-projection calculées pour chacun des trois cercles 3D et pour les 35 points projetés dans l'image omnidirectionnelle. Le graphique de couleur verte correspond au cercle 3D en  $z = 282\text{mm}$ , celui en rouge au cercle 3D positionné en  $z = 702\text{mm}$  et celui en bleu correspond au cercle en  $z = 1302\text{mm}$

La figure 3.14 présente les erreurs de reconstruction en millimètre pour les 24 points test de l'échantillon. Une erreur moyenne de reconstruction est ensuite calculée et est égale à :  $\bar{\epsilon} = 14.93 \pm 12.45\text{mm}$ . Il est également intéressant de pondérer ces distances par rapport aux distances aux points test. Il s'agit de la figure 3.15 qui présente les erreurs de reconstruction en pourcentage des distances aux points test, ce qui conduit à une erreur moyenne en pourcent de :  $\bar{\epsilon}_{\%} = 0.46$ .

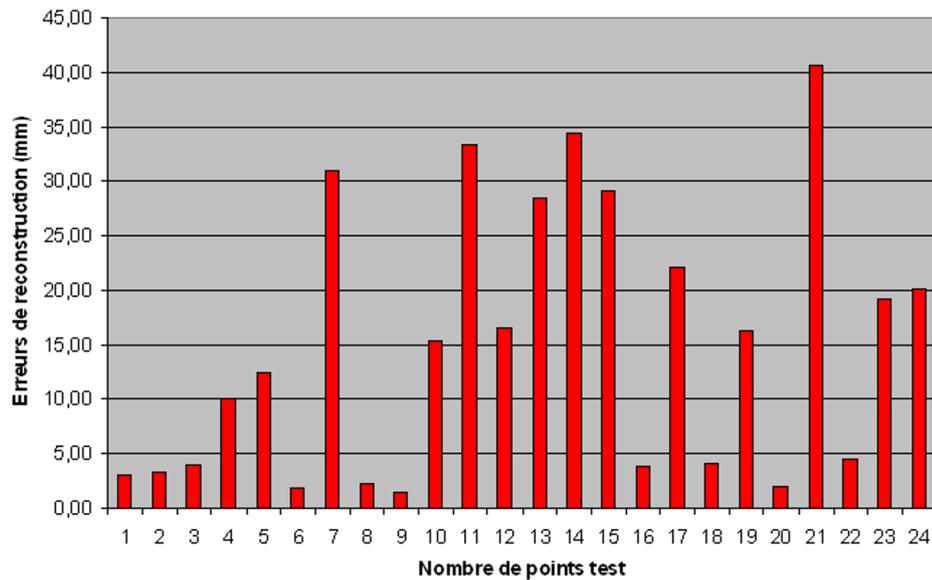


FIG. 3.14 – Erreur de reconstruction en millimètre estimée par intersection des lignes de vue des deux capteurs (données d'étalonnage obtenues en simulation) pour une distribution régulière des points d'intérêt 3D. Ces graphiques illustrent les erreurs de reconstruction estimées pour un échantillon de 24 points test répartis aléatoirement dans une scène 3D.

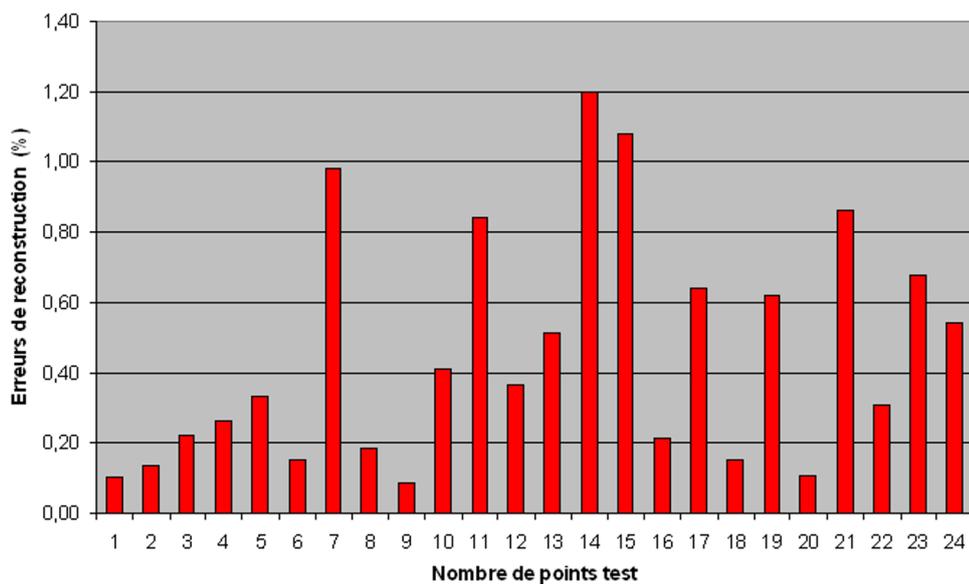


FIG. 3.15 – Erreur de reconstruction estimée par intersection des lignes de vue des deux capteurs (données d'étalonnage obtenues en simulation) et exprimée en pourcentage des distances aux points test. Il s'agit de résultats obtenus pour une distribution régulière des points d'intérêt 3D

### 3.4.2.2 Distribution régulière des point d'intérêt 3D pour le système réel

360 points réels définissent la mire d'étalonnage auxquels s'ajoutent 16200 correspondances générées par le sur-échantillonnage des projections des cercles 3D. Au final, le maillage illustré sur la figure 3.16 est donc constitué de 16560 couples point 3D-point 2D.

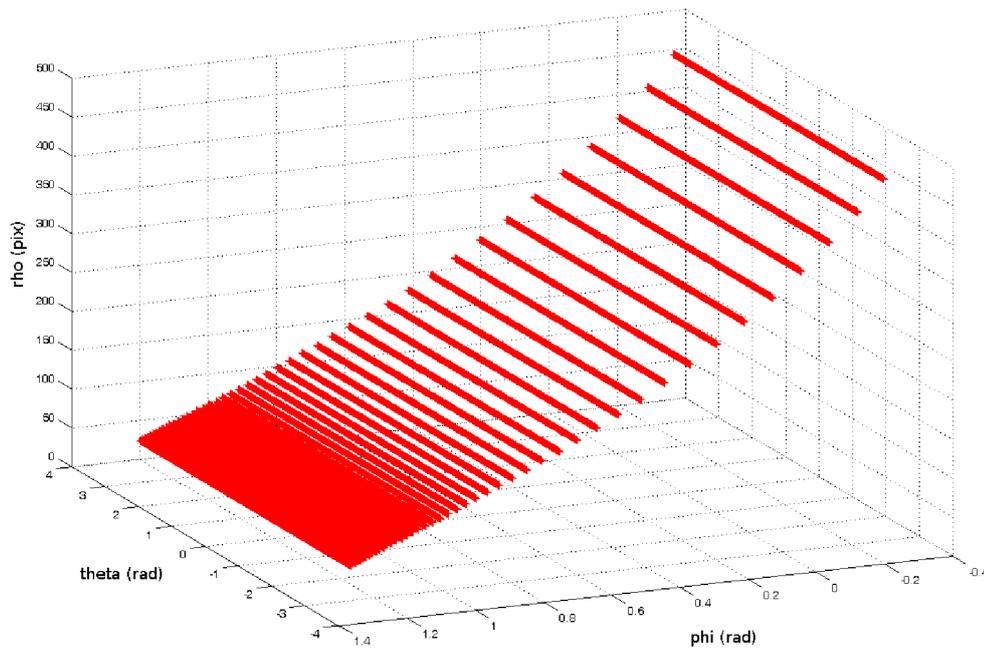


FIG. 3.16 – Maillage pour une distribution régulière des points d'intérêt 3D obtenu pour le système réel. Les points rouges modélisent les points  $\mathbf{V}$  du maillage (données obtenues pour le système mire-capteur réel).

La figure 3.17 illustre les erreurs de rétro-projection pour chaque cercle et pour chaque point test. L'erreur de rétro-projection moyenne estimée sur l'ensemble des 105 points test est égale à :  $\bar{\epsilon} = 2.08 \pm 2.01\text{mm}$ .

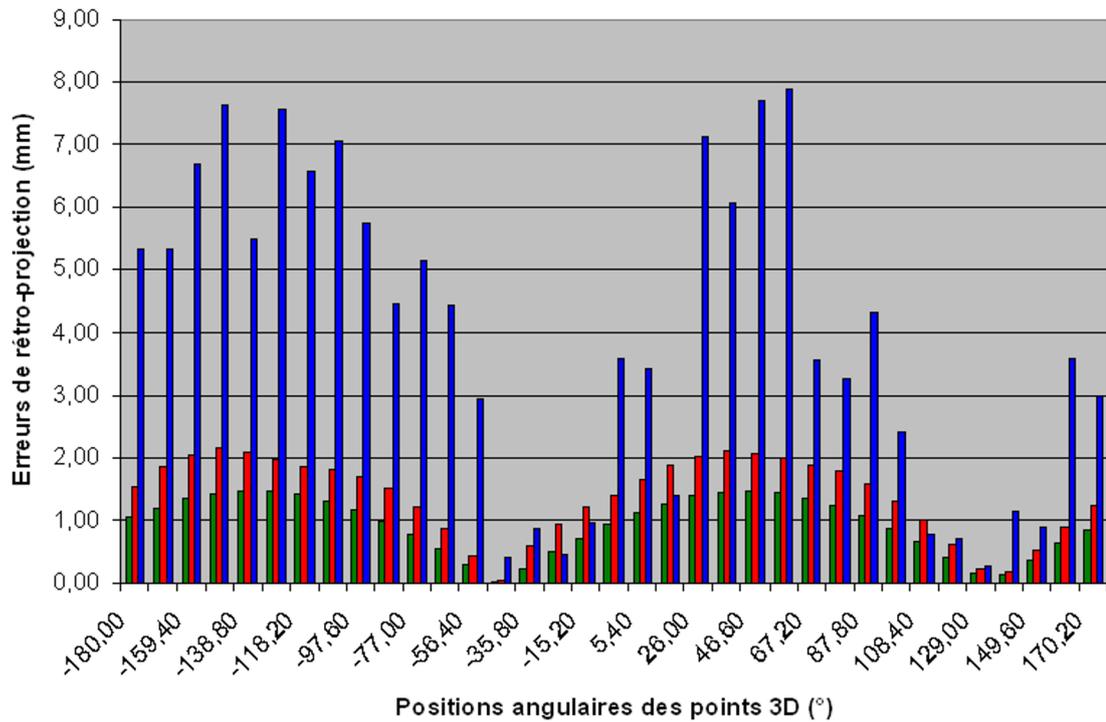


FIG. 3.17 – Erreur de rétro-projection pour une répartition uniforme des points d'intérêt 3D (données d'étalonnage pour le système mire-capteur réel). Ces graphiques illustrent les erreurs de rétro-projection calculées pour chacun des trois cercles 3D et pour les 35 points projetés dans l'image omnidirectionnelle. Le graphique de couleur verte correspond au cercle 3D en  $z = 282\text{mm}$ , celui en rouge au cercle 3D positionné en  $z = 702\text{mm}$  et celui en bleu correspond au cercle en  $z = 1302\text{mm}$

### 3.4.2.3 Distribution adaptée des points d'intérêt en simulation

L'utilisation d'une mire adaptée introduit des nouvelles positions des cercles directeurs 3D et induit un maillage uniforme tel que présenté en la figure 3.18.

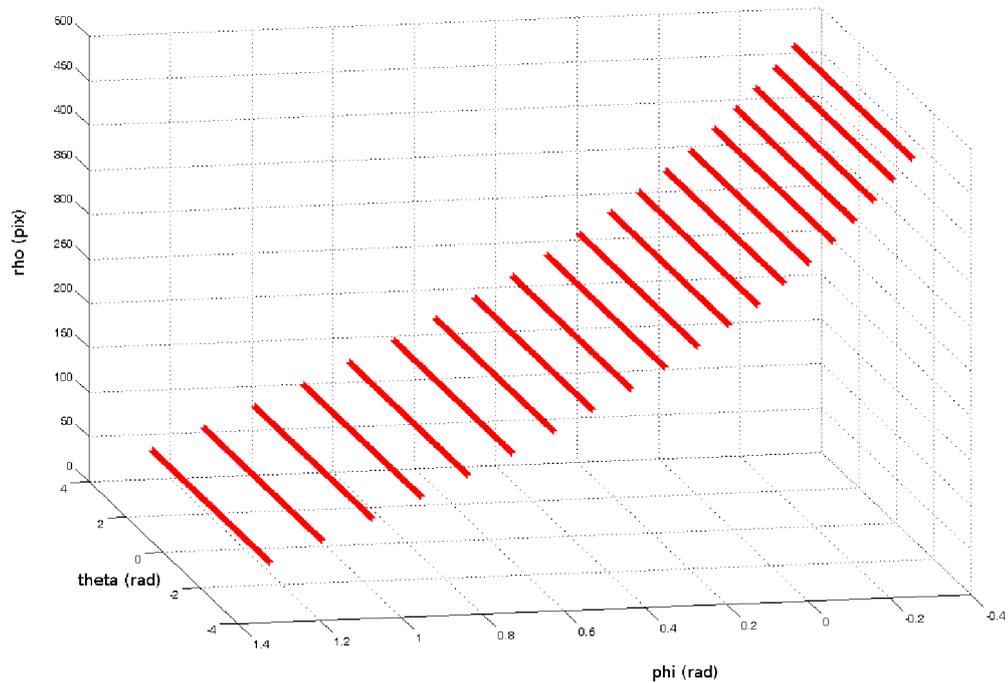


FIG. 3.18 – Vue du maillage obtenu en simulation et correspondant à une répartition adaptée des points d'intérêt 3D sur la mire d'étalonnage. Les points rouges modélisent les points  $\mathbf{V}$  du maillage (données obtenues en simulation).

Une erreur de projection est estimée en faisant varier le nombre et la position des cercles 3D sur la mire. La figure 3.19 et le tableau 3.1 résume les résultats obtenus.

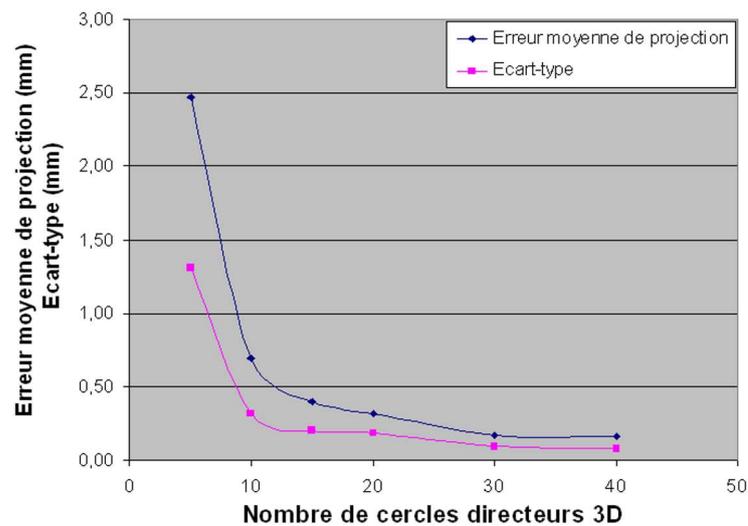


FIG. 3.19 – Evolution de l'erreur moyenne de projection (en pixels) et de l'écart-type associé en fonction du nombre de cercles directeurs 3D sur le cylindre d'étalonnage.

	5	10	15	<b>20</b>	30	40
$\bar{\varepsilon} \pm \sigma(\text{pix})$	$2.47 \pm 1.31$	$0.69 \pm 0.32$	$0.40 \pm 0.20$	<b><math>0.32 \pm 0.19</math></b>	$0.17 \pm 0.10$	$0.16 \pm 0.10$

TAB. 3.1 – Evolution de l’erreur moyenne de projection (en pixels) et de l’écart-type associé en fonction du nombre de cercles 3D sur le cylindre d’étalonnage.

La distribution des points d’intérêt sur la mire, c’est-à-dire leur nombre et leur position verticale sur le cylindre, doit être un compromis entre une précision requise pour l’étalonnage du capteur et la réalisation matérielle de la mire. Ainsi, au vu du tableau 3.1, un nombre de 20 cercles directeurs a été choisi, chacun étant constitué de 8 LEDs répartis iso-angulairement sur le cercle. Ainsi 160 points d’intérêt réels sont répartis sur la mire d’étalonnage. Le pas de sur-échantillonnage des cercles 2D est choisi égal à  $1^\circ$  ce qui conduit à 7200 couples point 3D-point 2D supplémentaires. Au total, 7360 couples point 3D-point 2D définissent le maillage pour une distribution adaptée des cercles sur la mire (cf. figure 3.18).

La figure 3.20 illustre les erreurs de rétro-projection pour chaque cercle et pour chaque point test. L’erreur moyenne sur l’ensemble des données d’évaluation est alors égale à :  $\bar{\varepsilon} = 1.70 \pm 1.17\text{mm}$

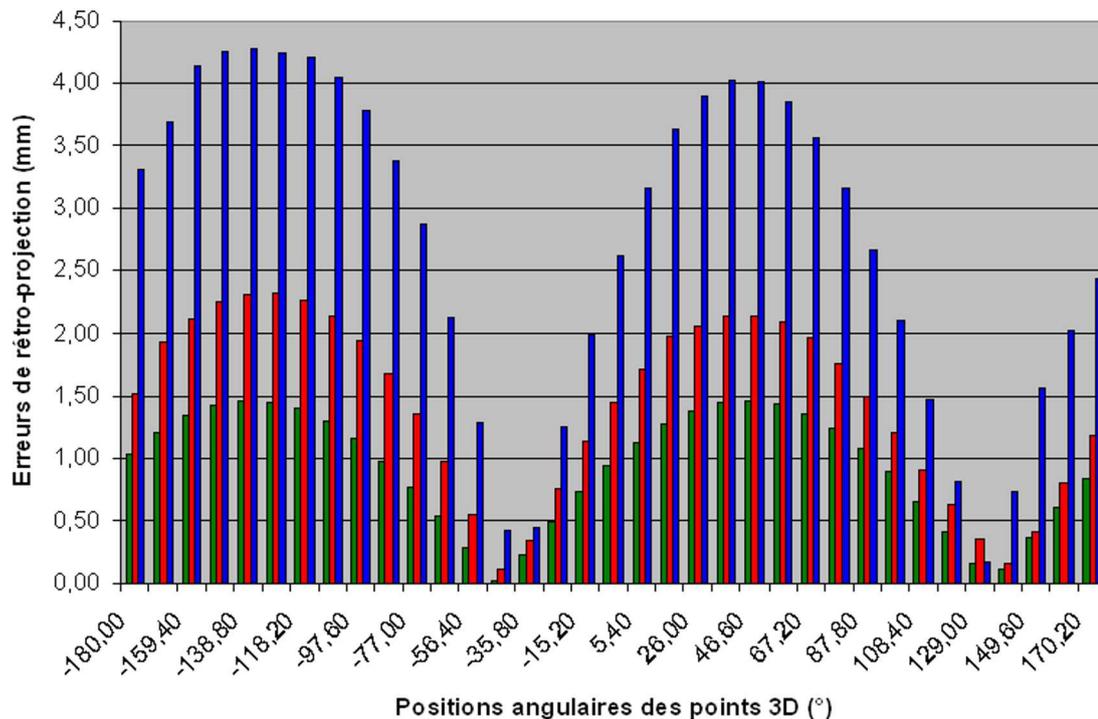


FIG. 3.20 – Erreur de rétro-projection pour une distribution adaptée des points d'intérêt 3D (données d'étalonnage obtenues en simulation). Ces résultats sont obtenus en simulation. Ces graphiques illustrent les erreurs de rétro-projection calculées pour chacun des trois cercles 3D et pour les 35 points projetés dans l'image omnidirectionnelle. Le graphique de couleur verte correspond au cercle 3D en  $z = 282\text{mm}$ , celui en rouge au cercle 3D positionné en  $z = 702\text{mm}$  et celui en bleu correspond au cercle en  $z = 1302\text{mm}$

A l'instar de ce qui a été proposé en simulation pour une distribution régulière des points 3D sur la mire, une erreur de reconstruction est estimée pour le banc stéréoscopique dans sa globalité. La figure 3.21 illustre les erreurs de reconstruction en millimètre. L'erreur moyenne estimée est égale à :  $\bar{\epsilon} = 90.25 \pm 76.08\text{mm}$ . La figure 3.22 montre les erreurs de reconstruction en pourcentage des distances aux points test. L'erreur moyenne en pourcent est alors égale à :  $\bar{\epsilon}_{\%} = 2.74$ .

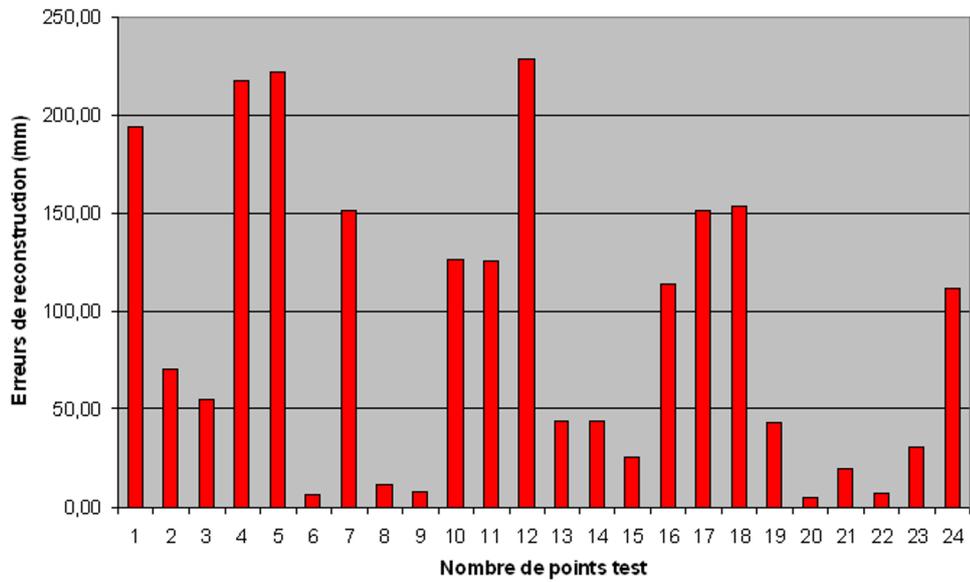


FIG. 3.21 – Erreur de reconstruction en millimètre estimée par intersection des lignes de vue des deux capteurs (données d'étalonnage obtenues en simulation) pour une distribution adaptée des points d'intérêt 3D. Ces graphiques illustrent les erreurs de reconstruction estimées pour un échantillon de 24 points test répartis aléatoirement dans une scène 3D.

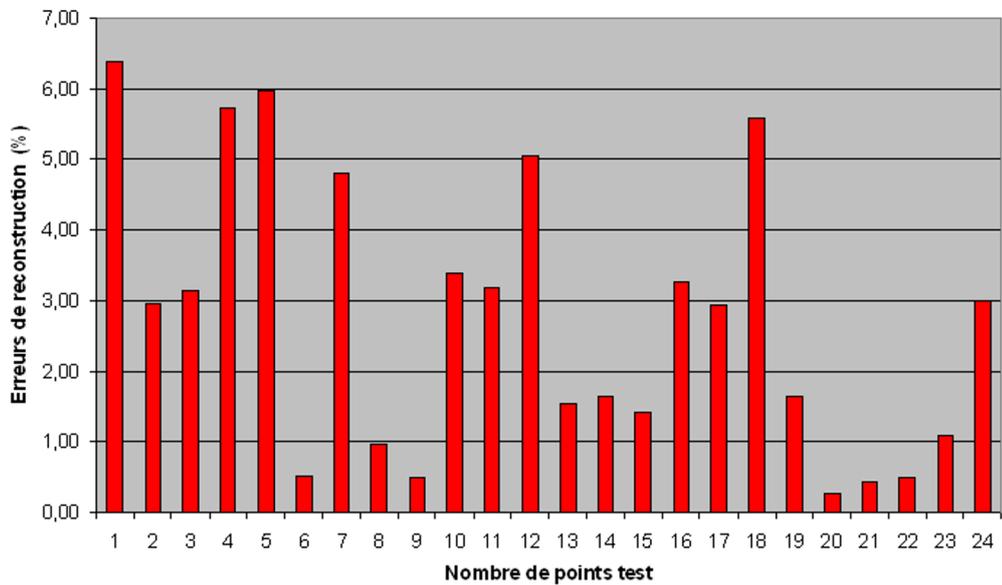


FIG. 3.22 – Erreur de reconstruction estimée par intersection des lignes de vue des deux capteurs (données d'étalonnage obtenues en simulation) et exprimée en pourcentage des distances aux points test. Ces erreurs sont obtenues pour une distribution adaptée des points d'intérêt 3D

## 3.5 Discussion

Les résultats proposés sont principalement fondés sur des mesures d'erreur de rétro-projection. Une première mesure ne s'intéresse qu'à un seul capteur ce qui permet d'évaluer l'étalonnage, c'est-à-dire les processus de projection et de rétro-projection (cf. figure 3.13, figure 3.17 et figure 3.20). Une seconde mesure se rapporte à des erreurs de reconstruction estimées pour le banc de stéréovision. Ceci permet d'évaluer l'étalonnage non-paramétrique dans son ensemble et de vérifier la précision du banc stéréoscopique pour des applications en reconstruction de scènes 3D (cf. figure 3.14, figure 3.15, figure 3.21, figure 3.22). Des erreurs moyennes complètent ces graphiques et permettent d'apprécier globalement la précision des données d'étalonnage.

Pour les données d'étalonnage obtenues pour une distribution régulière des points d'intérêt (en simulation ou pour le système réel, cf. figure 3.13 et figure 3.17), les erreurs de rétro-projection sont faibles au regard des distances aux points d'évaluation. Si l'on rapporte l'erreur à la distance à la cible, on constate que le pourcentage d'erreur moyen calculé sur l'intégralité des 105 points test est faible :  $\bar{\epsilon}_{\%} = 0.21$  pour des données d'étalonnage obtenues en simulation et  $\bar{\epsilon}_{\%} = 0.23$  pour des données obtenues sur le système réel. Si l'on s'intéresse à l'erreur maximale exprimée en pourcent, elle est de :  $\max(\epsilon_{\%}) = 0.35$  pour les données de simulation et de :  $\max(\epsilon_{\%}) = 0.59$  pour les données obtenues sur le système réel. Ces résultats confirment la pertinence de la méthode d'étalonnage non-paramétrique proposée. Ils illustrent également le fait que l'étalonnage est sensiblement moins précis pour le système réel. Ces différences sont principalement dues aux problèmes liés à l'intégration matérielle (bruit numérisation, défaut d'alignement mire-capteur, erreurs sur les coordonnées des points 3D, etc.). Ces différences sont d'autant plus accentuées que la répartition des cercles directeurs sur la mire n'est pas identique ( $\Delta z = 5\text{mm}$  en simulation et  $\Delta z = 25\text{mm}$  pour le cas réel).

La méthode d'étalonnage non-paramétrique induit des erreurs de reconstruction faibles : l'erreur de reconstruction maximale est de :  $\max(\epsilon) = 40.5\text{mm}$  et l'erreur moyenne est de :  $\bar{\epsilon} = 14.9 \pm 12.5\text{mm}$ , ce qui rapporté aux distances aux points test représente une erreur moyenne en pourcent de :  $\bar{\epsilon}_{\%} = 0.46$ . Ces résultats sont particulièrement encourageants. Ils confirment la pertinence de l'étalonnage non-paramétrique.

Même si les erreurs de rétro-projection et reconstruction sont faibles, des distributions régulières telles que celles proposées en simulation ou pour le système réel n'optimisent pas le nombre de points d'intérêt sur la mire et rendent la conception matérielle de cette dernière difficile. Il est donc judicieux et pertinent de proposer une méthode de pré-

dimensionnement de l'outil d'étalonnage afin d'optimiser le nombre et la distribution des points d'intérêt sur la mire.

Pour des données d'étalonnage obtenues pour une distribution adaptée des points d'intérêt (cf. figure 3.20), l'erreur moyenne de rétro-projection en pourcent calculée pour les 105 points test est égale à celle obtenue en simulation pour une distribution régulière des points d'intérêt ( $\bar{\epsilon}_{\%} = 0.21$ ). L'erreur maximale en pourcent est également identique ( $\max(\epsilon_{\%}) = 0.35$ ).

Les données obtenues en simulation ont été réalisées pour un cas spécifique, celui où 20 cercles judicieusement répartis permettent l'étalonnage du capteur. Ce choix a été motivé suite à une étude de l'erreur moyenne de projection en fonction du nombre de cercles directeurs 3D (cf. figure 3.19 et tableau 3.1). Ce graphique et ce tableau mettent en évidence l'importance de la densité des points d'intérêt et d'une répartition appropriée des cercles sur la mire afin que l'étalonnage soit précis. Comme le soulignent ces résultats, il ne semble pas pertinent de multiplier le nombre de point d'intérêt. Pour 20 cercles dont la distribution est adaptée sur la mire, l'erreur de projection est faible et ne diminue plus significativement lorsque le nombre de points d'intérêt augmente.

On constate qu'une distribution adaptée de 160 points d'intérêt donne une erreur de rétro-projection moyenne ( $\bar{\epsilon} = 1.70 \pm 1.17\text{mm}$ ) comparable à une celle issue d'une distribution régulière de 360 points ( $\bar{\epsilon} = 1.77 \pm 1.27\text{mm}$ ). Le dimensionnement au préalable de l'outil d'étalonnage est donc un pré-requis nécessaire : un compromis entre l'erreur de rétro-projection et le nombre de points d'intérêt simplifie la réalisation matérielle de la mire pour une précision de l'étalonnage équivalente à une distribution régulière.

Les erreurs de reconstruction obtenues pour cette distribution adaptée des points d'intérêt sont plus élevées que celles obtenues pour une distribution régulière. L'erreur moyenne est de  $\bar{\epsilon} = 90.3 \pm 76.1$ . En rapportant les erreurs de reconstruction aux distances aux cibles, on constate une erreur moyenne en pourcent de :  $\bar{\epsilon}_{\%} = 2.74$ .

Les graphiques proposés sur les figures 3.13, 3.17 et 3.20 montrent des profils d'erreurs de rétro-projection tout à fait caractéristiques. Si l'on considère l'erreur de rétro-projection en fonction de la position angulaire du point test, on constate des erreurs faibles lorsque les points test sont proches des points de référence du maillage (ex :  $129.0^{\circ}/\epsilon = 0.16\text{mm}$ ). La projection dans l'image omnidirectionnelle d'un cercle d'étalonnage 3D ne décrit pas idéalement un cercle. En s'intéressant aux diamètres mesurés entre les barycentres d'ovoïdes diamétralement opposés, on constate de légères variations de sorte que la projection d'un cercle devient une ellipse. Par exemple, pour des données de simulation, le diamètre mesuré entre deux ovoïdes diamétralement opposés peut varier de quelques dixièmes de

pixels. Pour des données obtenues sur le système mire-captur réel, ces variations peuvent atteindre deux pixels (cf. tableau 3.2).

Données de simulation				
	$\phi_1$	$\phi_2$	$\phi_3$	$\phi_4$
cercle en $z = 350\text{mm}$	512.83	512.59	512.58	512.62
cercle en $z = 670\text{mm}$	197.55	197.75	197.75	197.47
Données du système réel				
	$\phi_1$	$\phi_2$	$\phi_3$	$\phi_4$
cercle en $z = 350\text{mm}$	512.20	511.59	512.68	513.88
cercle en $z = 670\text{mm}$	196.73	196.99	198.92	198.96

TAB. 3.2 – Diamètres des projections des cercles 3D mesurés entre des barycentres d’ovoïdes diamétralement opposés

Pour les données de simulation, ces déformations sont essentiellement dues aux approximations effectuées lors de l’estimation des centres de gravité (cf. paragraphe 2.5.2). Du fait du processus d’interpolation, les points sont donc répartis sur une ellipse et non un cercle (cf. figure 3.23). L’erreur est donc distribuée de manière non-statistique sous forme de « cloches » traduisant la rétro-projection d’une ellipse sur le cylindre.

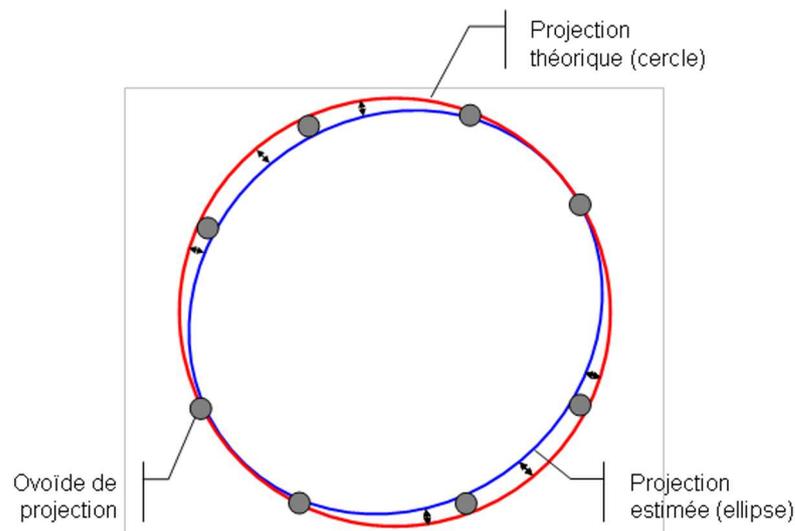


FIG. 3.23 – Projection théorique et estimée des ovoïdes. La projection théorique est un cercle. Du fait d’approximations dans l’estimation des centres de gravité, les projections effectivement estimées sont des ellipses ce qui explique l’allure spécifique des erreurs de rétro-projection

Pour les données obtenues sur le système réel, aux approximations lors de l'estimation des barycentres des ovoïdes s'ajoutent des défauts de positionnement de la mire au regard du capteur. L'axe de révolution de la mire et l'axe optique du capteur ne sont pas idéalement confondus ce qui conduit à ce que la projection d'un cercle ne soit pas un cercle mais une ellipse dans l'image omnidirectionnelle.

L'étalonnage non-paramétrique revêt un inconvénient majeur : les coordonnées des points d'intérêt 3D sont supposées parfaitement connues dans les repères des capteurs (c'est-à-dire les repères miroirs). En d'autres termes, les matrices de changement de base caractérisant les paramètres extrinsèques aux capteurs sont supposées parfaitement connues. Objectivement, cette hypothèse est difficilement formulable, la plupart des méthodes d'étalonnage ré-estimant d'ailleurs la pose entre la mire et le capteur au cours du processus de minimisation. Le caractère non-paramétrique de la méthode d'étalonnage rend ce problème d'estimation de pose délicat.

La méthode proposée (mais non implémentée) offre une solution au problème d'estimation de pose pour un étalonnage non-paramétrique. Elle utilise les spécificités de la mire et plus particulièrement le fait qu'un cercle en 3D se projette en un cercle lorsque l'axe de révolution de la mire et l'axe optique du capteur sont confondus. Lorsqu'un biais existe, le projeté d'un cercle est une ellipse. L'algorithme envisagé (cf. figure 3.24) se propose de modifier itérativement les coordonnées des points d'intérêt 3D afin que les projections décrivent des cercles dans l'image omnidirectionnelle.

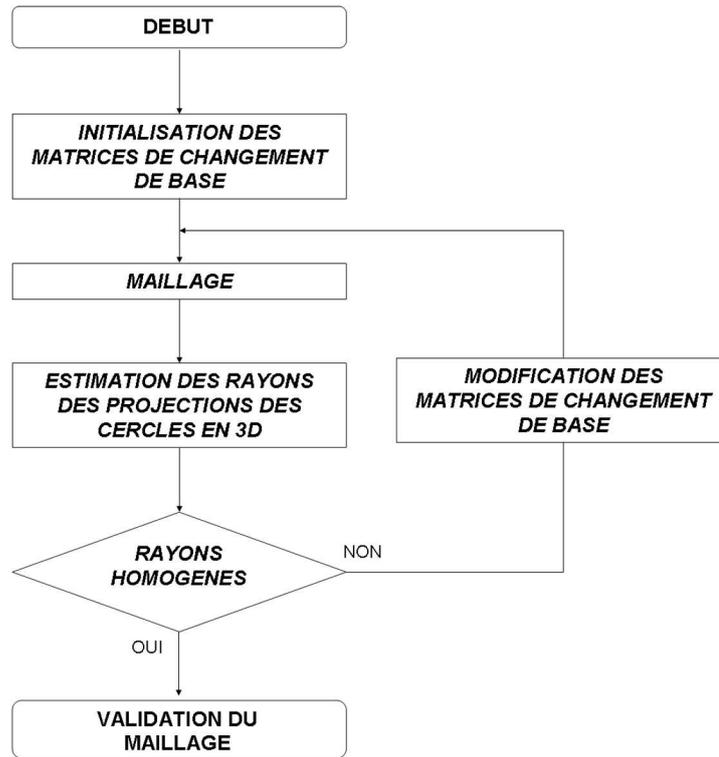


FIG. 3.24 – Estimation de la pose entre la mire et le capteur pour un étalonnage non-paramétrique.

Soit un maillage obtenu pour une estimée initiale de la pose entre la mire et le capteur. Ce maillage associe des points 3D  $\mathbf{Q}$  répartis sur la mire à leurs projections  $\mathbf{q}$  dans l'image omnidirectionnelle. Lorsque la mire est parfaitement positionnée par rapport au capteur, un cercle directeur  $C$  défini par huit points 3D cocycliques, se projette en un cercle  $c$  dans l'image omnidirectionnelle. Lorsqu'un défaut de positionnement apparaît, la projection d'un cercle directeur  $C$  définit une ellipse  $e$ . Soit  $\mathbf{Q}_{m,i}$  et  $\mathbf{Q}_{m,j}$  des points 3D appartenant à un cercle directeur de la mire d'étalonnage de sorte que  $\mathbf{Q}_{m,i}$  et  $\mathbf{Q}_{m,j}$  soient contigus. On note  $\mathbf{q}_{m,i}$  et  $\mathbf{q}_{m,j}$  leurs projetés respectifs. En présence d'un léger défaut de positionnement, les distances du centre de l'image aux deux points 2D contigus  $\mathbf{q}_{m,i}$  et  $\mathbf{q}_{m,j}$  ne sont pas égales :  $\rho_{m,i} \neq \rho_{m,j}$ . Les coordonnées des points 3D  $\mathbf{Q}_{m,i}$  et  $\mathbf{Q}_{m,j}$  sont alors modifiées de sorte qu'une variation angulaire  $\Delta\boldsymbol{\alpha} = (\Delta\alpha_X, \Delta\alpha_Y)^\top$  soit appliquée aux deux angles  $\alpha_X$  et  $\alpha_Y$  de la matrice de rotation  $R_M$  (aucune variation n'est appliquée sur l'angle  $\alpha_Z$  puisqu'il peut être facilement estimé). Une variation élémentaire  $\Delta\mathbf{t} = (\Delta t_M^X, \Delta t_M^Y, \Delta t_M^Z)^\top$  est également appliquée sur le vecteur de translation  $\mathbf{t}_M$ . Si  $\mathbf{Q}'_{m,i}$  est le point d'intérêt 3D dont les coordonnées ont été modifiées,  $\mathbf{Q}'_{m,i}$  et  $\mathbf{Q}_{m,i}$  sont alors reliés par la relation suivante :

$$\forall i \in \{1, \dots, I\}, \quad \mathbf{Q}'_{m,i} = (\mathbf{R}_M(\alpha_X + \Delta\alpha_X, \alpha_Y + \Delta\alpha_Y) | \mathbf{t}_M + \Delta\mathbf{t}) \mathbf{Q}_{m,i} \quad (3.19)$$

Les points d'intérêt 3D  $\mathbf{Q}'_{m,i}$  dont les coordonnées ont été modifiées se projettent en les  $\mathbf{q}'_{m,i}$  grâce aux données d'étalonnage à l'itération précédente. Si l'ellipse  $e$  tend à devenir un cercle, c'est-à-dire que les rayons des points 2D tendent à devenir égaux,  $\rho_{m,i} \approx \rho_{m,j}$ , le maillage est validé pour la suite des processus. Sinon les coordonnées des points d'intérêt 3D sont une nouvelle fois modifiées. Le maillage est définitivement validé lorsque les rayons sont égaux pour les huit points 2D décrivant la projection du cercle directeur. Ceci doit être également vérifié pour les projections de l'ensemble des cercles directeurs.

## 3.6 Conclusions

Au travers de ce **CHAPITRE 3**, notre contribution au problème de l'étalonnage de capteurs catadioptriques a été exposée. Il s'agit d'une méthode discrète et non-paramétrique qui définit les relations espace 3D-plan image dans un nouvel espace de représentation et pour un nombre fini de correspondances. Afin d'étendre les relations à tous les couples point 3D-point 2D, une interpolation numérique permet d'approximer localement les fonctions de projection et rétro-projection. Il s'agit d'une interpolation linéaire qui utilise une triangulation de Delaunay pour déterminer l'équation d'un plan. Le choix de cette méthode a été principalement motivé par sa simplicité d'implémentation et son adéquation pour des traitements temps-réel.

La précision de l'étalonnage est donc fonction de la densité du maillage et de la distribution des points d'intérêt 3D sur la mire. La suite du **CHAPITRE 3** s'est donc intéressée à une méthode pour le pré-dimensionnement de la mire. Cette procédure permet d'une part de définir le nombre de points d'intérêt par cercle directeur. Elle permet d'autre part d'estimer le nombre de cercles et leurs positions sur la mire de sorte à ce que cette nouvelle distribution soit la mieux adaptée pour le capteur.

Le **CHAPITRE 3** se termine par une présentation des résultats pour l'étalonnage non-paramétrique obtenus en simulation et pour le système mire-capteur réel. Les résultats sont encourageants puisque les différentes erreurs mesurées avoisinent celles obtenues pour la procédure d'étalonnage paramétrique. Ceci permet donc de conclure quant à la pertinence de la méthode d'étalonnage proposée.

Finalement, au regard des exigences des applications du programme *OmniViss*, la procédure d'étalonnage offre une solution pertinente pour établir les relations espace 3D-plan image. Les données d'étalonnage peuvent être organisées sous forme de « *Look Up Table* » qui est un processus de stockage particulièrement adapté pour des traitements temps-réel. De plus, la procédure offre une grande souplesse d'utilisation puisque l'ensemble des données d'étalonnage peut être plus ou moins dense en fonction du degré de précision requis par l'application.

L'étalonnage du capteur de vision n'étant pas une finalité en soi mais un préalable pour l'obtention de modèles 3D, à partir des données d'étalonnage organisées sous forme de « *Look Up Table* », nous proposons l'implémentation d'algorithmes pour des applications en reconstruction de scènes 3D.

# Chapitre 4

## Applications à la reconstruction de scènes 3D

### Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>123</b>
<b>4.2</b>	<b>Reconstruction 3D volumétrique pour le capteur de stéréovision omnidirectionnelle</b>	<b>125</b>
4.2.1	Principe général	125
4.2.2	Applications au capteur de stéréovision omnidirectionnelle	127
<b>4.3</b>	<b>Résultats</b>	<b>132</b>
4.3.1	Reconstructions 3D statiques pour une scène réelle	132
4.3.2	Reconstruction 3D dynamique par déplacements successifs du capteur pour une scène de synthèse	137
<b>4.4</b>	<b>Discussion</b>	<b>138</b>
<b>4.5</b>	<b>Conclusions</b>	<b>141</b>

---

## 4.1 Introduction

Les méthodes d'obtention d'un modèle 3D d'un objet en particulier ou d'une scène en général sont appelées « *shape from x* » où  $x$  tient lieu de « *shading* » [70], [172], [122], [123], [124], [34], « *texture* » [153], [90], [89], « *contour* » [34], « *stereo* », etc. Ces méthodes de modélisation 3D sont très différentes de part les procédés mis en jeu, les données initiales requises et les algorithmes utilisés. Elles sont, de ce fait, utilisées dans des cadres d'application différents.

Pour le cas du système de stéréovision binoculaire omnidirectionnelle, une méthode traditionnelle s'appuie sur les données 2D pour reconstruire la 3D. Nous faisons ici référence au « *shape from stereo* » qui consiste à trianguler les lignes de vue 3D associées à des primitives 2D extraites des images pour estimer la primitive 3D correspondante. Le principe de cette méthode est d'utiliser des images d'une même scène prises sous différents points de vue, d'apparier des primitives 2D stéréo-correspondante en exploitant la géométrie épipolaire des deux capteurs (cf. paragraphe 1.4) et de trianguler les lignes de vue 3D associées aux primitives appariées afin d'obtenir le modèle 3D. Les images sont soit acquises simultanément par les deux capteurs, soit obtenues préalablement aux traitements.

Les modèles 3D obtenus par cette méthode peuvent être métriques. La pose relative entre les deux capteurs est supposée avoir été estimée et chaque capteur calibré. L'algorithme est alors décomposé en deux phases : une première concerne l'appariement des points stéréo-correspondants, c'est-à-dire l'identification des points 2D correspondant au même point dans l'espace 3D. La seconde se rapporte à la triangulation des lignes de vue pour estimer la position du point 3D [66], [67], [13].

Les reconstructions 3D obtenues peuvent être connues à un facteur d'échelle près. L'algorithme consiste alors à estimer la matrice fondamentale du capteur de stéréovision, à déterminer la carte de disparité associée aux deux images. Une échelle est alors choisie pour établir le modèle 3D. Cette échelle ne tient compte d'aucune métrique de la scène 3D réelle, ce qui conduit à obtenir un modèle 3D à un facteur d'échelle près. A ces méthodes sont généralement associés des processus d'ajustement de surfaces par des maillages polygonaux afin d'affiner le modèle 3D [41], [143], [139].

Une alternative au « *shape from stereo* » réside dans l'utilisation d'un tableau 3D dont chaque cellule est un volume élémentaire appelé *voxel*<sup>20</sup> [36], [140], [148], [78]. Il s'agit de méthodes de reconstruction dites par « *shape from photoconsistency* » ou « méthode de

---

<sup>20</sup>*voxel* : abréviation de « *volume element* ». Il s'agit de l'homologue du pixel en 3D

---

reconstruction volumétrique », qui sont principalement dédiées à des applications multi-vues, c'est-à-dire lorsque plusieurs images d'un même objet sont disponibles (Seitz & al. [139] utilisent un jeu de données constitué de 317 images d'un même objet). Ces méthodes sont donc fondamentalement différentes de celles utilisant l'appariement de primitives bidimensionnelles.

Les travaux présentés dans ce chapitre se positionnent volontairement sur l'implémentation et l'évaluation d'une méthode de reconstruction volumétrique dédiée au capteur de stéréovision binoculaire omnidirectionnelle. Le **CHAPITRE 4** est structuré de la manière suivante : une première partie expose notre contribution pour la reconstruction de scènes 3D. Une deuxième partie présente les résultats obtenus pour des modèles 3D statiques et dynamiques. Une discussion et une conclusion viennent clore ce chapitre.

## 4.2 Reconstruction 3D volumétrique pour le capteur de stéréovision omnidirectionnelle

### 4.2.1 Principe général

Les méthodes de reconstruction par « *shape from photoconsistency* » [114], également connues sous l'appellation de « *volumetric methods* » (**méthodes volumétriques**) du fait des processus mis en jeu, sont apparues dans les années 1970 pour des applications médicales [63]. Les temps de calculs des processeurs et les capacités mémoires des composants actuels ont permis à ces méthodes de modélisation 3D d'être des alternatives pertinentes aux méthodes de reconstruction 3D traditionnelles basées sur l'appariement de primitives, triangulation des lignes de vue 3D et ajustement de surfaces [143].

Les méthodes par « *shape from photoconsistency* » utilisent le fait que les projections d'un même point 3D dans les images de capteurs orientés différemment dans l'espace 3D doivent avoir une cohérence colorimétrique [114]. En d'autres termes, les projections d'un point 3D appartenant à une surface doivent avoir une similarité colorimétrique, c'est-à-dire être consistants colorimétriquement ou être photo-consistants pour différents capteurs de l'espace 3D [138] (cf. figure 4.1). Ce postulat, à la base des algorithmes de « *shape from photoconsistency* », prend en compte les différentes perturbations (condition d'éclairage, propriétés surfaciques des matériaux, etc.) qui conduisent à ce que les projections d'un même point 3D observé sous des points de vue différents ne possèdent pas des caractéristiques parfaitement identiques. Ce postulat considère également que les capteurs sont calibrés et que la scène 3D est lambertienne.

L'estimation de la similarité colorimétrique des projections est donc un paramètre fondamental pour ces méthodes de reconstruction 3D. Ozun & al. [114] proposent une synthèse des mesures de photo-consistance disponibles dans la littérature. Certaines estiment l'écart-type des couleurs des projections du voxel pour statuer sur son appartenance à une surface [138], [140], [144]. D'autres utilisent les normes de Minkowski (cf. annexe E) et d'autres encore évaluent la similarité des histogrammes des projections du voxel.

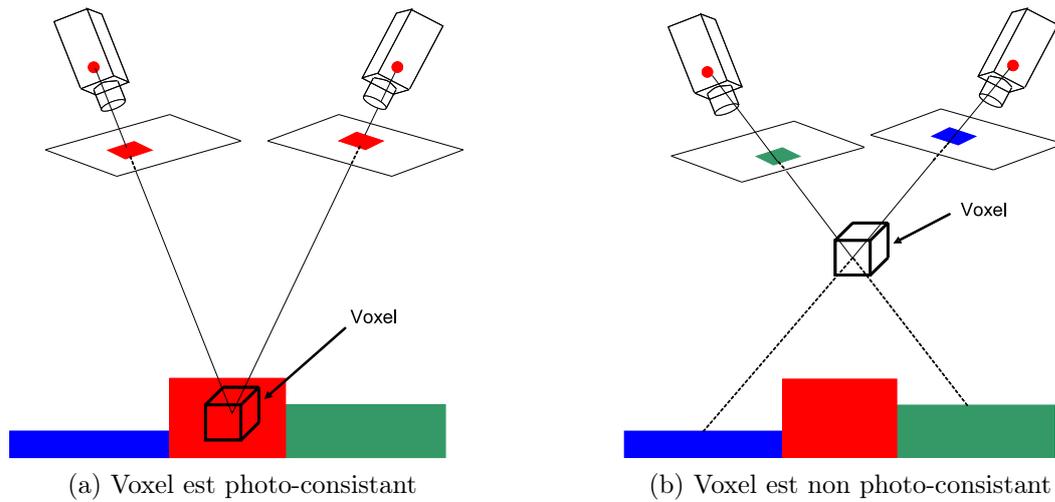


FIG. 4.1 – Principe de la photo-consistance d'un voxel. a) le voxel appartient à une surface et il est donc photo-consistant. b) le voxel n'est pas photo-consistant car il n'appartient pas à une surface [143].

Les méthodes de reconstruction par « *shape from photoconsistency* » modélisent la scène 3D par un volume fermé régulier, généralement cubique, uniformément échantillonné. Chaque échantillon est une cellule élémentaire appelée voxel [41] qui peut être défini par son centre de gravité ou par ses huit sommets [148]. Le volume de représentation est ensuite balayé dans un ordre précis et chaque voxel est projeté dans les images des différents capteurs. Le test de photo-consistance est alors appliqué sur le pixel (cas où le voxel est défini par son centre de gravité) ou l'ensemble de pixels définissant les projections. En fonction du résultat de photo-consistance trouvé, une décision binaire (transparent-opaque) ou ternaire (transparent-opaque-non vu) statue sur la nature du voxel projeté [41].

L'algorithme classique de « *shape from photoconsistency* » se rapporte au « *voxel coloring* » proposé par Seitz & Dyer [138], [140]. Le « *space carving* » proposé par Kutulakos & Seitz [77], le « *generalized voxel coloring* » décrit par Culbertson & al. [36], et le « *multi-hypothesis voxel coloring* » exposé par Eisert & al. [43] entre autres, peuvent être vus comme des variantes de l'algorithme d'origine puisque pour une même mesure de photo-consistance, des modèles 3D comparables sont obtenus [114]. Ozun & al. [114] mentionnent que la mesure de cohérence colorimétrique est le seul paramètre influant pour ces types d'algorithmes. A ce titre, ils proposent un état de l'art ainsi qu'une évaluation des différentes mesures de photo-consistance et leurs implications pour ces types de reconstructions 3D.

## 4.2.2 Applications au capteur de stéréovision omnidirectionnelle

L'approche de reconstruction 3D par stéréovision passe par une étape primordiale : l'appariement de primitives 2D. Pour un capteur catadioptrique, cette étape est délicate car le voisinage n'est pas constant. Une approche volumétrique constitue une alternative pertinente car elle préserve la cohérence géométrique entre la scène et le capteur. L'adéquation entre les données d'étalonnage stockées sous forme de « *Look Up Table* » et la structure algorithmique des méthodes volumétriques motive également le choix de recourir à cette solution pour les applications en modélisation de scènes 3D.

L'évaluation de cette méthode s'est donc déroulée en deux phases :

- une première s'est rapportée à l'évaluation de la procédure pour une configuration statique du capteur, c'est-à-dire pour une acquisition simultanée de la paire d'images ;
- la seconde s'est intéressée à l'évaluation de cette méthode pour des mouvements élémentaires du capteur. Il s'agit d'une reconstruction volumétrique dynamique qui est définie comme l'intersection des reconstructions statiques réalisées à des instants successifs.

### 4.2.2.1 Reconstruction volumétrique statique

Pour un instant  $t$  donné, considérons un volume de reconstruction  $V$  fermé de forme cubique constitué d'un ensemble de voxels  $\{\nu_1, \dots, \nu_i, \dots, \nu_I\}$ . Le capteur de stéréovision omnidirectionnelle admettant une symétrie axiale d'axe  $\mathbf{Z}$ , sa position est définie au centre du volume de reconstruction de sorte que les axes  $\mathbf{Z}$  du référentiel  $\mathfrak{R}_W$  associé au volume  $V$  et des capteurs du banc stéréoscopique :  $\mathfrak{R}_M^B$  et  $\mathfrak{R}_M^H$  soient confondus. Un voxel  $\nu_i$  est une cellule élémentaire cubique qui peut être approximé [148] :

- par son centre de gravité lorsque la cellule est infinitésimale. Les dimensions du voxel sont considérées comme infinitésimales par rapport à la résolution du capteur lorsque le voxel, défini par ses huit sommets, se projette en une surface dont l'aire est inférieure ou égale au pixel ;
- par ses huit sommets lorsque la cellule ne peut être considérée comme infinitésimale [138]. La projection des huit sommets du voxel définit un nuage de huit points dans l'image qui couvre deux pixels au minimum.

**Approximation du voxel par son centre de gravité :** soit un voxel  $\nu_i$  de dimension infinitésimale approximé par son centre de gravité  $\mathbf{CG}_i$  (cf. figure 4.2). Notons  $\mathbf{q}_i^B$  et  $\mathbf{q}_i^H$

les projections du voxel sur les plans images des capteurs bas et haut (réalisées grâce aux données d'étalonnage des capteurs bas et haut). La mesure de photo-consistance pour statuer de l'appartenance du voxel à une surface est la norme  $L^2$  dans l'espace des couleurs  $R, G, B$ . Aux points 2D  $\mathbf{q}_i^B$  et  $\mathbf{q}_i^H$  sont associés les paramètres colorimétriques  $R^B, G^B, B^B$  et  $R^H, G^H, B^H$ . La mesure de photo-consistance est alors donnée par :

$$\text{photo-consistant } (\nu_i) : \begin{cases} \text{vrai si } \sqrt{(R^B - R^H)^2 + (G^B - G^H)^2 + (B^B - B^H)^2} < s \\ \text{faux sinon} \end{cases} \quad (4.1)$$

$s$  est ici un seuil fixé par l'utilisateur pour confirmer ou infirmer l'appartenance du voxel  $\nu_i$  à la surface d'un objet.

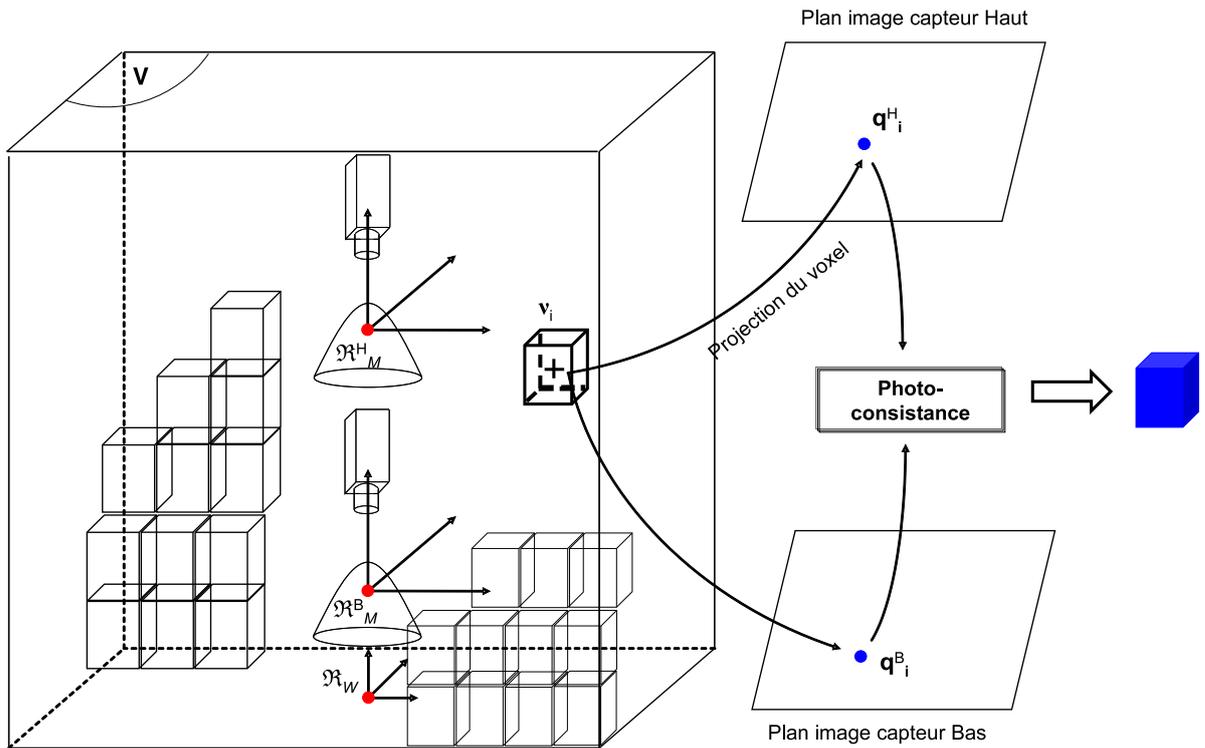


FIG. 4.2 – Reconstruction 3D volumétrique pour le capteur de stéréovision omnidirectionnelle. Chaque voxel est défini par son centre de gravité. Les projections sont donc ponctuelles dans les images omnidirectionnelles.

Approximation du voxel par ses huit sommets : soit  $\nu_i$  un voxel quelconque du volume de reconstruction défini par ses huit sommets  $\{\mathbf{S}_{i,1}, \dots, \mathbf{S}_{i,j}, \dots, \mathbf{S}_{i,8}\}$  (cf. figure 4.3). La

projection du voxel se résume à projeter les huit sommets dans les images des capteurs bas et haut en les  $\{\mathbf{q}_{i,1}^B, \dots, \mathbf{q}_{i,j}^B, \dots, \mathbf{q}_{i,8}^B\}$  et  $\{\mathbf{q}_{i,1}^H, \dots, \mathbf{q}_{i,j}^H, \dots, \mathbf{q}_{i,8}^H\}$ . Ces points 2D définissent des nuages de points qui sont approximés par des rectangles englobants  $\mathbf{R}^B$  et  $\mathbf{R}^H$ .

La mesure de la photo-consistance implémentée pour cette définition du voxel est reprise des travaux menés par Seitz & Dyer [140]. Soit  $\{c_1, \dots, c_i, \dots, c_I\}$  l'ensemble des couleurs associées aux rectangles englobants  $\mathbf{R}^B$  et  $\mathbf{R}^H$ . Notons également  $\sigma$  l'écart-type de cet ensemble colorimétrique. Le critère de photo-consistance pour un voxel  $\nu_i$  est donc défini par :

$$\text{photo-consistant } (\nu_i) : \begin{cases} \text{vrai si } \frac{\sigma^2}{\sigma_0^2} < s \\ \text{faux sinon} \end{cases} \quad (4.2)$$

$s$  est ici un seuil fixé par l'utilisateur pour confirmer ou infirmer l'appartenance du voxel  $\nu_i$  à la surface d'un objet.  $\sigma_0^2$  peut être obtenu en calculant l'écart-type d'une surface de couleur homogène vue par les deux capteurs.  $\sigma_0$  permet donc de prendre en compte le fait que la numérisation d'une même surface en 3D puisse être sensiblement différente pour les deux capteurs.

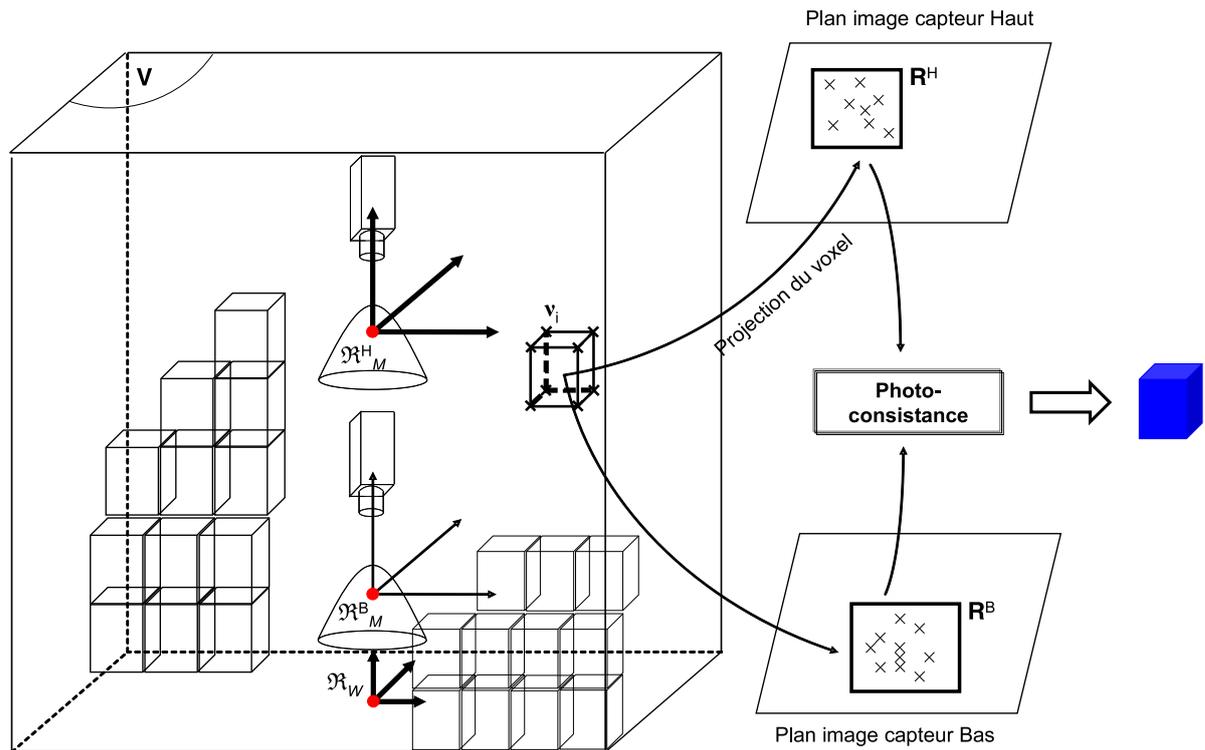


FIG. 4.3 – Reconstruction 3D volumétrique pour le capteur de stéréovision omnidirectionnelle. Chaque voxel est défini par ses huit sommets. A partir des données d'étalonnage, les sommets sont projetés dans les images omnidirectionnelles. Les nuages de points 2D sont approximés par des rectangles englobants.

Pseudo-code de l'algorithme de reconstruction 3D statique :

→ Pour un voxel défini par son centre de gravité :

PROGRAMME PRINCIPAL

DÉBUT

POUR chaque voxel  $\nu_i$  du volume de reconstruction  $V$

POUR chaque image des capteurs bas et haut

Projeter le centre de gravité  $CG_i$

Estimer la norme  $L^2$  pour les paramètres  $R^B, G^B, B^B$  et  $R^H, G^H, B^H$

SI  $L^2 < s$

Remplir le voxel avec la couleur du capteur haut

SINON laisser le voxel vide

FIN POUR

FIN POUR

FIN

→ Pour un voxel défini par ses huit sommets :

```
PROGRAMME PRINCIPAL
```

```
DÉBUT
```

```
  POUR chaque voxel  $\nu_i$  du volume de reconstruction  $V$ 
```

```
    POUR chaque image des capteurs bas et haut
```

```
      Projeter les huit sommets du voxel sur les images
```

```
      Définir les rectangles englobant les huit sommets :  $R^B$  et  $R^H$ 
```

```
      Calculer l'écart-type  $\sigma$  de couleurs  $\{c_1, \dots, c_i, \dots, c_I\}$  associées aux rectangles englobants  $R^B$  et  $R^H$ 
```

```
      SI  $\frac{\sigma^2}{\sigma_0^2} < s$ 
```

```
        Remplir le voxel avec la couleur du capteur haut
```

```
      SINON laisser le voxel vide
```

```
    FIN POUR
```

```
  FIN POUR
```

```
FIN
```

#### 4.2.2.2 Reconstruction volumétrique dynamique

La reconstruction volumétrique dynamique  $R3D_{dyn}(t)$  est définie comme l'intersection des reconstructions volumétriques statiques obtenues aux instants  $\{t, \dots, t + n\Delta t, \dots, t + N\Delta t\}$ , pour des déplacements successifs et connus du capteur. La reconstruction volumétrique dynamique est alors définie par :

$$R3D_{dyn}(t) : \bigcap_{n=0}^N R3D_{stat}(t + n\Delta t) \quad (4.3)$$

Le référentiel choisi, c'est-à-dire celui dans lequel sont ré-exprimées les différentes reconstructions volumétriques statiques de la séquence, est celui pour l'instant  $t$  d'origine. Ce référentiel d'origine peut néanmoins être recalculé en fonction du déplacement du capteur et de la taille de la zone mémoire que l'on souhaite allouer pour cette opération.

## 4.3 Résultats

### 4.3.1 Reconstructions 3D statiques pour une scène réelle

La reconstruction statique est obtenue pour une scène d'intérieur réelle (ex. un bureau). Le volume de reconstruction  $V$  est constitué de  $500 \times 500 \times 200$  voxels. Un voxel est une cellule cubique élémentaire dont les dimensions sont de  $10 \times 10 \times 10\text{mm}^3$ . Les résultats sont illustrés par plusieurs vues du volume de reconstruction :

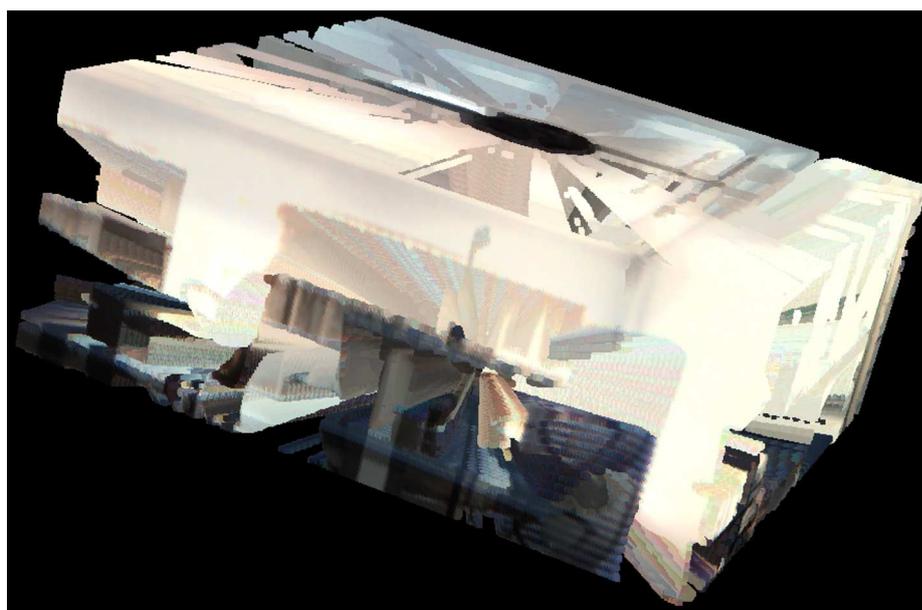
- la figure 4.4 présente une vue d'extérieur du volume de reconstruction tel qu'il apparaît après projection de l'intégralité des voxels ;
- la figure 4.5 représente le volume de reconstruction vu sous le même angle d'observation et pour lequel les voxels, dont la valeur de luminance est élevée sont supprimés. Il s'agit ici de supprimer les voxels modélisant le plafond afin d'apprécier plus finement la zone utile à reconstruire ;
- la figure 4.6, il s'agit d'une portion du volume de reconstruction de la figure 4.5 découpé suivant le plan  $(\mathbf{Y}_W \mathbf{O}_W \mathbf{Z}_W)$  ;
- enfin la figure 4.7 montre une vue en coupe de cette dernière portion de volume.



(a) Image capteur bas



(b) Image capteur haut



(c) Reconstruction 3D statique

FIG. 4.4 – Reconstruction 3D statique d'une scène d'intérieur. Le voxel est considéré comme une cellule infinitésimale qui est approximée par son centre de gravité.

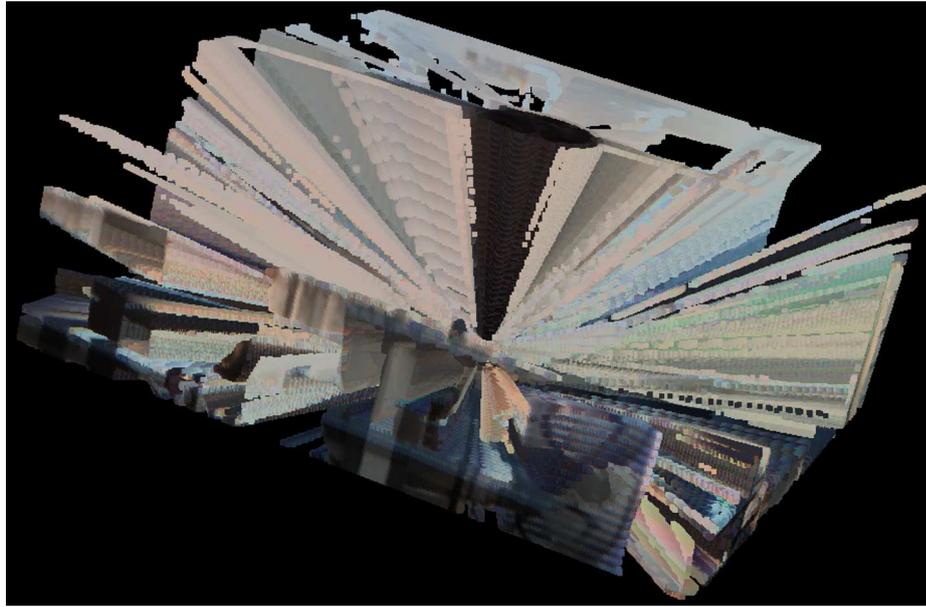


FIG. 4.5 – Reconstruction 3D statique d’une scène d’intérieur. Le volume de reconstruction est observé sous un angle de vue similaire à la figure 4.4. Les voxels dont la luminance est élevée ont été supprimés pour permettre une interprétation plus aisée du modèle 3D.

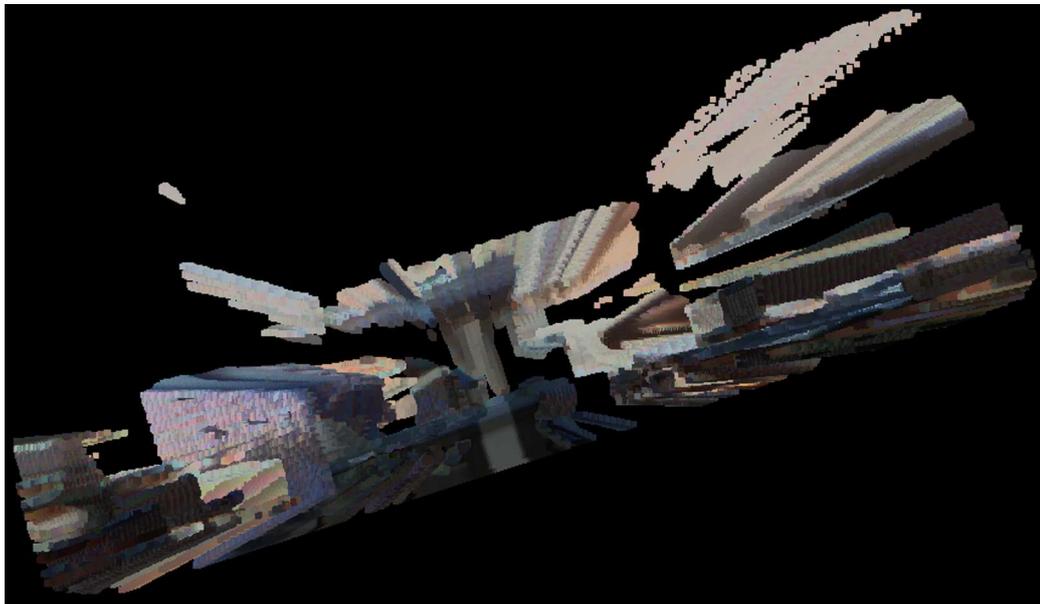


FIG. 4.6 – Vue d’une portion du volume de reconstruction découpé dans le plan  $(\mathbf{Y}_W \mathbf{O}_W \mathbf{Z}_W)$ , l’observateur étant placé à l’intérieur du volume 3D.

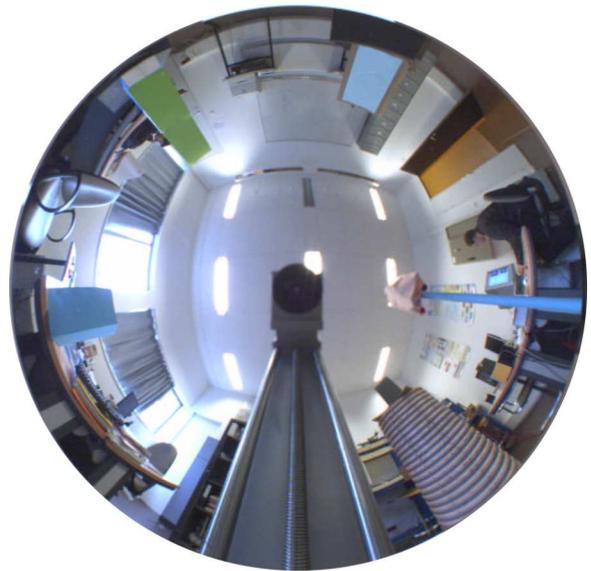


FIG. 4.7 – Vue en coupe dans le plan  $(Y_W O_W Z_W)$  du volume de reconstruction, l'observateur étant placé à l'intérieur du volume 3D.

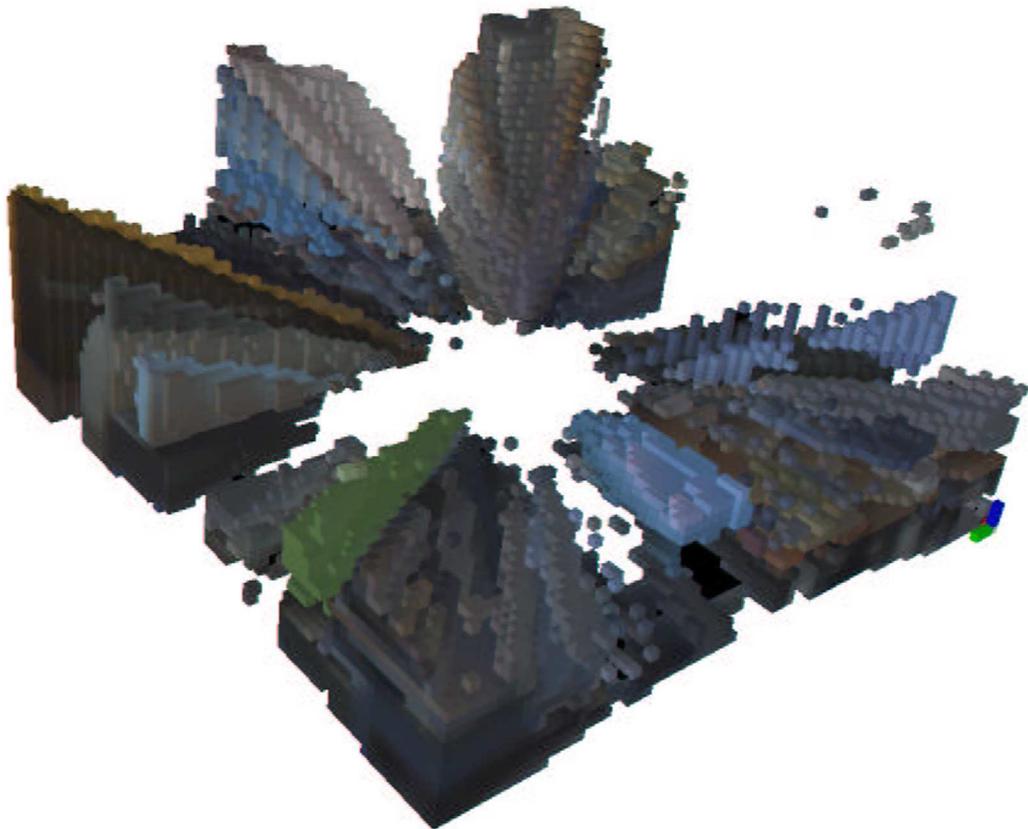
La reconstruction 3D statique est obtenue pour une scène d'intérieur. Le volume de reconstruction  $V$  est constitué de  $200 \times 200 \times 100$  voxels. Chaque voxel est une cellule élémentaire de dimensions  $50 \times 50 \times 50 \text{mm}^3$ . La figure 4.8 présente les deux images des capteurs bas et haut acquises par le banc stéréoscopique à un instant  $t$  donné et la reconstruction 3D statique correspondante.



(a) Image capteur bas



(b) Image Capteur haut



(c) Reconstruction 3D statique

FIG. 4.8 – Reconstruction statique pour une scène 3D d'intérieur. Les dimensions du voxel ne permettent pas de le considérer comme une cellule infinitésimale. Le voxel est par conséquent défini par ses huit sommets.

### 4.3.2 Reconstruction 3D dynamique par déplacements successifs du capteur pour une scène de synthèse

La reconstruction dynamique présentée en figure 4.9 est réalisée à partir d'une séquence de six paires d'images acquises à des instants successifs. Cette séquence est générée en simulation à l'aide du logiciel POV-Ray. La scène est constituée de trois colonnes de couleurs verte, violette et jaune, de murs rouges et d'un plafond bleu. Le volume de reconstruction  $V$  associé aux reconstructions statiques possède des dimensions analogues à celles décrites précédemment pour la reconstruction 3D statique présentée en figure 4.8, à savoir :  $200 \times 200 \times 100$  voxels et un voxel est défini comme étant un cube de  $50 \times 50 \times 50$  mm<sup>3</sup>.

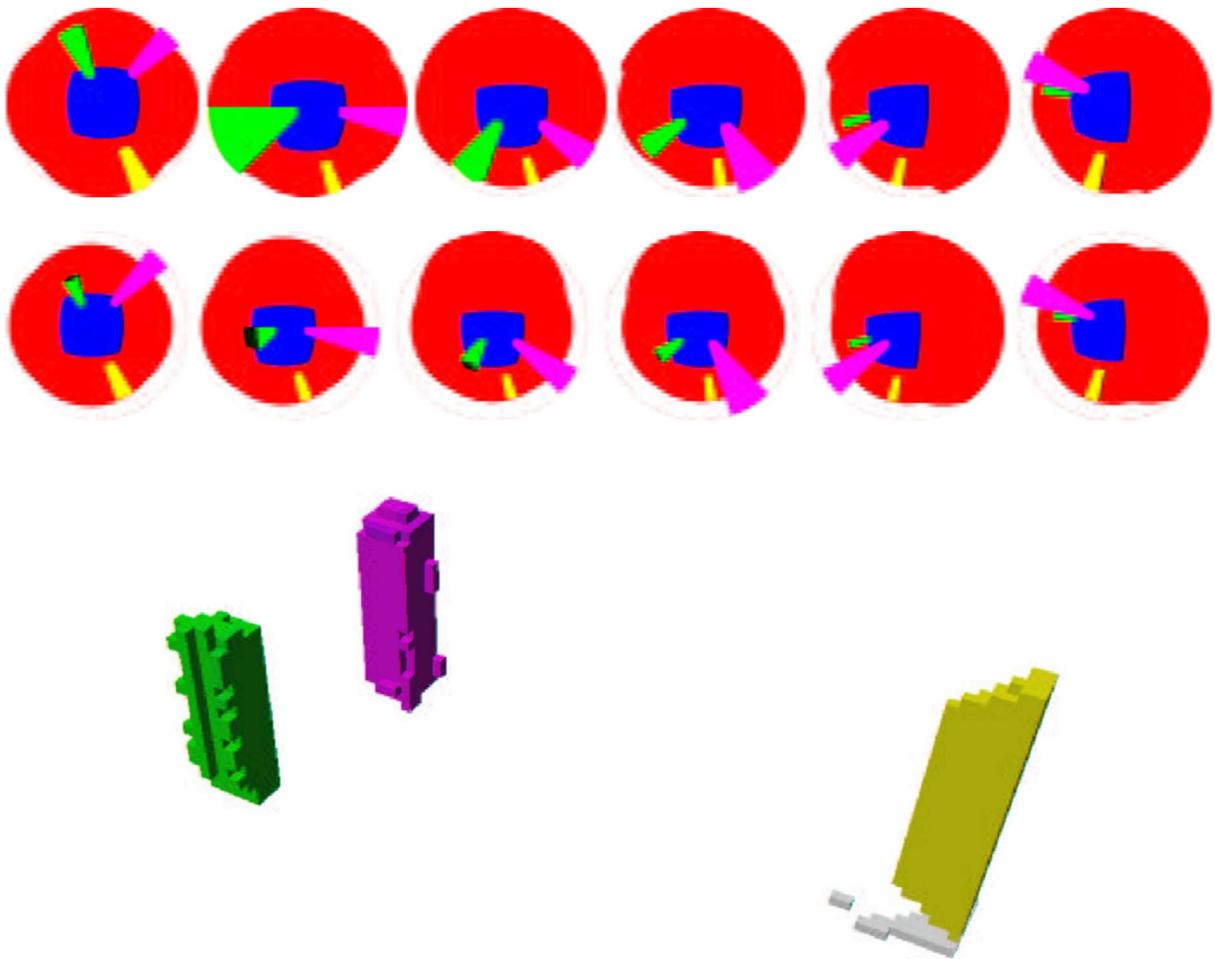


FIG. 4.9 – Reconstruction dynamique pour une scène 3D de synthèse modélisée sous POV-Ray. Cette représentation est définie comme l'intersection de six reconstructions 3D statiques pour cinq déplacements successifs du capteur dans la scène 3D. Les déplacements du capteur en rotation et en translation sont supposés parfaitement connus.

## 4.4 Discussion

Les figures 4.4, 4.5, 4.6, 4.7 et 4.8 sont des modèles 3D statiques d'une scène réelle d'intérieur obtenus pour deux définitions du voxel différentes. En particulier, pour les figures 4.4, 4.5, 4.6, 4.7, le voxel est approximé à un point unique, son centre de gravité, dans l'espace 3D de sorte que sa projection est ponctuelle dans l'image omnidirectionnelle. Implicitement, cette hypothèse considère que les dimensions du voxel sont infinitésimales au regard de la résolution du capteur. En d'autres termes, un voxel quelconque du volume de reconstruction se projette dans l'image omnidirectionnelle sur une surface inférieure au pixel au plus.

A l'inverse, la figure 4.8 considère le voxel comme une entité non-ponctuelle qui est définie par ses huit sommets. Le choix d'une approximation des projections des huit sommets par un rectangle englobant défini en valeurs entières de pixel est principalement motivé par les contraintes temps-réel des applications ciblées. Il s'agit d'une approximation simple qui ne requiert pas de traitements particuliers et qui, de ce fait, reste compatible avec le cahier des charges fixé.

Les mesures de photo-consistance implémentées sont respectivement une distance euclidienne sur les paramètres  $R, G, B$  lorsque le voxel est assimilé à un point dans l'espace 3D. Lorsque le voxel est défini par ses huit sommets, la mesure porte sur un écart-type colorimétrique des pixels constituant les rectangles englobants. Il s'agit d'une mesure inspirée des travaux de Seitz & Dyer [138], [140] qui est particulièrement adaptée lorsque l'objet est faiblement texturé. Néanmoins, l'inconvénient principal de ces deux types de mesure est la forte influence de la valeur du seuil sur les modèles 3D obtenus.

Les résultats illustrés sur les figures 4.4 et 4.5 sont encourageants puisque les modèles 3D obtenus laissent apparaître nettement des objets 3D reconstruits : un cylindre (objet strié à droite sur les figures) et une table où l'on distingue nettement le pied et le plateau (au centre sur les figures). La figure 4.4 en particulier est une représentation complète du volume de reconstruction, c'est-à-dire une représentation de l'intégralité des voxels qui ont été jugés photo-consistants. Le volume de reconstruction est globalement plein, avec la présence de zones qui ont été creusées (à gauche sur la figure 4.4).

Afin d'apprécier plus nettement les objets 3D reconstruits, les voxels modélisant le plafond ont volontairement été enlevés. Il s'agit de l'illustration présentée en la figure 4.5. Du fait des modèles 3D obtenus et plus particulièrement des profils des objets 3D qui apparaissent

sous forme de cônes dans le volume de reconstruction, la suppression du plafond ne peut se résumer à effacer les voxels à partir d'une certaine hauteur. Une identification par le paramètre de luminance a semblé être un critère pertinent pour caractériser ces voxels. La figure 4.5 laisse également clairement apparaître des cônes 3D associés aux objets reconstruits. Ces cônes de reconstruction prennent naissance au centre du volume de reconstruction et se prolongent à l'infini. La position du sommet du cône dans le volume de reconstruction est étroitement lié à la « *baseline* » du capteur stéréoscopique : plus la « *baseline* » est importante et plus le sommet du cône sera éloigné du centre du volume.

Les figures 4.6 et 4.7 sont des vues de l'intérieur du volume de reconstruction. La figure 4.6 en particulier est une vue d'une portion complète du volume de reconstruction et la figure 4.7 est une vue en coupe. Ces deux figures font apparaître les objets 3D reconstruits : le cylindre strié, le pied et le plateau de la table. Elles mettent également en évidence le fait que par endroits (notamment près du pied de la table), le volume de reconstruction est creusé.

La figure 4.8 présente le modèle 3D vu de dessus d'une scène réelle dans laquelle des objets spécifiques ont été disposés : un cube vert, des cubes bleus, etc. De manière analogue à la reconstruction statique présentée précédemment, les voxels modélisant le plafond ont été retirés pour faciliter l'interprétation du modèle 3D. Cette figure 4.8 montre une reconstruction 3D similaire à celles précédemment présentées : des objets 3D ont été reconstruits et apparaissent sous forme de cônes dans le volume de reconstruction (cube vert, cubes bleus, etc.). Des parties vides (principalement au centre du volume de reconstruction) illustrent le fait que le volume a été creusé par endroits.

Au vu des deux reconstructions 3D statiques présentées, que les voxels soient définis par leur centre de gravité ou par les huit sommets, des conclusions similaires peuvent être tirées : le volume de reconstruction est globalement plein et peu de zones ont été creusées. Les objets 3D sont reconstruits dans le volume de reconstruction et apparaissent sous forme de cônes en 3D. La nature de l'algorithme de reconstruction implémenté étaye ces conclusions. Il s'agit d'une méthode volumétrique qui est particulièrement adaptée lorsque plusieurs vues d'un même objet sont disponibles. Les deux images acquises par le capteur stéréoscopique ne suffisent donc pas à définir précisément la silhouette des objets en 3D.

Aussi, afin d'améliorer le modèle 3D statique, un processus de multi-vision par déplacements du capteur est mis en place. Il s'agit d'une multi-vision incrémentale puisque la séquence de paires d'images n'est pas disponible a priori. La figure 4.9 présente une

---

reconstruction volumétrique dynamique obtenue pour une scène 3D de synthèse. Des objets simples de part leurs formes et leurs couleurs sont placés dans cette scène. Afin de pouvoir interpréter plus aisément le modèle 3D final, les murs sont effacés manuellement. Les résultats obtenus sont particulièrement encourageants puisque le modèle 3D laisse clairement apparaître des silhouettes d'objets 3D nettement définies. Néanmoins, les paramètres ayant permis l'obtention de ce modèle 3D dynamique sont choisis de sorte à faciliter le processus d'intersection des reconstructions statiques acquises à des instants successifs : le capteur se déplace dans un plan, il n'est soumis à aucune rotation et le déplacement en translation est égal à un nombre entier de voxels.

## 4.5 Conclusions

Ce **CHAPITRE 4** a concerné l'évaluation d'une méthode alternative de modélisation 3D pour le dispositif développé pour le programme *OmniViss*. Il s'agit d'une méthode volumétrique principalement dédiée à des applications multi-vues, qui modélise la scène par un tableau 3D dont chaque cellule élémentaire est un voxel. Un critère de similarité colorimétrique entre les projections du voxel permet de statuer sur l'appartenance du voxel à la surface d'un objet. Le choix de cette solution de modélisation 3D a été principalement motivé par l'adéquation entre l'organisation des données d'étalonnage et la structure algorithmique de la procédure.

Deux scénarios de reconstruction ont été proposés : le premier concerne le capteur de stéréovision en position statique et le second considère le capteur en déplacement. Les résultats obtenus pour le modèle statique laissent apparaître des objets 3D reconstruits sous formes de cônes dans le volume de reconstruction. Le modèle dynamique pallie ces problèmes en proposant un processus de multi-vision incrémentale. Les résultats laissent apparaître des silhouettes d'objets 3D nettement définies. Néanmoins, ce modèle 3D dynamique est obtenu pour des hypothèses simples et des déplacements élémentaires.

Au final, les modèles 3D obtenus, tant pour les reconstructions statiques que dynamiques, ont permis de confirmer la cohérence de la méthode d'étalonnage proposée au **CHAPITRE 3** et du mode de stockage par « *Look Up Table* » afin de favoriser les traitements temps-réel (par exemple, le temps de traitement pour une reconstruction 3D statique qui modélise la scène 3D par un volume de  $200 \times 200 \times 100$  voxels est de 131 secondes).

# Conclusions et perspectives

## Conclusions

Les travaux de thèse présentés dans ce manuscrit s'inscrivaient dans un programme de recherche européen Interreg IIIA *OmniViss* associant l'Institut de Recherche en Systèmes Électroniques EMbarqués et le département d'Électronique de l'université du Kent en Grande-Bretagne pour **la conception et l'évaluation d'un outil de reconstruction et d'analyse de scènes 3D inconnues a priori**. L'objectif principal de l'étude était le déploiement d'une solution technologique répondant au problème de reconstruction 3D d'environnements sans connaissance a priori. Deux applications étaient alors envisagées : une première concernait la reconstruction 3D d'environnements pour un capteur en position fixe, une seconde se rapportait à la reconstruction de scènes 3D pour un capteur en mouvement. Les choix techniques pour la réalisation matérielle du capteur devaient donc prendre en compte les particularités des deux applications. Le système développé devait permettre une acquisition des données non intrusive, une observation large champ de l'environnement pour des applications temps-réel et intégrer des éléments bas-coût pour être compatible avec une solution industrialisable.

La solution matérielle retenue pour les applications du programme *OmniViss* s'est orientée vers un système compact, non-intrusif et permettant une observation large champ de l'environnement. De ce fait, une architecture combinant deux capteurs catadioptriques à miroirs hyperboloïdes positionnés verticalement a été retenue.

Un préalable nécessaire avant de modéliser la scène en 3D se rapporte à l'étalonnage du capteur. Il s'agit de définir les relations entre l'image et l'espace 3D. Ainsi, une méthode paramétrique utilisant un modèle de transformation ad hoc et une mire 3D a tout d'abord été implémentée.

Du fait d'une symétrie axiale du capteur, la mire choisie est un cylindre de révolution

---

dont les points d'intérêt sont des marqueurs lumineux de type LED. Les points d'intérêt sont répartis par cercle directeur et par génératrice et se projettent en des ovoïdes dans l'image omnidirectionnelle. La projection ponctuelle du marqueur est estimée par une moyenne pondérée par les niveaux de luminance associés aux pixels constituant l'ovoïde. Une carte électronique ainsi qu'une interface graphique ont été développées pour faciliter la procédure d'étalonnage du capteur.

L'étalonnage paramétrique peut être vu comme un problème de minimisation multivariable. La stratégie de résolution adoptée a consisté en une combinaison de méthodes linéaire (moindres carrés) et itérative sous contrainte (Levenberg-Marquardt) afin d'estimer le vecteur optimal des paramètres. Les résultats obtenus en simulation traduisent une précision remarquable de l'étalonnage ce qui permet de valider d'une part le modèle et d'autre part le processus d'estimation des projections des points d'intérêt 3D.

Le modèle ad hoc utilisé formule explicitement l'hypothèse d'un centre de projection unique. Cette contrainte est difficile à réaliser en pratique ce qui peut conduire à des dérives importantes dans la justesse de l'étalonnage. Une procédure d'étalonnage est donc proposée qui puisse s'affranchir des contraintes liées à l'utilisation des modèles. Il s'agit d'une méthode non-paramétrique qui définit les fonctions de projection et rétro-projection dans un nouvel espace de représentation et pour un nombre fini de points. Un maillage est donc réalisé sur ce nouvel espace et une interpolation locale au voisinage des points du maillage permet d'en fournir une estimation pour tous les couples point 3D-point 2D.

Une méthode classique de reconstruction 3D pour un capteur de stéréovision utilise l'appariement de primitives bidimensionnelles puis une triangulation des primitives appariées afin de recouvrir la primitive 3D. Notre attention s'est focalisée sur une méthode alternative qui modélise la scène par un tableau 3D dont chaque cellule est un voxel. Il s'agit d'une méthode de reconstruction volumétrique qui est classiquement utilisée pour des applications de multi-vision. L'algorithme de modélisation 3D est évalué dans deux cas d'étude :

- le premier est une reconstruction statique de la scène. Le capteur est fixe et la scène est reconstruite pour une acquisition simultanée d'images ;
- le second est une reconstruction dynamique. Le capteur est mobile et la scène 3D reconstruite est définie comme l'intersection des reconstructions statiques obtenues à des instants successifs.

Les modèles 3D obtenus pour le cas statique laissent clairement apparaître des silhouettes d'objets 3D reconstruites sous forme de 'cônes' dans le volume de reconstruction. La reconstruction dynamique obtenue pour des déplacements élémentaires du capteur permet

---

de définir plus finement les silhouettes des objets et donc d'améliorer significativement les modèles 3D.

## Perspectives

La distribution adaptée des points d'intérêt 3D a été évaluée en simulation et fournit des résultats encourageants. Il est donc nécessaire d'évaluer cette distribution adaptée sur un cas réel et donc de réaliser la mire en tenant compte des nouvelles coordonnées verticales des cercles directeurs 3D.

La méthode d'étalonnage non-paramétrique suppose une connaissance exacte des coordonnées des points d'intérêt 3D dans le repère capteur (c'est-à-dire le repère miroir). Il sera donc intéressant d'implémenter et d'évaluer la méthode d'estimation de pose proposée en fin de **CHAPITRE 3**, ce qui permettra à terme de statuer quant à sa faisabilité et sa pertinence.

A terme, il serait intéressant de proposer une comparaison et une synthèse des différentes méthodes d'étalonnage référencées dans ce rapport : paramétrique par un modèle ad hoc, paramétrique par le modèle unifié de Mei, paramétrique par le modèle unifié de Scaramuzza, non-paramétrique par le modèle de Ramalingam, etc.

Plus généralement, cette méthode d'étalonnage non-paramétrique pourrait être transposée à des miroirs de forme complexe dont le modèle n'est pas connu. Cette étape renforcerait l'intérêt de cette approche pour des systèmes bas-coût possédant des déformations locales générées lors des processus de fabrication.

La reconstruction 3D dynamique telle qu'elle est présentée dans ce rapport devrait être évaluée sur une séquence d'images réelles. De plus, le scénario de reconstruction dynamique tel qu'il est présenté actuellement formule des hypothèses contraignantes et restrictives : les déplacements sont simples et doivent être connus a priori, les déplacements s'effectuent selon un nombre entier de voxels, etc. Si l'on s'intéresse au cas particulier d'un robot mobile terrestre autonome, les déplacements sont très généralement fournis par l'odométrie. Un axe de recherche pourrait consister en l'implémentation d'algorithmes qui permettent l'estimation des déplacements par les seules informations de vision. Il serait également pertinent de proposer une solution au problème d'intersection des reconstructions statiques lorsque le capteur ne se déplace pas d'un nombre entier de voxels.

# Bibliographie

- [1] G. Adorni, S. Cagnoni, M. Mordonini, and A. Sgorbissa. Omnidirectional stereo systems for robot navigation. In *IEEE Workshop on Omnidirectional Vision and Camera Networks (in conjunction with IEEE Conference on Computer Vision and Pattern Recognition)*, pages 79–89, Madison, USA, June 2003.
- [2] I. Amidror. Scattered data interpolation methods for electronic imaging systems : a survey. *Journal of Electronic Imaging*, 11(2) :157–176, April 2002.
- [3] N. Ayache. *Vision Stéréoscopique et Perception Multisensorielle. Applications à la Robotique Mobile*. Masson, 1989.
- [4] N. Ayache and C. Hansen. Rectification of images for binocular and trinocular stereovision. Rapport de recherche RR-860, Institut National de Recherche en Informatique et Automatique, June 1988.
- [5] S. Baker and S. K. Nayar. A theory of catadioptric image formation. In *IEEE International Conference on Computer Vision*, pages 35–42, Bombay, India, January 1998.
- [6] S. Baker and S. K. Nayar. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2) :175–196, 1999.
- [7] H. Bakstein and T. Pajdla. Non-central cameras : a review. Research report CTU-CMP-2000-14, Czech Technical University, October 2000.
- [8] H. Bakstein and T. Pajdla. Panoramic mosaicing with a 180 field of view lens. In *IEEE Workshop on Omnidirectional Vision (in conjunction with European Conference on Computer Vision)*, pages 60–68, Copenhagen, Danemark, June 2002.
- [9] J. P. Barreto. *General Central Projection Systems : Modeling, Calibration and Visual Servoing*. Phd, University of Coimbra, October 2003.
- [10] J. P. Barreto and H. Araujo. Paracatadioptric camera calibration using lines. In *International Conference on Computer Vision*, volume 2, pages 1359–1365, Nice, France, October 2003.

- 
- [11] J. P. Barreto and H. Araujo. Geometric properties of central catadioptric line images and their application in calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8) :1327–1333, August 2005.
- [12] A. Bartoli. *Reconstruction et Alignement en Vision 3D : Points, Droites, Plans et Caméras*. Thèse de doctorat, Institut National Polytechnique de Grenoble, Septembre 2003.
- [13] A. Bartoli and J. T. Lapresté. Triangulation for points on lines. In *European Conference on Computer Vision*, volume 3, pages 189–200, Graz, Austria, May 2006.
- [14] R. Benosman. *Vision Omnidirectionnelle : Théorie, Capteurs et Applications*. Habilitation à diriger les recherches, Université Pierre et Marie Curie, Septembre 2006.
- [15] R. Benosman and J. Devars. *Panoramic Vision : Sensors, Theory and Applications*, chapter 9 : Panoramic Stereovision Sensor, pages 161–164. Springer Verlag, 2001.
- [16] R. Benosman and S. B. Kang, editors. *Panoramic Vision : Sensors, Theory and Applications*. Springer, 2001.
- [17] S. Betge Brezetz. *Modélisation Incrémentale et Localisation par Amers pour la Navigation d'un Robot Mobile Autonome en Environnement Naturel*. Thèse de doctorat, Université de Toulouse 3, 1996.
- [18] P. Biber, H. Andreasson, T. Duckett, and A. Schilling. 3d modeling of indoor environments by a mobile robot with a laser scanner and panoramic camera. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 4, pages 3430–3435, Sendai, Japan, September 2004.
- [19] S. Bigot-Marchand. *Outils de traitement d'images adaptés au traitement d'images omnidirectionnelles*. Thèse de doctorat, Université de Picardie Jules Verne, Octobre 2008.
- [20] S. Bosch. *Contribution à la modélisation d'environnements par vision monoculaire dans un contexte de robotique aéro-terrestre*. Thèse de doctorat, Université Toulouse III - Paul Sabatier, Octobre 2007.
- [21] E. Boyer. Détection de points d'intérêts - mise en correspondance. Cours de Master C4, UfR Informatique et Mathématiques Appliquées, 2008.
- [22] E. Brassart, L. Delahoche, C. Cauchois, C. Drocourt, C. Pegard, and E. M. Mouaddib. Experimental results got with the omnidirectional vision sensor : Syclop. In *IEEE Workshop on Omnidirectional Vision (in conjunction with IEEE Conference on Computer Vision and Pattern Recognition)*, pages 145–152, Hilton Head Island, USA, June 2000.

- 
- [23] R. Bunschoten and B. Krose. Robust scene reconstruction from an omnidirectional vision system. *IEEE Transactions on Robotics and Automation*, 19(2) :351–357, April 2003.
- [24] C. Cauchois. Calibration du capteur de vision omnidirectionnelle : Syclop. Rapport de dea, Université de Picardie Jules Verne, Août 1998.
- [25] C. Cauchois, E. Brassart, B. Marhic, and C. Drocourt. An absolute localization method using a synthetic panoramic image base. In *IEEE Workshop on Omnidirectional Vision (in conjunction with European Conference on Computer Vision)*, pages 128–135, Copenhagen, Danemark, June 2002.
- [26] C. Cauchois, E. Brassart, C. Pegard, and A. Clerentin. Technique for calibrating an omnidirectional sensor. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 1, pages 166–171, Korea, 1999.
- [27] E. Cecchi. *Reconstruction 3D pour la Volcanologie : Apports d'une Méthode Multi-Vues par Photogrammétrie Numérique*. Thèse de doctorat, Université Clermont-Ferrand II, 2003.
- [28] C. A. Chambon S. Mesures de corrélation pour des images couleur. *Traitement du Signal*, 21(6) :635–659, 2004.
- [29] J. L. Charron. Mesures sans contact : Méthodes magnétiques et capacitives. In *Techniques de l'Ingénieur, Mesures Mécaniques et Dimensionnelles*. Techniques de l'Ingénieur, 2003.
- [30] J. L. Charron. Mesures sans contact : Autres méthodes. In *Techniques de l'Ingénieur, Mesures Mécaniques et Dimensionnelles*. Techniques de l'Ingénieur, 2004.
- [31] J. L. Charron. Mesures sans contact : Méthodes optiques (partie 2). In *Techniques de l'Ingénieur, Mesures Mécaniques et Dimensionnelles*. Techniques de l'Ingénieur, 2004.
- [32] F. Comby, O. Strauss, and C. Caderas de Kerleau. Etalonnage de caméras catadioptriques hyperboloïdes. *Traitement du Signal*, 22(5) :419–431, Décembre 2005.
- [33] P. I. Corke. A robotics toolbox for matlab. *IEEE Robotics and Automation Magazine*, 3(1) :24–32, March 1996.
- [34] F. Courteille. *Vision Monoculaire : Contributions Théoriques et Application à la Numérisation des Documents*. PhD thesis, Université Paul Sabatier Toulouse III, 2006.
- [35] P. Cubaud, J. F. Haas, and A. Topol. Numérisation 3d de documents par photogrammétrie. In *Colloque International Francophone sur l'Ecrit et le Document*, page 6, 2004.

- 
- [36] W. Culbertson, T. Malzbender, and G. Slabaugh. Generalized voxel coloring. In *IEEE Workshop on Vision Algorithms Theory and Practice (in conjunction with IEEE International Conference on Computer Vision)*, pages 100–115, 1999.
- [37] K. Daniilidis and R. Klette. *Imaging Beyond the Pinhole Camera*. Springer, 2006.
- [38] A. J. Davison, I. Reid, M. N., and O. Stasse. Monoslam : Real-time single camera slam. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6) :1052–1067, June 2007.
- [39] D. Devernay. *Vision Stéréoscopique et Propriétés Différentielles des Surfaces*. Thèse de doctorat, Institut National Polytechnique, Février 1997.
- [40] C. Drocourt, L. Delahoche, C. Pegard, and A. Clerentin. Mobile robot localization based on an omnidirectional stereoscopic vision perception system. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1329–1334, Detroit, USA, May 1999.
- [41] C. R. Dyer. *Foundations of Image Understanding*, chapter 16 : Volumetric Scene Reconstruction from Multiple Views, pages 469–489. Kluwer, 2001.
- [42] T. Ea, O. Romain, C. Gastaud, and P. Garda. Un capteur de sphéro-vision stéréoscopique couleur. In *Congrès francophone de Vision par Ordinateur*, pages 301–308, Cahors, France, Juin 2001.
- [43] P. Eisert, E. Steinbach, and B. Girod. Multi-hypothesis, volumetric reconstruction of 3-d objects from multiple calibrated camera views. In *International Conference of the Acoustics, Speech, and Signal Processing*, volume 6, pages 3509–3512, Phoenix, USA, March 1999.
- [44] S. F. El Hakim, J. A. Beraldin, and F. Blais. A comparative evaluation of the performance of passive and active 3-d vision systems. In *SPIE Conference on Digital Photogrammetry*, 1995.
- [45] J. Fabrizio, J. Tarel, and R. Benosman. Calibration of panoramic catadioptric sensors made easier. In *IEEE Workshop on Omnidirectional Vision (in conjunction with European Conference on Computer Vision)*, pages 45–52, Copenhagen, Denmark, June 2002.
- [46] A. Fitzgibbon and R. B. Fisher. A buyer’s guide to conic fitting. In *British Machine Vision Conference*, volume 2, pages 513–522, Birmingham, England, 1995.
- [47] A. Fitzgibbon, M. Pilu, and R. B. Fisher. Direct least square fitting of ellipses. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 21(5) :476–480, May 1999.

- 
- [48] D. Fofi. *Navigation d'un Véhicule Intelligent à l'aide d'un Capteur de Vision en Lumière Structurée et Codée*. Thèse de doctorat, Université de Picardie Jules Verne, 2001.
- [49] A. Fusiello, E. Trucco, and A. Verri. A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, 12(1) :16–22, March 2000.
- [50] T. Gandhi and M. M. Trivedi. Vehicle mounted wide fov stereo for traffic and pedestrian detection. In *IEEE International Conference on Image Processing*, volume 2, pages 121–124, September 2005.
- [51] L. Gatet. *Intégration de Réseaux de Neurones pour la Télémétrie Laser*. Thèse de doctorat, Institut National Polytechnique de Toulouse, Septembre 2007.
- [52] C. Geyer. *Catadioptric Projective Geometry : Theory and Applications*. Phd, University of Pennsylvania, 2003.
- [53] C. Geyer and K. Daniilidis. Catadioptric camera calibration. In *IEEE International Conference on Computer Vision*, volume 1, pages 398–404, Kerkyra, Greece, September 1999.
- [54] C. Geyer and K. Daniilidis. Paracatadioptric camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5) :687–695, 2002.
- [55] C. Geyer and K. Daniilidis. Properties of the catadioptric fundamental matrix. In *European Conference on Computer Vision*, pages 140–154, 2002.
- [56] C. Geyer and K. Daniilidis. Conformal rectification of omnidirectional stereo pairs. In *IEEE Workshop on Omnidirectional Vision (in conjunction with IEEE Conference on Computer Vision and Pattern Recognition)*, volume 7, pages 73–78, Madison, USA, June 2003.
- [57] C. Geyer and K. Daniilidis. Mirrors in motion : Epipolar geometry and motion estimation. In *IEEE International Conference on Computer Vision*, 2003.
- [58] J. Gluckman and S. K. Nayar. Rectified catadioptric stereo sensors. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, pages 380–387, 2000.
- [59] J. Gluckman, S. K. Nayar, and K. K. Thorek. Real-time omnidirectional and panoramic stereo. In *Workshop of Image Understanding*, 1998.
- [60] N. Goncalves and H. Araujo. Estimating parameters of non-central catadioptric systems using bundle adjustment. In *IEEE Workshop on Omnidirectional Vision, Camera Networks and Non classical cameras (in conjunction with IEEE International Conference on Computer Vision)*, page 8, Beijing, China, October 2005.

- 
- [61] J. J. Gonzalez-Barbosa. *Vision Panoramique pour la Robotique Mobile : Stéréovision et Localisation par Indexation d'Images*. Thèse de doctorat, Université Paul Sabatier Toulouse III, Janvier 2004.
- [62] J. J. Gonzalez Barbosa and S. Lacroix. Stéréovision panoramique dense : Calibrage précis et rectification d'images panoramiques. *Traitement du Signal*, 22(5) :497–512, Décembre 2005.
- [63] J. Greenleaf, T. Tu, and E. Wood. Computer generated 3-d oscilloscopic images and associated techniques for display and study of the spatial distribution of pulmonary blood flow. *IEEE Transactions on Nuclear Science*, 17(3) :353–359, 1970.
- [64] K. Gremban, C. Thorpe, and T. Kanade. Geometric camera calibration using systems of linear equations. In *IEEE Conference on Robotics and Automation*, volume 1, pages 562–567, Philadelphia, USA, Avril 1988.
- [65] C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey Vision Conference*, pages 147–151, Manchester, England, September 1988.
- [66] R. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2) :146–157, November 1997.
- [67] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2004.
- [68] R. A. Hicks. Designing a mirror to realize a given projection. *Journal of the Optical Society of America*, 22(2) :323–330, 2005.
- [69] R. Horaud and O. Monga. *Vision par Ordinateur : Outils Fondamentaux*. Traité des Nouvelles Technologies, 1995.
- [70] B. K. P. Horn. *Shape From Shading : a Method for Obtaining the Shape of a Smooth Opaque Object from One View*. Phd, Massachusetts Institute of Technology, 1970.
- [71] K. Huang and M. M. Trivedi. Driver head pose and view estimation with single omnidirectional video stream. In *International Workshop on In-Vehicle Cognitive Computer Vision Systems*, Graz, Austria, April 2003.
- [72] S. H. Ieng and R. Benosman. Les surfaces caustiques par la géométrie. applications aux capteurs catadioptriques. *Traitement du Signal*, 22(5) :433–442, Décembre 2005.
- [73] H. J. *Calcul des probabilités*, chapter Analyse statistique des dépendances, pages 321–376. 2004.
- [74] F. Jacquy. *Traitements d'Images Omnidirectionnelles*. Thèse de doctorat, Université Montpellier II, Décembre 2007.

- 
- [75] G. Jang, S. Kim, and I. Kweon. Single camera catadioptric stereo system. In *IEEE Workshop on Omnidirectional Vision, Camera Networks and Non classical cameras (in conjunction with IEEE International Conference on Computer Vision)*, page 7, Beijing, China, October 2005.
- [76] S. Kang. Catadioptric self-calibration. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 201–207, Hilton Head Island, USA, June 2000.
- [77] K. Kutulakos and S. Seitz. A theory of shape by space carving. *International Journal of Computer Vision*, 38(3) :199–218, July 2000.
- [78] Y. Kuzu and O. Sinram. Volumetric modelling using voxel coloring. In *International Congress*, Washington DC, 2002.
- [79] D. Lanman, D. Crispell, M. Wachs, and G. Taubin. Spherical catadioptric arrays : Construction, multi-view geometry and calibration. In *International Symposium on 3D Data Processing, Visualization and Transmission*, pages 81–88, Chapel Hill, USA, June 2006.
- [80] P. Lasserre. *Vision pour la Robotique Mobile en Environnement Naturel*. Thèse de doctorat, Université Paul Sabatier, Septembre 1996.
- [81] M. Laurent. Acquisition au service de l’industrie 3d. *J’Automatise*, 19 :28–32, 1997.
- [82] C. Laugeau. Perception géométrique tridimensionnelle en robotique. In *Techniques de l’Ingénieur*, volume 3, pages 1–10. Techniques de l’Ingénieur, 1997.
- [83] J. M. Lavest. Calibrage des systèmes de vision. Cours de master, Ecole Normale Supérieure des Télécommunications, Mai 2005.
- [84] J. F. Layerle, X. Savatier, J. Y. Ertaud, and E. M. Mouaddib. Catadioptric sensor for a simultaneous tracking of the driver’s face and the road scene. In *IEEE Workshop on Omnidirectional Vision (in conjunction with European Conference on Computer Vision)*, page 7, Marseille, France, October 2008.
- [85] J. F. Layerle, X. Savatier, E. M. Mouaddib, and J. Y. Ertaud. Catadioptric vision system for an optimal observation of the driver face and the road scene. In *IEEE Intelligent Vehicles Symposium*, Eindhoven, Netherlands, June 2008.
- [86] V. Lemonde. *Stéréovision Embarquée sur Véhicule : de l’Auto-Calibrage à la Détection d’Obstacles*. Thèse de doctorat, Institut National des Sciences Appliquées de Toulouse, Novembre 2005.
- [87] K. Levenberg. A method for the solution of certain non-linear problems in least squares. *The Quarterly of Applied Mathematics*, 2 :164–168, 1944.

- 
- [88] M. Lhuillier. Reconstruction 3d automatique avec une caméra catadioptrique. In *Congrès francophone AFRIF-AFIA, Reconnaissance des Formes et Intelligence Artificielle*, Tours, France, Janvier 2006.
- [89] A. Loh. *The Recovery of 3D Structure Using Visual Texture Patterns*. Phd, University of Western Australia, February 2006.
- [90] A. M. Loh and R. Hartley. Shape from non-homogeneous, non-stationary, anisotropic, perspective texture. In *British Machine Vision Conference*, Oxford, England, September 2005.
- [91] C. Loop and Z. Zhang. Computing rectifying homographies for stereo vision. Technical report TR99-21, Microsoft Research, April 1999.
- [92] M. I. A. Lourakis. A brief description of the levenberg-marquardt algorithm implemented by levmar. February 2005.
- [93] Q. Luong and O. Faugeras. The fundamental matrix : Theory, algorithms, and stability analysis. *International Journal of Computer Vision*, 17(1) :43–75, January 1996.
- [94] Q. T. Luong, R. Deriche, O. Faugeras, and T. Papadopoulos. On determining the fundamental matrix : analysis of different methods and experimental results. Research Report RR-1894, INRIA, May 1993.
- [95] G. L. Mariottini and D. Prattichizzo. EGT for multiple view geometry and visual servoing. robotics and vision with pinhole and panoramic cameras. *IEEE Robotics and Automation Magazine*, 12(4) :26–39, December 2005.
- [96] C. Mc Glone, E. Mikhail, J. Bethel, and A. Editors. *Manual of Photogrammetry, 5th Ed.* Asprs, 2004.
- [97] C. Mei. *Couplage Vision Omnidirectionnelle et Télémétrie Laser pour la Navigation en Robotique*. Thèse de doctorat, Ecole Nationale des Mines de Paris, Février 2007.
- [98] C. Mei and P. Rives. Calibrage non biaisé d’un capteur central catadioptrique. In *Congrès francophone AFRIF-AFIA, Reconnaissance des Formes et Intelligence Artificielle*, Tours, France, 2006.
- [99] C. Mei and P. Rives. Single view point omnidirectional camera calibration from planar grids. In *IEEE International Conference on Robotics and Automation*, pages 3945–3950, Roma, Italia, April 2007.
- [100] B. Micusik. *Two-View Geometry of Omnidirectional Cameras*. Phd, Czech Technical University, June 2004.

- 
- [101] B. Micusik and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 485–490, Madison, USA, June 2003.
- [102] B. Micusik and T. Pajdla. Autocalibration & 3d reconstruction with non-central catadioptric cameras. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 1, pages 58–65, June 2004.
- [103] B. Micusik and T. Pajdla. Structure from motion with wide circular field of view cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7) :1135–1149, July 2006.
- [104] R. Molana and C. Geyer. *Motion Estimation with Essential and Generalized Essential Matrices*, chapter 2 : Motion, pages 107–125. Springer, 2006.
- [105] D. Moldovan and T. Wada. A geometrically calibrated pinhole camera model for single viewpoint omnidirectional imaging systems. In *British Machine Vision Conference*, page 9, London, England, September 2004.
- [106] E. M. Mouaddib. Introduction à la vision panoramique. *Traitement du Signal*, 22(5) :409–417, Décembre 2005.
- [107] E. M. Mouaddib, F. Comby, S. H. Ieng, C. Demonceaux, H. Hadj Abdelkader, P. Sturm, J. J. Gonzalez Barbosa, L. Smadja, J. Fabrizio, and O. Romain. Vision omnidirectionnelle. *Traitement du Signal*, 22(5) :407–562, Décembre 2005.
- [108] E. M. Mouaddib, G. Dequen, and L. Devendeville. A new omnidirectional stereo-vision sensor. In *IEEE Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras (in conjunction with IEEE International Conference on Computer Vision)*, pages 1–6, Rio De Janeiro, Brasil, October 2007.
- [109] D. Murray and J. Little. Using real-time stereo vision for mobile robot navigation. *Autonomous Robots*, 8(2) :161–171, April 2000.
- [110] V. Nalwa. A true omnidirectional viewer. Technical memorandum 0115500-960115-01, Bell Laboratories, February 1996.
- [111] S. K. Nayar. Catadioptric omnidirectional camera. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 482–488, San Juan, Puerto Rico, Juin 1997.
- [112] D. Nister. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26 :756–770, June 2004.
- [113] D. Nitzan. Three dimensional vision structure for robot applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(3) :291–309, May 1988.

- 
- [114] O. Ozun, U. Yilmaz, and V. Atalay. Comparison of photoconsistency measures used in voxel coloring. In *ISPRS Workshop (in conjunction with IEEE International Conference on Computer Vision)*, 2005.
- [115] J. Pagès and J. Salvi. Coded light projection techniques for 3d reconstruction. *Journal sur l'enseignement des sciences et technologies de l'information et des systèmes*, 4(3) :8, November 2005.
- [116] J. Pagès, J. Salvi, and C. Matabosch. Implementation of a robust coded structured light technique for dynamic 3d measurements. In *IEEE International Conference on Image Processing*, volume 2, pages 1073–1076, Barcelona, Spain, September 2003.
- [117] T. Pajdla. Epipolar geometry of some non-classical cameras. In *Winter Workshop on Computer Vision*, pages 159–180, February 2001.
- [118] T. Pajdla, T. Svoboda, and V. Hlavac. *Panoramic Vision : Sensors, Theory and Applications*, chapter 5 : Epipolar Geometry of Central Panoramic Cameras, pages 85–114. Springer Verlag, 2001.
- [119] J. P. Perez. *Optique Géométrique et Ondulatoire*. Masson, 1995.
- [120] E. Pervin and J.-A. Webb. Quaternions for computer vision and robotics. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 382–383, 1983.
- [121] T. Peynot. *Sélection et contrôle de modes de déplacement pour un robot mobile autonome en environnements naturels*. Thèse de doctorat, Institut National Polytechnique de Toulouse, Juillet 2006.
- [122] E. Prados. Brève introduction aux méthodes variationnelles. application au problème du "shape from shading". Cours de vision 3d, Université de Nice Sophia-Antipolis, 2001.
- [123] E. Prados and 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.
- [124] E. Prados and O. Faugeras. *Mathematical Models of Computer Vision : The Handbook*, chapter 5 : Shape From Shading, pages 377–390. Springer Verlag, 2005.
- [125] N. Ragot, J. Y. Ertaud, X. Savatier, and B. Mazari. Modelling of the environment of a mobile robot by catadioptric sensors. state of the art and proposal for an innovative architecture. In *International Conference on Modelling and Simulation*, Rouen, France, July 2005.
- [126] N. Ragot, J. Y. Ertaud, X. Savatier, and B. Mazari. Modelling of the environment of a mobile robot by catadioptric sensors. state of the art and proposal for an innovative architecture. *Best ok book AMSE*, 1(1) :23–33, 2005.

- 
- [127] N. Ragot, J. Y. Ertaud, X. Savatier, and B. Mazari. Calibration of a panoramic stereovision sensor : Analytical vs interpolation-based methods. In *IEEE Conference of the Industrial Electronics Society*, volume 1, pages 4130–4135, Paris, France, November 2006.
- [128] N. Ragot, J. Y. Ertaud, X. Savatier, and B. Mazari. An efficient catadioptric sensor calibration based on a low-cost test-pattern. In *International Conference on Computer Vision Theory and Applications*, volume 1, pages 11–18, Setubal, Portugal, February 2006.
- [129] N. Ragot, R. Rossi, X. Savatier, J. Y. Ertaud, and B. Mazari. 3d volumetric reconstruction with a catadioptric stereovision sensor. In *IEEE International Symposium on Industrial Electronics*, volume 1, pages 1306–1311, Cambridge, England, July 2008.
- [130] S. Ramalingam, P. Sturm, and S. Lodha. Theory and calibration algorithms for axial cameras. In *Asian Conference on Computer Vision*, volume I, pages 704–713, Hyderabad, Inde, January 2006.
- [131] D. Rees. Panoramic television viewing system. *US Patent*, 3.505.465 :7, April 1970.
- [132] H. Rey. *Utilisation de la Modélisation 3D pour l'Analyse et la Simulation du Développement et de la Croissance Végétative d'une Plante de Tournesol en Conditions Environnementales Fluctuantes (température et rayonnement)*. Thèse de doctorat, Ecole Nationale Supérieure d'Agronomie de Montpellier, 2003.
- [133] O. Romain, T. Ea, C. Gastaud, and P. Garda. Un capteur multi-spectral de vision panoramique 3d. In *Congrès des Jeunes Chercheurs en Vision par Ordinateur*, page 8, Cahors, France, June 2001.
- [134] E. Salamin. Application of quaternions to computation with rotations. Working Paper, 1979.
- [135] X. Savatier. *Etude et Mise en Oeuvre d'un Système d'Observation de la Qualité Locomotrice chez l'Homme*. Thèse de doctorat, Université de Rouen, Juin 2003.
- [136] D. Scaramuzza. *Omnidirectional Vision : from Calibration to Robot Motion Estimation*. Phd, Swiss Federal Institute of Technology Zurich, February 2008.
- [137] D. Scaramuzza, A. Martinelli, and R. Siegwart. A flexible technique for accurate omnidirectional camera calibration and structure from motion. In *IEEE International Conference on Computer Vision Systems*, pages 45–52, New York, USA, January 2006.

- 
- [138] S. Seitz and C. Dyer. Photorealistic scene reconstruction by voxel coloring. In *IEEE Computer Vision and Pattern Recognition*, pages 1067–1073, San Juan, Puerto Rico, June 1997.
- [139] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *IEEE Conference on Vision and Pattern Recognition*, volume 1, pages 519–528, 2006.
- [140] S. M. Seitz and C. R. Dyer. Photorealistic scene reconstruction by voxel coloring. *International Journal of Computer Vision*, 35(2) :151–173, November 1999.
- [141] J. Shi and C. Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600, Seattle, USA, June 1994.
- [142] T. Skordas. Une revue des derniers progrès en autocalibration de caméras ccd. *Traitement du Signal*, 12(1) :31–59, 1995.
- [143] G. Slabaugh, B. Culbertson, T. Malzbender, and R. Schafer. A survey of methods for volumetric scene reconstruction from photographs. In *International Workshop on Volume Graphics*, pages 21–22, June 2001.
- [144] G. Slabaugh, W. Culbertson, T. Malzbender, M. Stevens, and R. Schafer. Methods for volumetric reconstruction of visual scenes. *International Journal of Computer Vision*, 57(3) :179–199, May 2003.
- [145] L. Smadja, R. Benosman, and J. Devars. Cylindrical sensor calibration using lines. In *IEEE International Conference on Image Processing*, volume 3, pages 1851–1854, Singapore, October 2004.
- [146] L. Smadja, R. Benosman, and J. Devars. Hybrid stereo configurations through a cylindrical sensor calibration. *Machine Vision and Applications*, 17(4) :251–264, September 2006.
- [147] L. Spacek. Omnidirectional perspective and stereopsis with conical mirrors. Research report CTU-CMP-2003-12, Czech Technical University, June 2003.
- [148] E. Steinbach, B. Girod, P. Eisert, and A. Betz. 3-d reconstruction of real-world objects using extended voxels. In *International Conference on Image Processing*, volume 1, pages 569–572, Vancouver, Canada, September 2000.
- [149] P. Sturm. *Vision non Calibrée : Contributions à la Reconstruction Projective et Etude des Mouvements Critiques pour l'Auto-Calibrage*. Thèse de doctorat, Institut National Polytechnique de Grenoble, 1992.
- [150] P. Sturm. Calibrage de caméras à des fins métrologiques. In *Journée GdR ISIS - Métrologie 3D par vision*, Lyon, France, Mars 2008.

- 
- [151] P. Sturm and S. Ramalingam. A generic calibration concept : Theory and algorithms. Rapport de recherche RR-5058, Institut National de Recherche en Informatique et en Automatique, December 2003.
- [152] P. Sturm and S. Ramalingam. A generic concept for camera calibration. In *European Conference on Computer Vision*, volume 2, pages 1–13, Prague, Czech Republic, May 2004.
- [153] B. J. Super and A. C. Bovik. Shape from texture using local spectral moments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(4) :333–343, April 1995.
- [154] T. Svoboda. *Central Panoramic Cameras : Design, Geometry, Egomotion*. Phd, Czech Technical University, 1999.
- [155] T. Svoboda and T. Pajdla. Epipolar geometry for central catadioptric cameras. *International Journal of Computer Vision*, 49(1) :23–37, August 2002.
- [156] R. Swaminathan, M. D. Grossberg, and S. K. Nayar. Caustics of catadioptric cameras. In *IEEE International Conference on Computer Vision*, volume 2, pages 2–9, Vancouver, Canada, July 2001.
- [157] J. P. Tarel. Calibrage radiométrique de caméra. Rapport de recherche RR-2509, Institut National de Recherche en Informatique et en Automatique, Mars 1995.
- [158] R. Tebourbi and Z. Belhadj. Rectification épipolaire d’images stéréoscopiques. In *International Symposium on Image/Video Communications over Fixed and Mobile Networks*, page 4, Brest, France, Juillet 2004.
- [159] L. Torresani, A. Hertzmann, and C. Bregler. Non-rigid structure-from-motion : Estimating shape and motion with hierarchical priors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5) :878–892, May 2008.
- [160] B. Vandeportaele. *Contributions à la Vision Omnidirectionnelle : Etude, Conception et Etalonnage de Capteurs pour l’Acquisition d’Images et la Modélisation 3D*. Thèse de doctorat, Institut National Polytechnique de Toulouse, Décembre 2006.
- [161] B. Vandeportaele, M. Cattoen, P. Marthon, and P. Gurdjos. A new linear calibration method for paracatadioptric cameras. In *International Conference on Pattern Recognition*, volume 4, pages 647–651, Hong Kong, China, August 2006.
- [162] B. Vandeportaele, C. Dehais, M. Cattoen, and P. Marton. Orient-cam, a camera that knows its orientation and some applications. In *11th Iberoamerican Congress on Pattern Recognition*, page 10, 2006.

- 
- [163] J. M. Vezien. *Techniques de Reconstruction Globale par Analyse de Paire d'Images Stéréoscopiques*. Thèse de doctorat, Université Paris 7, Avril 1995.
- [164] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10) :965–980, October 1992.
- [165] M. Wilczkowiak, P. Sturm, and E. Boyer. Calibrage et reconstruction à l'aide de parallélépipèdes et de parallélogrammes. In *Congrès francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle*, page 9, Angers, France, Janvier 2002.
- [166] N. Winters and J. Santos Victor. Mobile robot navigation using omni-directional vision. In *Irish Machine Vision and Image Processing Conference*, page 16, Dublin, Ireland, September 1999.
- [167] G. Xu and Z. Zhang. *Epipolar Geometry in Stereo, Motion and Object Recognition. A Unified Approach*. Springer, 1996.
- [168] Y. Yagi and S. Kawato. Panorama scene analysis with conic projection. In *IEEE International Workshop on Intelligent Robots and Systems*, volume 1, pages 181–187, Ibaraki, Japan, July 1990.
- [169] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional imaging with hyperboloidal projection. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, volume 2, pages 1029–1034, Yokohama, Japan, July 1993.
- [170] X. Ying and Z. Hu. Catadioptric camera calibration using geometric invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(10) :1260–1271, October 2004.
- [171] M. L. Zani. La photogrammétrie anime la mesure 3d en atelier. *Mesures*, 770 :31–34, December 2004.
- [172] R. Zhang, P. S. Tsai, J. E. Cryer, and M. Shah. Shape from shading : a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8) :690–706, August 1999.
- [173] Z. Zhang. Parameter estimation techniques : A tutorial with application to conic fitting. Research report RR-2676, Institut National de Recherche en Informatique et en Automatique, October 1995.
- [174] Z. Zhang. Determining the epipolar geometry and its uncertainty : A review. *International Journal of Computer Vision*, 27(2) :161–195, March 1998.

- [175] Z. Zivkovic and O. Booij. How did we built our hyperbolic mirror omnidirectional camera - practical issues and basic geometry. Technical report IAS-UVA-05-04, University of Amsterdam, April 2005.

# Annexe A

## Modèle sténopé

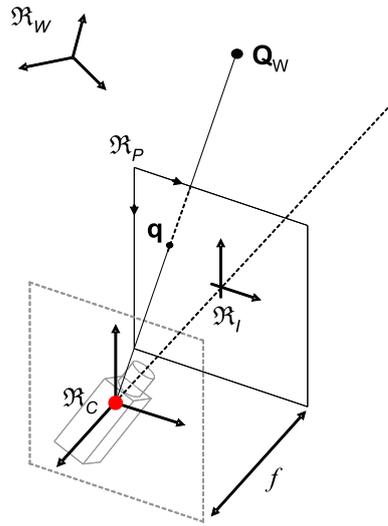


FIG. A.1 – Le modèle sténopé détaillé

Soit  $Q_W = (X_W, Y_W, Z_W, 1)^T$  un point 3D en coordonnées homogènes exprimé dans un référentiel monde  $\mathfrak{R}_W$ . Soit  $Q_C = (X_C, Y_C, Z_C, 1)^T$  le point  $Q_W$  exprimé dans le repère caméra  $\mathfrak{R}_C$ . Les vecteurs de  $\mathfrak{R}_W$  forment une base orthonormée directe et l'origine du repère est choisie de manière arbitraire.  $\mathfrak{R}_C$  est un trièdre directe pour lequel l'origine  $O_C$  est localisée en  $C$ , le point principal de la caméra, et l'axe  $Z_C$  est confondu avec l'axe optique (cf. figure A.1). La transformation entre ces deux points est un changement de base décrit par la matrice  $S_C$ . En coordonnées homogènes la transformation est donnée par l'équation A.1.  $R_C$  est une matrice de rotation et  $t_C$  un vecteur de translation. Ces deux entités décrivent les paramètres extrinsèques au processus de prise d'images.

$$\mathbf{Q}_C = S_C \mathbf{Q}_W = \begin{pmatrix} R_C & \mathbf{t}_C \\ \mathbf{0}_{3 \times 1}^\top & 1 \end{pmatrix} \mathbf{Q}_W \quad (\text{A.1})$$

Soit  $\mathbf{q}_C = (x_C, y_C, z_C, 1)^\top$  et  $\mathbf{q}_I = (x_I, y_I, 1)^\top$  les projetés de  $\mathbf{Q}_C$  exprimés respectivement dans le référentiel caméra  $\mathfrak{R}_C$  et le référentiel image  $\mathfrak{R}_I$ . Une étude rapide dans les deux plans de coupe, respectivement  $(\mathbf{Y}_C \mathbf{O}_C \mathbf{Z}_C)$  et  $(\mathbf{X}_C \mathbf{O}_C \mathbf{Y}_C)$ , en appliquant le théorème de Thalès, conduit à définir les coordonnées de  $\mathbf{q}_C$  et  $\mathbf{q}_I$  comme suit :

$$x_C = x_I = f \frac{X_C}{Z_C}, \quad y_C = y_I = f \frac{Y_C}{Z_C}, \quad z_C = f \quad (\text{A.2})$$

→ le passage de  $\mathbf{Q}_C$  à  $\mathbf{q}_C$  s'écrit alors comme suit :

$$s \begin{pmatrix} x_C \\ y_C \\ z_C \\ 1 \end{pmatrix} = P_C \mathbf{Q}_C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \mathbf{Q}_C \quad (\text{A.3})$$

Avec  $s = Z_C/f$  un facteur d'échelle et  $P_C$  est définie comme la matrice homogène de projection perspective lorsque  $\mathbf{Q}_C$  et  $\mathbf{q}_C$  sont définis dans  $\mathfrak{R}_C$ .

→ le passage de  $\mathbf{Q}_C$  à  $\mathbf{q}_I$  s'écrit alors comme suit :

$$s \begin{pmatrix} x_I \\ y_I \\ 1 \end{pmatrix} = P_I \mathbf{Q}_C = \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{Q}_C \quad (\text{A.4})$$

Avec  $s = Z_C$  un facteur d'échelle et  $P_I$  est définie comme la matrice homogène de projection perspective lorsque  $\mathbf{Q}_C$  est défini dans  $\mathfrak{R}_C$  et  $\mathbf{q}_I$  est défini dans  $\mathfrak{R}_I$ .

Soit  $\mathbf{q} = (u, v, 1)^\top$  le point 2D en coordonnées homogènes exprimé dans un référentiel pixel  $\mathfrak{R}_P$  défini au coin supérieur gauche du plan image. Les vecteurs de  $\mathfrak{R}_P$  sont pris de sorte que  $\mathbf{v}$  soit parallèle à l'axe  $\mathbf{y}_I$  et de même sens, le vecteur  $\mathbf{u}$  est de sens inverse à  $\mathbf{x}_I$ .

Le passage de  $\mathbf{q}_I$  à  $\mathbf{q}$  est défini par une matrice  $K'$  (cf. équation A.5) qui fait intervenir des paramètres liés au nombre de pixels par unité de longueur sur chacun des axes ( $k_u, k_v$ ) et les coordonnées du centre de l'image  $\mathbf{c} = (u_0, v_0)^\top$  dans  $\mathfrak{R}_P$ . Certaines formulations introduisent un paramètre supplémentaire : le défaut d'orthogonalité des axes  $\mathbf{u}$  et  $\mathbf{v}$  paramétré par un angle  $\theta$ . Néanmoins et afin d'éviter une sur-paramétrisation des modèles, il est courant de considérer ces deux axes orthogonaux ( $\theta = \pi/2$ ).

$$\mathbf{q} = \mathbf{K}'\mathbf{q}_I = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & kv & v_0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{q}_I \quad (\text{A.5})$$

Soit  $\mathbf{K}$  la matrice homogène définie comme le produit matriciel de  $\mathbf{K}'$  et  $\mathbf{P}_I$  :

$$\mathbf{K}_P = \mathbf{K}'\mathbf{P}_I = \begin{pmatrix} fk_u & 0 & u_0 & 0 \\ 0 & fk_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (\text{A.6})$$

$\mathbf{K}_P$  est définie comme la matrice homogène des paramètres intrinsèques pour une projection perspective. Elle modélise les caractéristiques internes au phénomène de prises d'images. Elle permet d'associer à un point 3D  $\mathbf{Q}_C$  défini dans  $\mathfrak{R}_C$  un point 2D  $\mathbf{q}$  exprimé dans  $\mathfrak{R}_P$ . En coordonnées homogènes, la relation entre  $\mathbf{q}$  et  $\mathbf{Q}_C$  est alors donnée par :

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{K}_P \mathbf{Q}_C \quad (\text{A.7})$$

Une autre manière de représenter la relation A.7 consiste à écrire le problème comme suit :

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \mathbf{K}_P \mathbf{P}_I \mathbf{Q}_C = \begin{pmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{pmatrix} \quad (\text{A.8})$$

# Annexe B

## Modèle orthographique

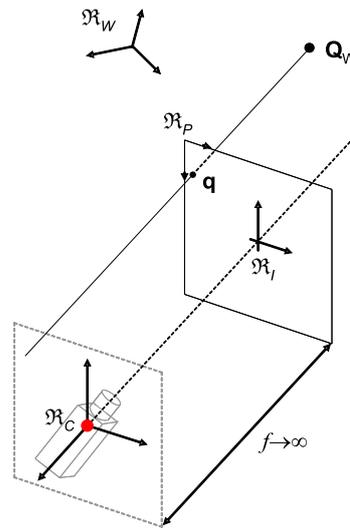


FIG. B.1 – Le modèle de projection orthographique détaillé

Le modèle de projection orthographique est un cas limite du modèle perspectif tel que  $f$  tende vers l'infini et par conséquent, que le rapport  $f/z$  tende vers 1. De l'équation A.2, le point 2D  $\mathbf{q}_I = (x_I, y_I, 1)^\top$  est alors défini par :

$$x_I = X_C, \quad y_I = Y_C \quad (\text{B.1})$$

Et sous forme matricielle homogène :

$$s \begin{pmatrix} x_I \\ y_I \\ 1 \end{pmatrix} = O_I \mathbf{Q}_C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{Q}_C \quad (\text{B.2})$$

$s = Z_C$  est un facteur d'échelle et  $O_I$  est définie comme la matrice homogène de projection orthographique.  $\mathbf{Q}_C$  est un point 3D défini dans le référentiel caméra  $\mathfrak{R}_C$ . Lorsque le point 3D est connu dans un référentiel monde  $\mathfrak{R}_W$ , une matrice de changement de base  $S_C$  analogue à celle présentée en l'équation A.1 permet d'exprimer ce point dans  $\mathfrak{R}_C$ .

Le passage de  $\mathbf{q}_I$  à  $\mathbf{q} = (u, v, 1)^\top$  est défini par la matrice  $K'$  introduite en l'équation A.5. La transformation de  $\mathbf{Q}_C$  en  $\mathbf{q}$  est donnée par :

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K' O_I \mathbf{Q}_C \quad (\text{B.3})$$

Le produit matriciel  $K' O_I$  définit la matrice homogène des paramètres intrinsèques  $K_O$  pour une projection orthographique telle que :

$$K_O = K' O_I = \begin{pmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} k_u & 0 & 0 & u_0 \\ 0 & k_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (\text{B.4})$$

Une autre formulation couramment rencontrée dans la littérature consiste à poser le problème sous la forme suivante [61] :

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K_O O_I \mathbf{Q}_C = \begin{pmatrix} k_u & 0 & 0 & u_0 \\ 0 & k_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{Q}_C \quad (\text{B.5})$$

## Annexe C

# Algorithme de Levenberg-Marquardt

L'algorithme de **Levenberg-Marquardt** est un standard en **Vision par Ordinateur** pour la résolution de problèmes de minimisation non linéaire au sens des moindres carrés [67], [73], [92], [86]. Il s'agit d'une technique itérative qui recherche le minimum d'une fonction coût multivariable définie comme la somme des carrés de fonctions non linéaires. L'algorithme de Levenberg-Marquardt combine la **méthode de descente de gradient** et celle de **Gauss-Newton**. Il s'agit tout d'abord, partant d'un point fixé arbitrairement, de se rapprocher aussi rapidement que possible du minimum de la fonction. Pour y parvenir, il est logique de suivre la ligne de plus grande pente. Pour une fonction monovariante, cette pente est donnée par la dérivée de la fonction. Lorsque la fonction dépend de plusieurs paramètres la dérivée est un vecteur, le gradient de la fonction, dont chaque élément est défini comme la dérivée partielle de la fonction selon l'un des paramètres. Ensuite, la méthode de Gauss-Newton est appliquée ce qui permet une convergence extrêmement rapide vers le minimum de la fonction, à condition que le vecteur initial des estimées soit proche du point à atteindre ; Si ce n'est pas le cas, la méthode ne converge plus du tout ou converge trop lentement pour être efficace [73].

Soit  $J$  la fonctionnelle qui associe un vecteur de paramètres  $\boldsymbol{\chi} = (\chi_1, \chi_2, \dots, \chi_i, \dots, \chi_I)^\top$  à un vecteur de mesures  $\mathbf{y} = (y_1, y_2, \dots, y_n, \dots, y_N)^\top$  telle que :  $\mathbf{y} = J(\boldsymbol{\chi})$ . Notons  $\boldsymbol{\chi}_0$  le vecteur initial des paramètres et  $\boldsymbol{\chi}^*$  le vecteur des paramètres minimisant la fonction coût  $J(\boldsymbol{\chi})$  définie comme la somme des distances au carré entre les mesures  $\mathbf{y}$  et les données estimées par le modèle  $\hat{\mathbf{y}}$  :

$$\begin{cases} J(\boldsymbol{\chi}) = \frac{1}{2} \boldsymbol{\epsilon}^\top \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} = \mathbf{y} - \hat{\mathbf{y}} \end{cases} \quad (\text{C.1})$$

(le terme en  $1/2$  permet de simplifier les calculs de la Jacobienne).

L'algorithme de Levenberg-Marquardt propose une approximation linéaire en **série de Taylor** de  $J$  au voisinage de  $\boldsymbol{\chi}$ . Pour une petite variation  $\Delta$ , la fonctionnelle  $J$  s'écrit :

$$\begin{cases} J(\boldsymbol{\chi} + \Delta) = J(\boldsymbol{\chi}) + \mathbf{G}\Delta \\ \mathbf{G} = \frac{\partial J(\boldsymbol{\chi})}{\partial \boldsymbol{\chi}} \end{cases} \quad (\text{C.2})$$

Avec  $\mathbf{G}$  qui définit la Jacobienne de la fonction coût, c'est-à-dire la dérivée de la fonction coût par rapport aux paramètres. L'algorithme de Levenberg-Marquardt est une méthode itérative qui débute le processus de minimisation d'une valeur initiale des paramètres  $\boldsymbol{\chi}_0$ , puis produit le vecteur  $\boldsymbol{\chi}_1$  et converge itérativement vers un minimum local de  $J$  noté  $\boldsymbol{\chi}^*$ . Pour chaque itération,  $\Delta$  est déterminé en résolvant une **équation de mise à jour** dite **équation normale augmentée** ou définie par :

$$(\mathbf{G}^\top \mathbf{G} + \lambda \mathbf{I}) \Delta = -\mathbf{G}^\top \boldsymbol{\epsilon} \quad (\text{C.3})$$

Cette équation diffère de l'**équation normale** ( $\mathbf{G}^\top \mathbf{G} \Delta = -\mathbf{G}^\top \boldsymbol{\epsilon}$ ) classiquement utilisée dans les méthodes de minimisation itérative newtonienne. Le terme en  $\lambda \mathbf{I}$  permet à l'algorithme de Levenberg-Marquardt d'avoir un comportement similaire à une méthode de Gauss-Newton ou à une méthode du Gradient. La valeur initiale  $\lambda_0$  est classiquement choisie telle que  $\lambda_0 = 10^{-3} \text{diag}(\mathbf{G}^\top \mathbf{G})$  et deux cas sont alors à envisager :

- Si la valeur de  $\Delta$  estimée en résolvant l'équation normale augmentée, entraîne une diminution de la fonction coût,  $\lambda$  est alors divisé par 10. Pour des faibles valeurs de  $\lambda$  le terme en  $\lambda \mathbf{I}$  ne devient plus significatif. L'équation normale augmentée est équivalente à l'équation normale et la méthode de minimisation s'approche à une méthode de Gauss-Newton.
- Si la valeur de  $\Delta$  conduit à une augmentation de la fonction coût,  $\lambda$  est alors multiplié par 10. Pour des valeurs importantes de  $\lambda$  le terme en  $\lambda \mathbf{I}$  est prépondérant par rapport au terme  $\mathbf{G}^\top \mathbf{G}$ . L'équation normale augmentée devient donc :  $\lambda \Delta = -\mathbf{G}^\top \boldsymbol{\epsilon}$ , qui est l'équation de mise à jour pour la méthode du Gradient.

## Annexe D

# Décomposition en valeurs singulières

La décomposition en valeurs singulières dont une abréviation courante est **SVD** (« *Singular Value Decomposition* ») est un outil mathématique particulièrement utilisé en **Vision par Ordinateur**. Elle permet par exemple d'orthonormaliser les matrices de rotation durant les processus de minimisation. Elle permet également de résoudre un système linéaire exact ou d'obtenir la solution au sens des moindres carrés d'un système linéaire sur-dimensionné [12], [20]. La décomposition en valeurs singulières est basée sur le théorème d'algèbre linéaire suivant : toute matrice  $A$  de taille  $(N \times M)$  telle que  $N > M$  peut être décomposée comme le produit suivant :

$$\begin{cases} A_{N \times M} = U_{N \times M} \Sigma_{M \times M} V_{M \times M}^T \\ \Sigma = \text{diag}_{M \times M} (\sigma_1, \sigma_2, \dots, \sigma_M) \end{cases} \quad (\text{D.1})$$

Où :

- $U$  possède des colonnes orthonormales ;
- $V$  est orthonormale et contient les vecteurs singuliers de  $A$  ;
- $\Sigma$  est diagonale et contient les valeurs singulières de  $A$ .

Et, une matrice  $U$  est orthonormale si elle satisfait :  $U^T U = I$

# Annexe E

## Normes de Minkowski

Soit  $\mathbf{x} = (x_1, x_2, \dots, x_i, \dots, x_I)^\top$  et  $\mathbf{y} = (y_1, y_2, \dots, y_i, \dots, y_I)^\top$  deux vecteurs tels que  $\mathbf{x} \in \mathbb{R}^I$  et  $\mathbf{y} \in \mathbb{R}^I$ . Les **normes de Minkowski** ou **distances de Minkowski** à l'ordre  $p$  des vecteurs  $\mathbf{x}$  et  $\mathbf{y}$ , également dénommées normes  $L^p$  sont données par :

→ la norme  $L^1$  est définie comme la différence en valeur absolue des coefficients :

$$\forall i \in \{1, \dots, I\}, \quad L^1 = \sum_{i=1}^I |x_i - y_i| \quad (\text{E.1})$$

→ la norme  $L^2$  est par définition la norme euclidienne :

$$\forall i \in \{1, \dots, I\}, \quad L^2 = \sqrt{\sum_{i=1}^I |x_i - y_i|^2} \quad (\text{E.2})$$

→ la norme  $L^p$  est une généralisation des normes  $L^1$ ,  $L^2$ , etc. :

$$\forall i \in \{1, \dots, I\}, \quad L^p = \sqrt[p]{\sum_{i=1}^I |x_i - y_i|^p} \quad (\text{E.3})$$

→ la norme  $L^\infty$  des vecteurs  $\mathbf{x}$  et  $\mathbf{y}$  est la limite de la norme  $L^p$  lorsque  $p$  tend vers l'infini :

$$\forall i \in \{1, \dots, I\}, \quad L^\infty = \max_{1 \leq i \leq I} (|x_1 - y_1|, \dots, |x_i - y_i|, \dots, |x_I - y_I|) \quad (\text{E.4})$$



## Résumé

Les travaux de thèse présentés dans ce mémoire ont pour objectif la conception et l'évaluation d'un système de vision 3D constitué de deux capteurs catadioptriques positionnés verticalement. Afin d'obtenir un modèle 3D, l'étalonnage des capteurs, c'est-à-dire l'estimation des relations entre l'espace 3D et le plan image, est un préalable indispensable. Une méthode, qualifiée de paramétrique a tout d'abord été implémentée. Elle utilise un modèle ad hoc et une mire 3D de forme cylindrique dont les points d'intérêt sont des diodes électroluminescentes. Cette méthode formule l'hypothèse d'un centre de projection unique qui peut s'avérer difficile à mettre oeuvre en pratique. Nous proposons donc une méthode non-paramétrique qui permet de s'affranchir des contraintes liées aux modèles. Les relations entre l'espace 3D et le plan image sont alors définies pour un nombre fini de points et des algorithmes d'interpolation numérique permettent d'approximer les fonctions de projection et rétro-projection pour tous les couples points 3D-points 2D. Nous proposons ensuite l'implémentation et l'évaluation d'une méthode de reconstruction 3D volumétrique qui modélise la scène sous forme d'un tableau 3D dont chaque cellule est un voxel. Une mesure de similarité colorimétrique des projections du voxel dans chaque une des images permet de statuer quant à l'appartenance du voxel à un objet de la scène. Cette méthode est évaluée pour une reconstruction statique de la scène 3D, c'est-à-dire pour une acquisition d'images simultanée. Puis cette méthode est étudiée pour des déplacements du capteur. Il s'agit d'une reconstruction dynamique définie comme l'intersection des reconstructions statiques réalisées à des instants successifs.

**MOTS-CLÉS** : vision omnidirectionnelle, stéréovision, capteur catadioptrique, étalonnage paramétrique et non-paramétrique, reconstruction 3D volumétrique

## Abstract

The work presented in this thesis deals with the design and the evaluation of a vision sensor made of two catadioptric sensors, one above the other. To obtain a 3D model, it is necessary to calibrate the sensors, that is to say to estimate the relationships between the 3D space and the image plane. A parametric method has been implemented. It uses an ad hoc model and a cylindrical 3D pattern whose points of interest are light-emitting diodes. This method is based on the assumption that the sensor respects the single effective viewpoint constraint, but which is difficult to put into practice. Thus, we propose a non-parametric method which eliminate the problem of model restrictions. The relationships are defined for a finite number of points and numerical interpolation algorithms allow us to approximate the projection and back-projection functions for all pairs of 3D points-2D points. Then, we propose the implementation and the evaluation of a 3D volumetric reconstruction method which models the scene by a 3D array whose cells are voxels. A colorimetric similarity measure of the voxel projections estimate whether or not the voxel is an object in 3D. This method is evaluated on a 3D static reconstruction, that is a simultaneous image acquisition. Then, this method is studied for sensor displacements. This is a 3D dynamic reconstruction which is defined as the intersection of 3D static reconstructions made at successive times.

**KEYWORDS** : omnidirectional vision, stereovision, catadioptric sensor, parametric and non-parametric calibration, 3D volumetric reconstruction