

Table des matières

Résumé	iii
Abstract.....	v
Table des matières.....	vii
Liste des tableaux	xii
Liste des figures	xiv
Dédicace.....	xvi
Remerciements	xviii
Introduction générale.....	1
Chapitre 1. Notions théoriques et revue de littérature.....	5
1.1 Les différents types d'objectifs.....	6
1.1.1 Les objectifs à focale fixe	6
1.1.2 Les objectifs à focale variable	7
1.1.3 Les objectifs grands angles.....	9
1.1.3.1 Objectifs catadioptriques.....	9
1.1.3.2 Objectif Fisheye et objectif panomorphe	10
1.2 La stéréovision	11
1.2.1 Définition.....	11
1.2.2 Calibrage de caméras.....	13
1.2.2.1 Calibrage des caméras panoramiques.....	13
1.2.2.2 Calibrage des caméras zoom	15
1.2.3.1 Géométrie épipolaire perspective	21
1.2.3.2 Géométrie épipolaire panoramique	21
1.2.4 Triangulation	22
1.3 Projection perspective	22
1.3.1 Modèle mathématique de la projection perspective.....	22
1.3.1.1 Transformation coordonnées/direction	23
1.3.1.2 Transformation direction/coordonnées normalisées.....	24
1.3.1.3 Transformation coordonnées normalisées/coordonnées image	24
1.3.2 Distorsions.....	25

1.3.2.1	Modèle de distorsions radiales et tangentielles	25
1.3.2.2	Modèle polynomial	26
1.3.2.3	Modèle logarithmique	27
1.3.2.4	Modèle de division	27
1.3.2.5	Modèle rationnel	27
1.3.3	Résolution du modèle	28
1.3.4	Mire de calibrage	29
	Problématique et objectifs.....	34
	Chapitre 2. Calibrage des caméras classiques à l'aide des éléments de l'optique diffractive	38
2.1	Lasers et réseaux de diffraction.....	39
2.1.1	Géométrie d'un faisceau laser	39
2.1.2	Réseaux de diffraction.....	40
2.1.3	Réseau de Dammann et éléments diffractifs	41
2.2	Banc de calibrage: DOE setup	42
2.3	Condition générale	43
2.4	Modèle mathématique de calibrage	44
2.5	Détection de points de contrôle dans l'image	46
2.5.1	Algorithme 1: Méthode binaire	46
2.5.2	Algorithme 2 : Méthode d'intensité	47
2.5.3	Comparaison des deux méthodes	48
2.6	Résultats pour des caméras aériennes.....	48
2.6.1	Conditions de travail	49
2.6.2	Caméras utilisées	49
2.6.3	Calcul des paramètres	49
2.7	Comparaison avec un objet de calibrage 3D	52
	Chapitre 3. Fisheye lens calibration using virtual grid.....	55
3.1	Résumé	56
3.2	Abstract.....	56
3.3	Introduction.....	56
3.4	Setup for fisheye calibration	59
3.5	Fisheye calibration with CDOE	60

3.5.1 Calibration pattern	60
3.5.2 Fisheye projection	61
3.5.3 Projection procedures	61
3.6 Calibration Algorithm.....	62
3.7 Results and Discussion	64
3.8 Conclusion	69
Chapitre 4. Calibration of zoom-lens with virtual optical pattern	72
4.1 Résumé.....	73
4.2 Abstract	73
4.3 Introduction	73
4.4 Cross-diffractive optical element setup	75
4.4.1 Grating Fabrication	75
4.4.2 Grating Characterization	76
4.4.3 Experimental Setup.....	77
4.5 Fixed camera model using CDOE.....	78
4.6 Mathematical model for zoom-lens camera	79
4.7 Calibration algorithm for zoom-lens camera.....	80
4.8 Results and discussion.....	81
4.8.1 Step 1: CDOE pattern and image recording	81
4.8.2 Step 2: Camera calibration with $L = 20$	82
4.8.3 Step 3: Fitting the models of internal parameters	83
4.8.4 Step 4: Global error optimization of the adjustable model	85
4.8.5 Step 5: Evaluation of the generality and repeatability of the calibration model ..	85
4.9 Conclusion	86
Chapitre 5. A new mathematical model for hybrid and panoramic stereovision systems: panoramic to rectilinear conversion model	89
5.1 Résumé	90
5.2 Abstract	90
5.3 Introduction	90
5.4 Mathematical formulation	93

5.4.1 <i>Classical model</i>	93
5.4.2 <i>Fisheye and Panomorph camera models</i>	95
5.4.3 <i>Composed cubic Splines functions (CCS functions)</i>	95
5.4.4 <i>Epipolar geometry and Fundamental matrix</i>	96
5.4.5 <i>Retroprojection and triangulation</i>	98
5.5 Stereovision calibration algorithm	98
5.6 Experimental results and discussion	100
5.7 Conclusions	105
Discussion générale, conclusions et perspectives de recherche.	107
Appendix A	111
Appendix B	112
Appendix C	113
Publications	115
Bibliographie	116

Liste des tableaux

TABLEAU 2.1 : TABLEAU COMPARATIF ENTRE LES DEUX MÉTHODES	48
TABLEAU 2.2 CARACTÉRISTIQUES DES CAMÉRAS UTILISÉES POUR TESTER LE BANC DE CALIBRAGE	49
TABLEAU 2.3 PARAMÈTRES INTRINSÈQUES DES DEUX CAMÉRAS.....	52
TABLEAU 2.4 TABLEAU COMPARATIF ENTRE LES DEUX TECHNIQUES DE CALIBRAGE	53
TABLEAU 3.1 CALIBRATION RESULTS USING THE EQUIDISTANT MODEL	66
TABLEAU 3.2 CALIBRATION RESULTS USING THE THIRD-ORDER MODEL	66
TABLEAU 3.3 CALIBRATION RESULTS USING THE FIFTH-ORDER MODEL	66
TABLEAU 3.4 CALIBRATION COMPARISON	69
TABLEAU 3.5 RMSE FOR FISHEYE CAMERA CALIBRATION USING CDOE AND 3D OBJECT	69
TABLEAU 4.1 SUBSTRATE SPECIFICATIONS.....	76
TABLEAU 4.2 PRINCIPAL SETUP COMPONENTS ILLUSTRATED AT FIGURE 4.2.....	78
TABLEAU 4.3 COEFFICIENTS OF POLYNOMIAL FUNCTIONS.....	85
TABLEAU 4.4 COMPARISON OF CALIBRATION PARAMETERS USING TWO DIFFERENT PATTERNS .	86
TABLEAU 5.1 INTRINSIC PARAMETERS FOR THE THREE CAMERAS.....	101
TABLEAU 5.2 RMSE FOR THE EPIPOLAR GEOMETRY COMPUTATION	104
TABLEAU 5.3 RMSE IN X, Y AND Z DIRECTIONS FOR THE THREE PAIRS.	104

Rapport-Gratuit.com

Liste des figures

FIGURE 1.1 CARACTÉRISTIQUES D'UN OBJECTIF	7
FIGURE 1.2. OBJECTIFS À FOCAL VARIABLE	8
FIGURE 1.3 OBJECTIF CATADIOPTRIQUE	10
FIGURE 1.4 OBJECTIFS RÉFRACTIFS	11
FIGURE 1.5 COMPARAISON ENTRE LE MODÈLE STÉNOPÉ ET LE MODÈLE DE LENTILLE ÉPAISSE...	17
FIGURE 1.6 MODÈLE DU CALIBRAGE PAR RÉSEAU DE NEURONES.....	19
FIGURE 1.7 PAN-TILT-ZOOM CAMERA	20
FIGURE 1.8 PROJECTION PERSPECTIVE	23
FIGURE 1.9 MOTIFS CLASSIQUES POUR CALIBRER UNE CAMÉRA	30
FIGURE 1.10 POINTS D'APPUI POUR CALIBRER DES CAMÉRAS AÉRIENNES.....	31
FIGURE 1.11 MOTIF SPÉCIAL POUR CALIBRER UNE CAMÉRA MUNIE D'UN ZOOM.....	31
FIGURE 2.1 GÉOMÉTRIE D'UN FAISCEAU LASER	39
FIGURE 2.2 RÉSEAU DE DIFFRACTION	41
FIGURE 2.3 EXEMPLE DE RÉSEAU BINAIRE OU RÉSEAU DE DAMMANN	41
FIGURE 2.4 BANC DE CALIBRAGE AVEC DES ÉLÉMENTS D'OPTIQUE DIFFRACTIVE MIS EN CASCADE	42
FIGURE 2.5 LE CHEVAUCHEMENT DES ORDRES DE DIFFRACTION DES DOES ET LA POSITION DE LA PUPILLE D'ENTRÉE DE LA CAMÉRA À CALIBRER.....	43
FIGURE 2.6 GRILLE FORMÉE PAR LES ÉLÉMENTS DIFFRACTIFS SUR UN CAPTEUR CCD	44
FIGURE 2.7 DÉTECTION D'UN CENTROÏDE PAR LA MÉTHODE BINAIRE	46
FIGURE 2.8 DÉTECTION D'UN CENTROÏDE PAR LA MÉTHODE D'INTENSITÉ.....	47
FIGURE 2.9 LES COORDONNÉES CALCULÉES (GRILLE PRINCIPALE) PAR LA MÉTHODE D'INTENSITÉ	50
FIGURE 2.10 LES COORDONNÉES CALCULÉES (GRILLE HYBRIDE) PAR LA MÉTHODE D'INTENSITÉ	51
FIGURE 2.11 MIRE D'ÉTALONNAGE DE L'UNIVERSITÉ LAVAL	53
FIGURE 3.1: CALIBRATION SETUP WITH ROTATABLE CAMERA	59
FIGURE 3.2 : GRID IMAGES RECORDED BY FISHEYE CAMERA IN 9 DIFFERENT ORIENTATIONS	60
FIGURE 3.3 FISHEYE PROJECTION.....	62
FIGURE 3.4 DIAGRAM OF CALIBRATION ALGORITHM	64
FIGURE 3.5 PIXEL COORDINATES FOR ALL IMAGES COMPUTED BY CENTROID ALGORITHM.....	65
FIGURE 3.6 REMAINING RESIDUALS FOR ALL IMAGES AFTER CALIBRATION (IN PIXELS)	67
FIGURE 3.7 TARGETS OF THE 3D CALIBRATION OBJECT	68
FIGURE 4.1 DIFFRACTIVE POWER DISTRIBUTION OF THE 33 DIFFRACTION ORDERS. THE ZERO- ORDER POWER IS LESS THAN THE DOUBLE OF THE REMAINING AVERAGE POWER FRACTION	77
FIGURE 4.2 SETUP OF THE PTZ CAMERA CALIBRATION USING CROSSED DIFFRACTIVE ELEMENTS	78
FIGURE 4.3 CDOE ORDERS CALIBRATION GRID (33X33).....	82
FIGURE 4.4 ZOOM SETTINGS (S_t) USED TO RECORD CALIBRATION DATA FOR $L=20$ (TOP). ZOOM SETTINGS (S_b) REQUIRED TO CALIBRATE THE ZOOM LENS FOR $L=10$ (BOTTOM).	82
FIGURE 4.5 INTERNAL CAMERA PARAMETERS VERSUS ZOOM SETTING S	84
FIGURE 4.6 POLYNOMIAL FIT OF THE INTERNAL CHARACTERISTICS OF THE CAMERA	84

FIGURE 5.1 RECTILINEAR AND EQUIDISTANT PROJECTION	94
FIGURE 5.2 EPIPOLAR GEOMETRY	97
FIGURE 5.3 GRID IMAGES RECORDED BY FISHEYE CAMERA IN 9 DIFFERENT ORIENTATIONS	99
FIGURE 5.4 DIAGRAM OF STEREOVISION CALIBRATION.....	100
FIGURE 5.5 STEREO PAIRS RECTIFICATIONS	102
FIGURE 5.6 GEOMETRY EPIPOLAR ESTIMATION.....	103

Dédicace

À mes parents, Aïd et Chedlia, pour votre amour, vos sacrifices. Que Dieu vous protège!

À la femme de ma vie, mon amour... À Khaoula. Merci chérie pour ton amour inconditionnel, ta présence à mes côtés dans les moments de joie et de peine. Merci pour ta patience et ton dévouement.

À mon fils Ahmed pour la joie que tu nous donnes. Grâce à vous je me lève tous les matins avec le sourire.

À ma sœur Leïla et mon frère Mourad. Merci pour votre encouragement tout au long de cette thèse. Merci d'être à côté de moi là où j'avais besoin. Merci pour les bons et inoubliables moments que nous avons passés ensemble.

À mes sœurs Amel et Malak, pour votre amour et votre encouragement sans limite malgré les distances qui nous séparent.

À mes beaux-parents Mounir et Henda et mon beau-frère Hamed pour votre encouragement et votre prière pour moi.

À tous mes professeurs qui m'ont appris à grimper l'échelle scientifique avec confiance, humilité et esprit critique.

À tout(e)s mes ami(e)s.

Remerciements

En tout premier lieu, je remercie Dieu, tout puissant, de m'avoir donné l'audace pour dépasser toutes les difficultés.

Et parce que « Celui qui ne remercie pas les gens ne remercie pas Dieu », j'aimerais remercier les nombreuses personnes qui ont rendu cette thèse possible.

Mes premiers remerciements vont d'abord au professeur Simon Thibault qui m'a offert l'opportunité de travailler dans son laboratoire de recherche. J'apprécie particulièrement ses encouragements, son support et ses précieux conseils qui m'ont été très utiles tout au long de ces cinq années. Son encadrement a été exemplaire et une excellente leçon d'apprentissage pour la suite de ma carrière.

Je remercie également Monsieur Michel Boulianne, qui a codirigé cette thèse, pour ses conseils judicieux et ses cours magistraux de photogrammétrie imprégnés de savoir et de connaissances.

Je tiens aussi à remercier Monsieur Ermanno F. Borra pour ses conseils avisés, particulièrement lors de la phase de la prélecture.

J'adresse aussi mes profonds remerciements à Monsieur Benoit Debaque pour avoir accepté avec enthousiasme d'être l'examineur externe de cette thèse, de même que Madame Sylvie Daniel et Monsieur Denis Laurendeau pour avoir accepté de former le reste du jury.

Je souhaiterais remercier également Monsieur Marc Cocard le directeur des études supérieures en sciences géomatiques qui m'a fait l'honneur de présider le Jury de cette thèse.

Mes sincères remerciements vont à Pierre Désaulniers et à tous les membres de l'équipe de recherche de Monsieur Simon Thibault.

Je tiens à remercier les personnes les plus importantes à mes yeux qui ont contribué par leur soutien et leur amour à la réussite de ce travail, ils s'agissent évidemment de ma famille, en particulier mon père Aid, ma mère Chedlia, mes deux sœurs Leila et Amel et mon frère Mourad, et de ma belle-famille.

Enfin, je voudrais adresser un remerciement tout particulier à l'amour de ma vie Khaoula qui m'a soutenu quotidiennement avec amour et patience pendant les moments les plus difficiles.

Introduction générale

De nos jours, la caméra est de plus en plus utilisée comme un outil précis de mesurage. Elle permet d'effectuer des mesures métriques et angulaires d'objets bidimensionnels et tridimensionnels contenus dans une scène. Le spectre d'utilisation des caméras couvre différentes applications comme la surveillance et/ou pour effectuer des relevés topographiques, concevoir des projets d'urbanisme, prospector de nouveaux gisements et observer des exploitations agricoles. Les caractéristiques et le type de la caméra utilisée dépendent de la nature de l'application et de propriétés de l'objet photographié.

La focale et le champ de vue sont les principales propriétés d'une caméra. La focale permet de distinguer les caméras classiques (ou standards), qui ont une focale fixe, des caméras à focale variable qui sont dotées d'un zoom, soit le grossissement apparent de l'image réalisé par le déplacement d'un bloc mobile de lentilles. Alors que le champ de vue permet de classer les caméras parmi les systèmes de vision classique ou panoramique. Contrairement à la vision classique, la vision panoramique est le procédé de vision qui fournit un hémisphère de vue du monde observé à partir de son centre.

Cette thèse s'intéresse à deux problématiques fondamentales en photogrammétrie et en vision numérique : le calibrage des caméras et la stéréovision.

Le calibrage d'une caméra est une étape importante pour de nombreux algorithmes et applications. C'est une opération qui consiste à estimer les paramètres intrinsèques et extrinsèques de la caméra comme la position et l'orientation, le centre optique, la distorsion et la distance principale.

Le calibrage d'une caméra munie d'un zoom par exemple doit prendre en compte deux modifications, soit la modification de la distance principale et celle de la distance entre le centre optique et l'objet. En ce qui concerne les caméras panoramiques, l'utilisation de certains modèles et algorithmes complexes complique la procédure de calibrage.

Deux caméras de caractéristiques identiques ou différentes peuvent être combinées pour former un système de stéréovision. La détermination d'un modèle qui décrit la géométrie de ce système est une grande problématique surtout lorsqu'une caméra panoramique est utilisée.

Afin de résoudre ces problèmes, cette thèse propose, en premier lieu, un nouveau banc de calibrage permettant, à l'aide des éléments d'optique diffractive, de calibrer toute sorte de caméras (classiques, munies d'un zoom, panoramiques). Elle présente ainsi un modèle simple et précis pour chaque caméra permettant d'extraire ses paramètres.

Ensuite, elle offre un nouveau modèle générique permettant de modéliser géométriquement, à l'aide d'une projection perspective et de splines cubiques composées, des systèmes de stéréovision mixte (formés par deux caméras, panoramique et classique) ou panoramiques (formés par deux caméras panoramiques) et de simplifier le calibrage d'une lentille aussi complexe que la lentille panomorphe.

Cette thèse comprend cinq chapitres, dont trois présentés sous forme d'articles scientifiques.

Le premier chapitre aborde successivement les concepts fondamentaux nécessaires pour comprendre les différents types de caméras et les notions théoriques de la stéréovision. Puis, il introduit le modèle de projection perspective (sténopé ou « pinhole »). Enfin, il présente les objectifs.

Le deuxième chapitre examine le banc de calibrage et le modèle mathématique permettant de calibrer les caméras classiques. En premier lieu, les caractéristiques théoriques d'un réseau de diffraction et d'un faisceau laser sont introduites. Dans un second lieu, le banc de calibrage et le modèle complet de projection sont décrits. Enfin, les résultats pour certaines caméras classiques sont présentés.

Le chapitre 3 s'intéresse à une méthode de calibrage des caméras panoramiques. Plus précisément, il propose une technique de calibrage d'une caméra fisheye à l'aide de notre banc d'étalonnage.

Dans le quatrième chapitre, une nouvelle méthode pour calibrer une caméra munie d'un zoom est offerte. Des nouveaux éléments d'optique diffractive qui ont été conçus spécialement dans LRIO (Laboratoire de recherche en Ingénierie Optique) pour calibrer ce type de caméras sont présentés.

Le chapitre 5 est consacré à la présentation d'un nouveau modèle générique qui permet de modéliser les caméras classiques et panoramiques. Il décrit les résultats de reconstruction tridimensionnelle pour trois systèmes de stéréovision: perspective/fisheye, perspective/panomorphe et fisheye/panomorphe.

Chapitre 1. Notions théoriques et revue de littérature

Ce chapitre est une revue de littérature qui présente :

- les caractéristiques de certains types de caméras.
- Les différentes techniques de calibrage
- Les différentes approches du problème de la stéréovision.

1.1 Les différents types d'objectifs

Quel que soit le mode d'enregistrement d'une image (prises de vue fixes ou vidéo) ou les types de captures d'images (analogique, numérique, couleur, monochrome), un objectif photographique est indispensable. L'objectif est un système convergent destiné à former une image réelle d'un objet sur une surface sensible (sels d'argent, capteur CCD ou CMOS,...). Il est l'élément qui permet de définir la qualité géométrique et chromatique de l'image.

1.1.1 Les objectifs à focale fixe

Malgré le grand potentiel d'une caméra zoom, la plupart des applications se limitent à l'utilisation des caméras à focale fixe (Mirisola et al., 2007, Schindler et al., 2008, Bennewitz et al., 2006, Börlin and Igasto, 2009). Cela est dû essentiellement à la difficulté de modéliser géométriquement une caméra avec des lentilles ajustables.

Un objectif est caractérisé par sa focale, son ouverture, son angle de champ ou champ de vue et ses paramètres de distorsion (Figure 1.1).

La focale est la distance à laquelle se forme l'image d'un point à l'infini, mesurée à partir du centre optique, soit son point focal émergent. Elle permet de distinguer les objectifs grands angulaires, ceux à focales moyennes ou standards et les téléobjectifs.

Le champ de vue désigne l'angle de prise de vue et permet de déterminer l'étendue de la scène photographiée. L'ouverture d'un objectif représente le rapport entre sa distance focale et le diamètre de sa pupille d'entrée et permet de déterminer la quantité de lumière qui atteint le capteur.

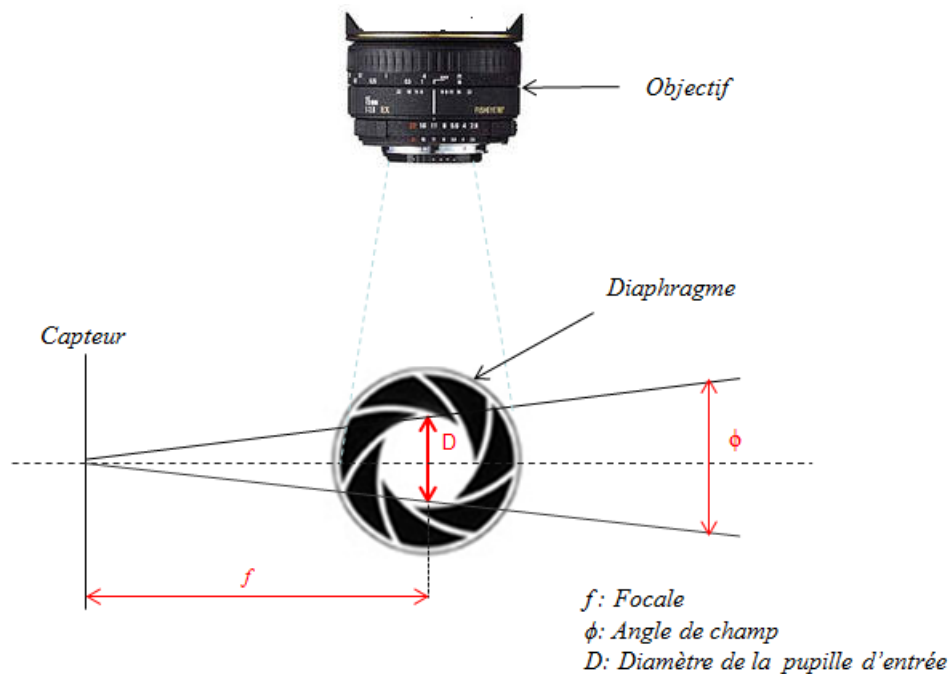


Figure 1.1 Caractéristiques d'un objectif

1.1.2 Les objectifs à focale variable

Le déplacement d'un groupe de lentilles permet la modification de la focale et donc l'agrandissement tout en maintenant la luminosité constante. Il permet aussi de changer le champ de vue. Généralement, la mise au point d'un objectif à focale variable est assurée par une seule bague et dépend de la distance par rapport à l'objet, de la focale et de l'ouverture.

Un zoom comprend (Smith, 1966) :

- Un bloc divergent, appelé variateur, qui permet de varier la distance focale et de modifier l'agrandissement de l'image.
- Un bloc convergent, appelé compensateur, qui compense la mise au point suite un changement de la distance focale.

Les zooms appartiennent à deux types de base :



- Zoom à compensation mécanique : un ou plusieurs groupes internes mobiles dont leurs déplacements non linéaires respectifs permettent la compensation de la mise au point pour toute valeur de la focale de zoom. Ces mouvements nécessitent un mécanisme à cames (Figure 1.2-a).
- Zoom à compensation optique : deux groupes de lentilles couplés se déplaçant en respectant d'autres groupes fixes. Ces déplacements permettent d'obtenir la mise au point pour les valeurs de la focale du zoom (Figure 1.2-b).

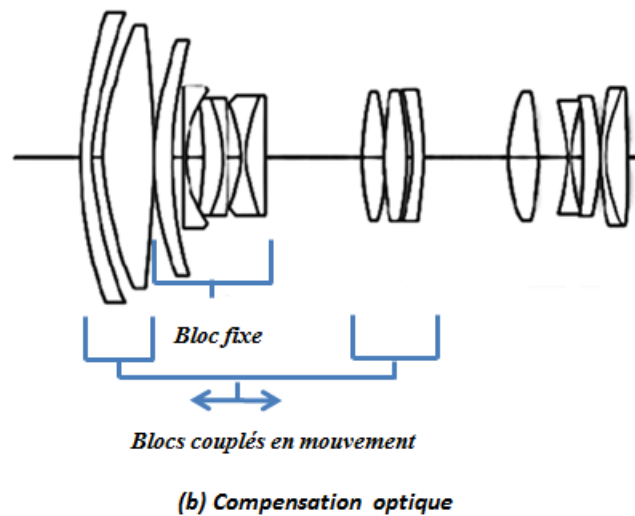
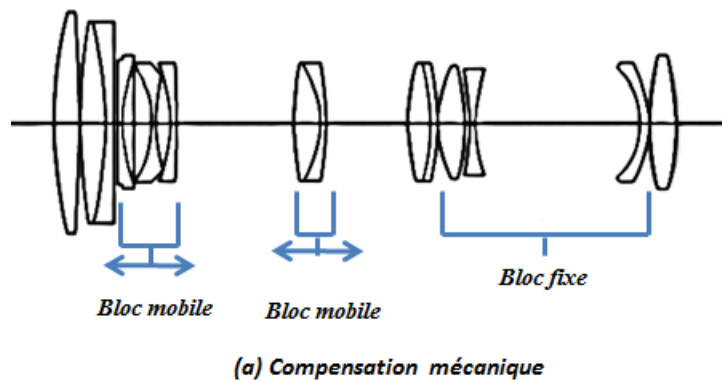


Figure 1.2. Objectifs à focale variable

(Inspirée de Smith (1966))

1.1.3 Les objectifs grands angles

Dans la vision panoramique, on peut distinguer deux grandes approches : la première est l'association de caméras perspectives ou affines et de miroirs de courbure connue (objectifs catadioptriques) (Baker and Nayar, 1998, Lhuillier, 2011, Schuster, 2011). La deuxième voie est la combinaison d'une surface réfractive (lentille) avec des imageurs conventionnels tels que les caméras fisheye et panomorphe (Kumler and Bauer, 2000, Aqi et al., 2011, Thibault, 2007).

1.1.3.1 Objectifs catadioptriques

L'objectif catadioptrique repose sur la combinaison d'un miroir et d'un objectif classique (Aliakbarpour et al., 2014, Xiang et al., 2013, Zhang et al., 2013). Le miroir est un système catoptrique qui permet de réfléchir les rayons lumineux provenant de toutes directions (Figure 1.3). Les rayons réfléchis sont ensuite convergés par une série de lentilles et forment une image omnidirectionnelle, une fois projetés sur le capteur.

L'inconvénient de ces objectifs est que la présence du miroir rend une partie importante du centre de l'image inexploitable.

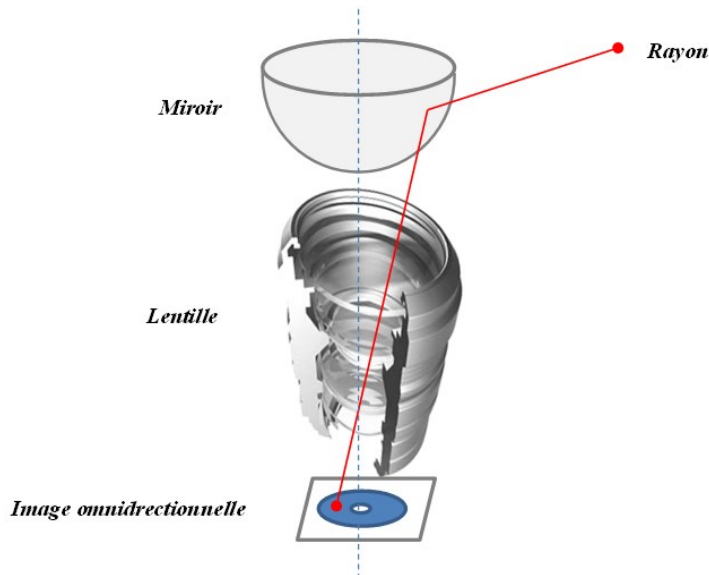


Figure 1.3 Objectif catadioptrique

1.1.3.2 Objectif Fisheye et objectif panomorphe

Les images omnidirectionnelles peuvent être acquises par un objectif fisheye (Figure 1.4-a) ou panomorphe (Figure 1.4-b) qui fournissent un hémisphère de vue du monde observé (Zhang et al., 2014, Thibaut, 2007). Ce sont des objectifs réfractifs ayant une distance focale courte et un large angle de champ jusqu'à 180 degrés. Ils permettent de résoudre le problème des caméras catadioptriques et d'exploiter la partie au centre de l'image (plus grande couverture optique par rapport aux caméras catadioptriques).

Néanmoins, ils présentent l'inconvénient de fortes distorsions dans l'image.

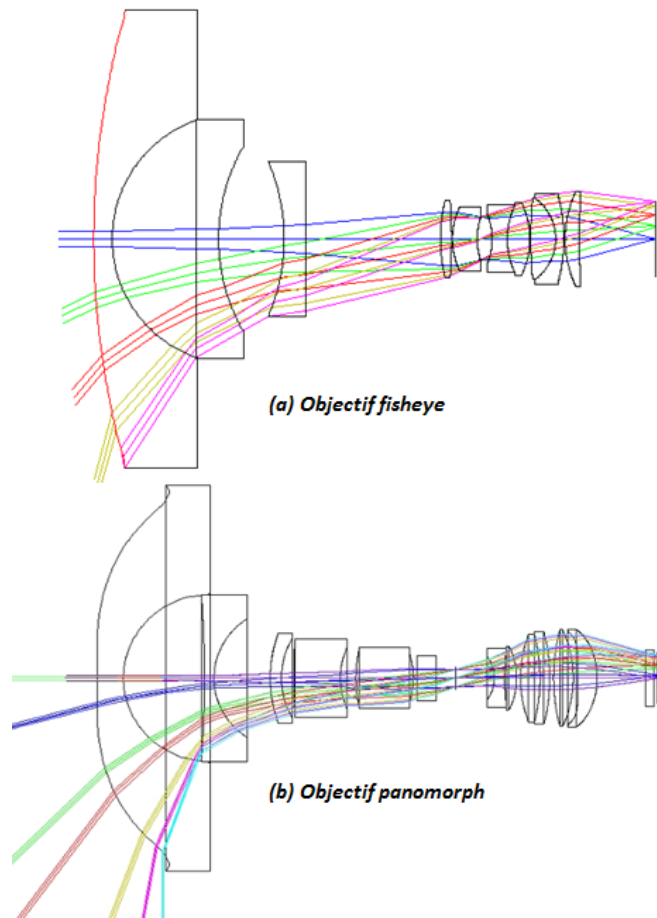


Figure 1.4 Objectifs réfractifs

1.2 La stéréovision

1.2.1 Définition

Le principe de la stéréovision (stéréoscopie) consiste à reconstruire tridimensionnellement une cible ou toute une scène en utilisant des images prises simultanément (prises de points de vue différents) ou séquentiellement. Plusieurs chercheurs ont étudiés des systèmes stéréoscopiques formés de deux (Dwyer and Southam, 2004, Wan and Zhou, 2008, Nishimoto and Yamaguchi, 2007) ou plusieurs caméras (Oskiper et al., 2007, Ahmad et al., 2006, Sundaresan and Chellappa, 2005).

Généralement, on peut distinguer deux approches : la stéréovision dynamique et la stéréovision statique. La stéréovision dynamique est une méthode de mesure tridimensionnelle d'une cible qui se déplace entre les prises de vue (Mouragnon et al., 2009, Waxman and Sinha, 1986). Dans ce cas une seule caméra pourrait être suffisante pour effectuer des mesures tridimensionnelles. Plusieurs images prises séquentiellement sont considérées comme des images acquises depuis différents points de vue.

La stéréovision statique impose que les objets à mesurer soient fixes par rapport au référentiel objet. Les techniques de reconstruction tridimensionnelle de cette approche utilisent des images acquises au moyen d'une seule caméra qui se déplace dans la scène (Gao and Ahuja, 2004, Davison, 2003) ou bien deux ou plusieurs caméras afin de prendre des images simultanément (Lauer et al., 2011, Denker and Umlauf, 2011).

Dans toutes les techniques utilisées, les cibles devraient être visibles dans au moins deux images.

Selon les champs de vue (FOV) des caméras utilisées, on peut distinguer 3 types de systèmes stéréoscopiques soit perspective, mixte et panoramique.

La stéréovision perspective utilise une ou plusieurs caméras classiques qui peuvent être modélisées par une projection rectiligne (modèle sténopé) afin d'obtenir une reconstruction tridimensionnelle (Nedevschi et al., 2004, Ding et al., 2009).

Le grand intérêt des caméras panoramiques est leur large champ de vue. La stéréovision mixte combine ces caméras avec des caméras classiques et exploite les vues perspectives et panoramiques pour faire des mesures tridimensionnelles (Eynard et al., 2010, Roberti et al., 2008).

La stéréovision panoramique vise à élargir le champ de vue commun des caméras utilisées pour recueillir un nombre plus important d'informations et réaliser des mesures dans l'ensemble des directions (Peleg et al., 2001, Moreau et al., 2012).

Les étapes importantes pour des nombreux algorithmes de stéréovision sont : le calibrage des caméras, la détermination des points en correspondance entre les images et la reconstruction par triangulation.

1.2.2 Calibrage de caméras

En vision numérique, le calibrage d'une caméra est une opération de modélisation. Il consiste à estimer les paramètres intrinsèques (focale, point principale, paramètres de la distorsion) et extrinsèques (position et orientation) de la caméra. En photogrammétrie, la détermination de la position et de l'orientation de la caméra ne fait pas partie du calibrage. Ces paramètres d'orientation externe sont déterminés lors du relèvement spatial. Dans cette thèse, c'est la définition utilisée en vision numérique qui sera considérée.

Dans la littérature, il y a plusieurs techniques de calibrage qui ont été présentées selon le type de la caméra. Un bref aperçu des méthodes de calibrage des caméras panoramiques et zoom sera donné ici.

1.2.2.1 Calibrage des caméras panoramiques

Pour les caméras catadioptriques, le modèle géométrique introduit une transformation supplémentaire (par rapport aux modèles classiques) qui décrit les caractéristiques du miroir. Il y a plusieurs modèles qui ont été développés suivant le type du miroir et de la caméra utilisée (Baker and Nayar, 1999, Barreto and Araujo, 2005, Gandhi and Trivedi, 2006, Mei and Rives, 2007, Zhang et al., 2009, Dunne et al., 2010). Cependant, Geyer and Daniilidis (2000) proposent un modèle unifié en modélisant la caméra par une première projection sur une sphère suivie d'une seconde projection perspective sur le plan image.

1.2.2.1.1 Autocalibrage

On désigne par l'auto-calibrage, toute méthode permettant d'étalonner automatiquement une caméra sans utiliser des objets spécifiques d'étalonnage. Dans la vision panoramique comme dans le cas des caméras classiques, la procédure consiste à estimer la matrice fondamentale à partir de la mise en correspondance de deux images d'une scène en même

temps que le raffinement des coordonnées du patron de points de calibrage (Micusik and Pajdla, 2004, El Rahman Shabayek et al., 2009, Li and Hartley, 2006).

1.2.2.1.2 Étalonnage en utilisant une mire de calibrage

Ces méthodes sont les plus utilisées et reposent essentiellement sur le calcul des paramètres intrinsèques et extrinsèques à l'aide d'un ensemble de points de calibrage. Ces points ne sont pas parfaitement connus, mais on possède un estimé de leur précision. Cela permet d'introduire un poids (degré de confiance) dans la compensation (Hill and Tiedeman, 2006).

Le principe est d'utiliser des points, dits points d'appui, marqués sur une grille ou un autre objet de calibrage et dont les coordonnées sont connues dans l'espace. En connaissant les coordonnées espace d'un ensemble de points, le problème revient alors à optimiser le modèle mathématique associé au capteur pour que leurs projections correspondent au mieux aux mêmes points extraits de l'image (Sturm and Maybank, 1999, Mei and Rives, 2007, Du and Zhu, 2011).

1.2.2.1.3 Calcul du modèle générique

Une méthode qui a été développée par Sturm et Ramalingam (Sturm and Ramalingam, 2003, Sturm and Ramalingam, 2004, Ramalingam et al., 2005) et qui englobe la plupart des modèles de projection utilisés dans la vision numérique ou en photogrammétrie.

Chaque pixel de l'image capte la lumière qui se déplace le long d'un rayon en 3D. La caméra est décrite par les coordonnées de ce rayon et la projection entre les rayons et les pixels. La méthode consiste à acquérir N images ($N \geq 2$) d'un objet de référence (mire de calibrage) depuis des points de vue inconnus. En considérant le même pixel dans les deux N images, il y a N points dans la scène qui sont vus par ce pixel. Le calibrage consiste à calculer les coordonnées de ces N points et par la suite déduire le rayon vu par le pixel.

1.2.2.1.4 Calibrage basé sur des points de fuite

L'information d'un objet dans l'espace tel que les lignes permet d'un côté de réduire le nombre de points d'appui utilisés pour calibrer et d'un autre côté d'offrir des conditions géométriques supplémentaires utilisées pour obtenir un calibrage précis (He and Li, 2008).

Un point de fuite est la projection de l'intersection à l'infini de deux ou plusieurs droites parallèles. Théoriquement, les projections de ces lignes sur l'image sont des lignes. Sous l'effet de la distorsion, elles présentent des courbures dans l'image qui convergent en deux points de fuite. Cette propriété a permis d'estimer le point principal, la focale et les paramètres de distorsion en utilisant une grille formée d'un ensemble de droites horizontales et verticales parallèles (Grammatikopoulos et al., 2007, Hughes et al., 2010a).

1.2.2.2 Calibrage des caméras zoom

Les objectifs à focale variable ou zoom présentent la particularité d'autoriser le déplacement d'un bloc de lentilles convergentes de façon à configurer le changement de la focale et provoquer un grossissement apparent de l'image. Le calibrage d'un zoom estime les paramètres intrinsèques et extrinsèques de la caméra ainsi que les paramètres de la distorsion à des multiples distances focales.

Plusieurs travaux ont traité le calibrage relatif à ce genre d'objectif et la transformation subie par l'image en faisant varier la focale. L'objectif était d'établir une relation entre les paramètres de contrôle de la lentille (le changement de focale et la mise au point) et les paramètres de formation de l'image (paramètres intrinsèques et extrinsèques). Remondino and Fraser (2006) considèrent qu'en cas d'une caméra zoom, deux transformations sont présentes. Une première transformation qui modélise la relation entre le processus de formation de l'image (projection perspective) et la configuration géométrique des lentilles. Et une deuxième transformation qui relie la configuration des lentilles aux paramètres de contrôle des lentilles. Pour une position donnée de la focale (ou dans le cas d'une caméra à focale fixe), la caméra est modélisée par une seule transformation (la projection perspective ou rectiligne).

1.2.2.2.1 Relation entre Focale et focus

Le focus, ou la mise au point, permet d'obtenir une image nette en déplaçant les lentilles. Son réglage dépend de la distance par rapport à l'objet, de la distance focale et de l'ouverture. La distance focale est la distance entre le centre optique de l'objectif et la surface du capteur lorsque la mise au point est à l'infini. Laval et al. (1992) ont prouvé expérimentalement que pour un zoom parfaitement réglé, la défocalisation ne modifie pas la netteté de l'image. Ils ont montré que, pour un objectif où on considère que le changement de focale est réalisé par une translation pure, les coordonnées du point principal restent stables. En faisant des approximations, Enciso et al. (1993) ont considéré que le changement de mise au point ou de focale revient à effectuer une transformation affine dans le plan image. Cette transformation permet de calculer les nouveaux paramètres intrinsèques. Puisque la taille de pixel ne change pas, les paramètres de cette transformation sont au nombre de trois et il suffit de suivre deux points pour les calculer. Cependant, cette méthode ne permet pas de donner des mesures précises.

1.2.2.2.2 Modèle optique à lentille épaisse

Fayman et al. (2001) ont montré qu'un zoom ne pouvait pas se modéliser par un modèle sténopé. Le modèle à lentille épaisse permet d'expliquer mieux son comportement. Il permet de représenter la lentille par deux plans principaux (Figure 1.5). Il est équivalent au modèle sténopé tout en considérant le changement de focale comme étant un mouvement axial virtuel.

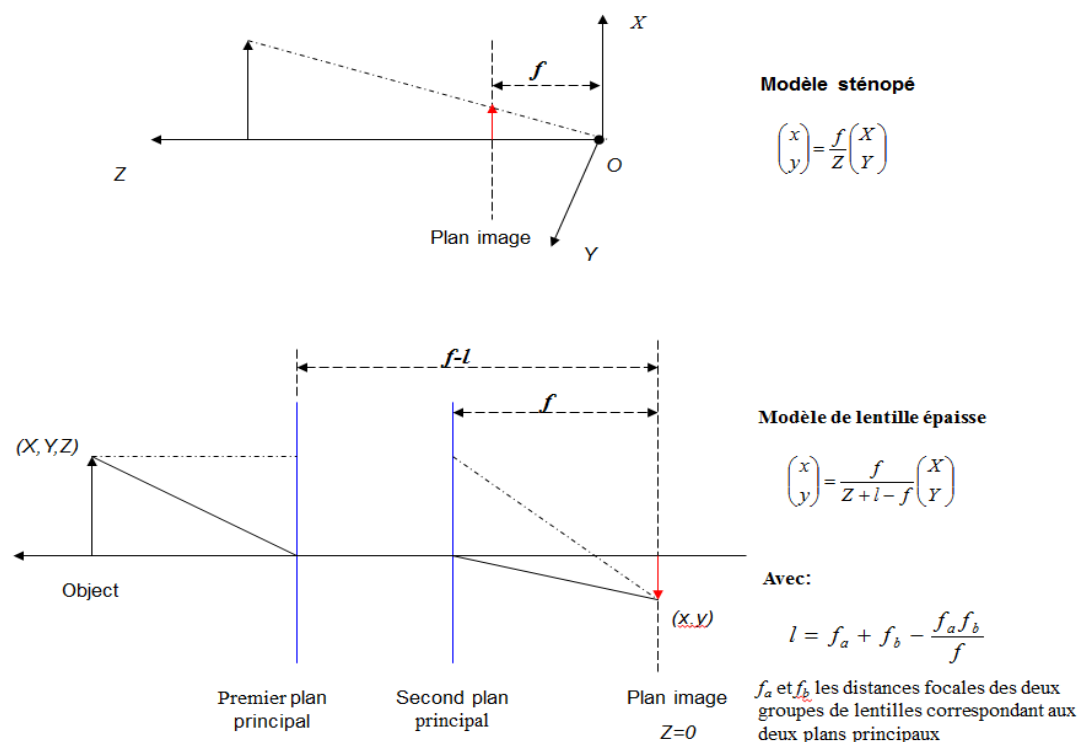


Figure 1.5 Comparaison entre le modèle sténopé et le modèle de lentille épaisse
(inspirée de Fayman et al. (2001))

1.2.2.2.3 Table de consultation et interpolation polynomiale

Dans le cas d'un zoom, plusieurs travaux ont estimé les paramètres lors de calibrage pour différentes configurations focale/mise au point afin d'établir une table de consultation (Atienza and Zelinsky, 2001, Park and Hong, 2001). Le point principal ainsi que le grossissement de l'image peuvent être estimés pour multiples zooms. Ensuite, une table de consultation est calculée avec interpolation. La variation des paramètres de distorsion peut être estimée en se basant sur la fonction de grossissement (Collins and Tsin, 1999).

Plusieurs travaux ont proposé des fonctions polynomiales pour estimer la distorsion lorsqu'on varie la focale. Fraser and Al-Ajlouni (2006), par exemple, ont préconisé une variation linéaire du point principal en fonction de la focale. Låbe and Förstner (2004) ont montré un comportement non linéaire et ont prouvé expérimentalement que l'interpolation



linéaire conduit à des erreurs plus au moins importantes. Alvarez et al. (2012) ont proposé une approximation polynomiale pour estimer la variation de la distorsion radiale.

1.2.2.2.4 Autocalibrage

L'auto-calibrage pour les caméras zoom est réalisé en calculant les projections projectives 2D ou homographies pour différentes focales, ce qui permet d'estimer la matrice de paramètres intrinsèques (Seo and Hong, 1998, Agapito et al., 2001, Cao et al., 2006). L'interpolation permet, par la suite, d'estimer les paramètres de calibrage pour un niveau de zoom arbitraire.

1.2.2.2.5 Neurocalibrage

Cette méthode consiste à déterminer la relation entre un point 3D de la scène et sa projection sur l'image sans connaître la géométrie de la caméra et sans recours à une procédure expérimentale complexe. Un réseau de neurones artificiels permet de résoudre le système non-linéaire qui lie les coordonnées espace aux coordonnées image et par la suite d'obtenir l'information 3D sans connaître les paramètres de la caméra.

Dans la littérature, on retrouve plusieurs travaux qui ont utilisé cette méthode pour calibrer des caméras classiques (Jun and Kim, 1999, Ahmed and Farag, 2000, Woo and Park, 2009). Ils ont comparé les résultats obtenus par réseaux de neurones artificiels et ceux obtenus avec des méthodes de calibrage explicites tel que la méthode de Tsai (Tsai, 1987). Ils ont montré qu'un réseau de neurones permet une modélisation plus précise. L'architecture d'un réseau de neurones consiste en la représentation des neurones en couches (Figure 1.6). Pour calibrer une caméra à focale fixe, la première couche est la couche d'intrants constituée de 4 neurones (3 coordonnées espace et un neurone fixé à 1). La dernière couche est la couche de sortie formée de 3 neurones (2 coordonnées pixel et un neurone fixé à 1). Les couches intermédiaires sont appelées la couche cachée qui se comporte comme une fonction qui transforme les paramètres d'entrée en paramètres de sortie. L'apprentissage ou la phase de mise à jour de poids vise à modifier le comportement de réseau afin d'obtenir les paramètres de sortie désirés. À chaque neurone correspond un poids. La matrice des poids de la couche cachée correspond aux paramètres intrinsèques de

18

la caméra alors que les poids de la couche de sortie correspondent aux paramètres extrinsèques.

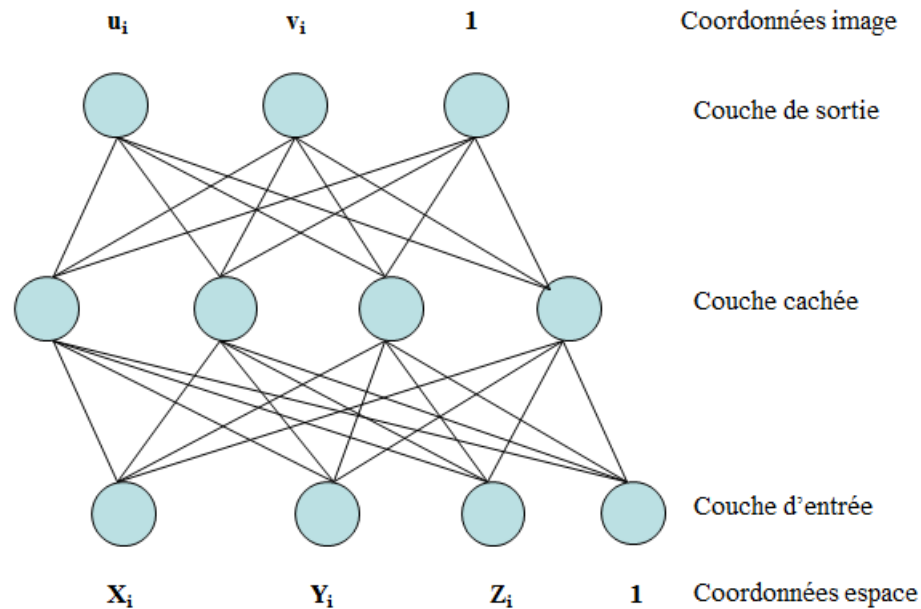


Figure 1.6 Modèle du calibrage par réseau de neurones

Pour calibrer un zoom, chaque paramètre supposé constant en variant la focale est donné par le poids correspondant. Tandis que chaque paramètre qui varie en fonction de la focale est représenté par un réseau de neurones. Ainsi l'estimation de ce dernier revient à déterminer le nombre de couches et les poids de neurones dans son réseau (Ahmed and Farag, 2002).

1.2.2.2.6 Calibrage de caméras de type Pan-Tilt-Zoom camera

Le Pan est la rotation de la caméra autour de l'axe optique. Le tilt désigne l'inclinaison ou le déplacement de la caméra dans le plan vertical (vers le haut ou vers le bas). Le zoom est le mouvement de la lentille le long de l'axe optique (Figure 1.7). La possibilité d'orienter la caméra dans différentes directions permet de couvrir toute la scène. Le zoom quant à lui

permet de surveiller une zone de détails ajustable. Dans un intervalle de temps, la caméra peut être considérée comme une caméra panoramique à haute résolution.

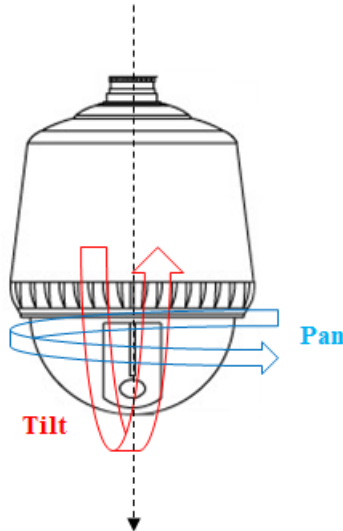


Figure 1.7 Pan-Tilt-Zoom camera

Dans la littérature, il y a généralement deux approches de calibrage de la Pan-Tilt-Zoom caméra. La première approche consiste à optimiser les paramètres intrinsèques et extrinsèques de la caméra considérés comme variables dans multiples images prises en variant le pan, le tilt et le zoom (Liao and Cho, 2008, Szwoch et al., 2013). L'utilisation de mires spécifiques et précises constitue la première contrainte de cette méthode. Une deuxième contrainte consiste à détecter les points cibles dans les images avec précision. La deuxième approche utilise le modèle d'autocalibrage qui, en variant le pan et le tilt, permet de construire un panorama de la scène pour chaque zoom. La variation des paramètres intrinsèques est, donc, déterminée en optimisant les homographies pour chaque couple d'images et pour chaque position de zoom (Sinha and Pollefeys, 2006, Leite et al., 2011).

1.2.3 Calcul de la géométrie épipolaire

La géométrie épipolaire décrit la contrainte liant les projections d'un seul point de l'espace dans deux images. Elle définit le déplacement d'un point entre deux prises de vues différentes.

1.2.3.1 Géométrie épipolaire perspective

Dans le cas d'un système de stéréovision perspective, les données nécessaires pour estimer la géométrie épipolaire (généralement on parle de droites épipolaires) sont les projections de certains points vus par les deux caméras. La première méthode permettant d'estimer la géométrie épipolaire est la méthode des 8 points (méthode linéaire) qui impose que les coordonnées de 8 points soient connues dans deux images (Hartley and Zisserman, 2003).

La deuxième technique est la méthode d'optimisation linéaire utilisée lorsqu'on a plus que 8 paires de points en correspondance (Zhang, 1998). La troisième technique est basée sur une méthode itérative ou non-linéaire (Zhang, 1998) qui consiste à minimiser la distance entre les points et les droites épipolaires correspondantes et qui nécessite une initialisation des paramètres (généralement en utilisant les données calculées par la méthode linéaire).

1.2.3.2 Géométrie épipolaire panoramique

La géométrie épipolaire est difficile à exprimer analytiquement dans le cas d'un système de stéréovision mixte ou panoramique. Les modèles de projection des deux caméras sont différents et généralement on ne parle pas de droites épipolaires mais plutôt de coniques épipolaires (intersection d'un plan avec le miroir).

Le plus souvent, les coniques épipolaires dépendent de types de caméras utilisées et de leurs configurations. Gluckman and Nayar (1998) ont proposé un système formé de deux caméras catadioptriques placées l'une au-dessus de l'autre et ont montré que les courbes épipolaires dans ce cas sont des droites radiales. Une formulation générale pour les caméras catadioptriques centrales a été établie par (Svoboda and Pajdla, 2002, Fujiki et al., 2007).

1.2.4 Triangulation

Une fois que les coordonnées d'une cible sont estimées dans une ou plusieurs images sous des angles de vue différents, la triangulation consiste à calculer l'intersection des rayons issus de chaque caméra passants par ces points. Chaque rayon est construit à partir de la matrice de projection correspondante. Cependant, à cause des erreurs de mesure dues au calibrage, les rayons ne se croisent pas. La première technique proposée était de calculer le point au milieu du plus petit segment reliant les deux rayons (Beardsley et al., 1997, Beardsley et al., 1994). Des techniques plus stables ont été proposées ultérieurement. En effet, en minimisant la somme des carrés des distances entre les points projetés en utilisant le modèle théorique et les points mesurés ou observés dans l'image, on peut obtenir une solution plus précise. Une revue des méthodes classiques de triangulation avec optimisation linéaire et non linéaire a été présentée par Hartley and Sturm (1997).

1.3 Projection perspective

Le modèle simplifié généralement utilisé pour représenter une caméra est le modèle « pinhole » (ou rectiligne ou perspective ou sténopé) où tous les rayons passent par un seul point (centre de projection ou centre optique). Le modèle dépend de deux types de paramètres: intrinsèques et extrinsèques.

Les paramètres intrinsèques décrivent les propriétés internes de la caméra comme la distance principale (la distance entre le centre du repère caméra et le plan image) et la taille d'un pixel. Les paramètres extrinsèques décrivent la position et l'orientation de la caméra par rapport au repère espace. Ils se décomposent en une matrice de rotation et un vecteur de translation.

1.3.1 Modèle mathématique de la projection perspective

Le modèle sténopé est une transformation projective de $\mathbb{P}^3 \rightarrow \mathbb{P}^2$, donc tout point $\mathbf{P}$ de l'espace de coordonnées (X,Y,Z) se projette en un point $\mathbf{p}$ de l'image de coordonnées (u,v) . Cette projection est caractérisée par trois transformations.

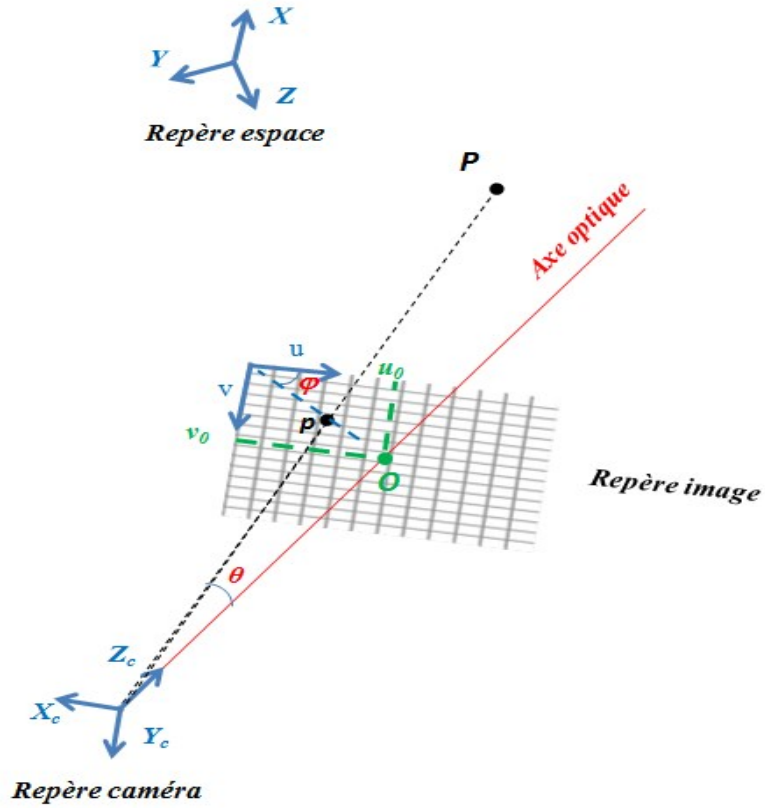


Figure 1.8 Projection perspective

1.3.1.1 Transformation coordonnées/direction

Cette transformation permet d'exprimer la direction d'un rayon (θ, φ) provenant du point P dans un repère lié à la caméra. θ est l'angle que fait ce rayon avec l'axe optique et φ l'angle entre ce rayon et l'axe u de l'image. La transformation est composée d'une rotation (R) et d'une translation (T).

$$(\sin \theta \cos \varphi \quad \sin \theta \sin \varphi \quad \cos \theta) = (RP + T) / \|RP + T\| \quad (1.1)$$

où:

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \text{ et } T = \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix}.$$

1.3.1.2 Transformation direction/coordonnées normalisées

Les coordonnées normalisées (x,y) dans l'image peuvent être déterminées à partir de la direction du rayon (θ, φ) de la façon suivante :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \tan \theta \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}. \quad (1.2)$$

1.3.1.3 Transformation coordonnées normalisées/coordonnées image

Le repère caméra est un repère intermédiaire dont l'origine est le centre de projection. L'axe Z_C est l'axe optique. Les axes X_C et Y_C sont choisis comme étant parallèles au plan image et perpendiculaire entre eux. Alors que le repère image est un repère dont l'unité de mesure est le pixel et son origine est confondu avec le point principal, soit intersection de l'axe optique avec le plan image, et ses axes u et v sont parallèles aux axes X_C et Y_C du repère caméra.

Les coordonnées image (u_p, v_p) sont obtenues en appliquant une transformation affine aux coordonnées normalisées:

$$\begin{pmatrix} u_p \\ v_p \end{pmatrix} = \begin{pmatrix} f_u & 0 \\ 0 & f_v \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}. \quad (1.3)$$

où:

- (u_0, v_0) sont les coordonnées du point principale noté O .
- (f_u, f_v) : la distance principale (focale) mesurée respectivement suivant l'axe horizontal (u) et vertical (v) de l'image.

L'expression finale de la projection perspective qui lie la direction d'un rayon (θ, φ) aux coordonnées image est la suivante :

$$\begin{pmatrix} u_p \\ v_p \end{pmatrix} = \begin{pmatrix} f_u & 0 \\ 0 & f_v \end{pmatrix} \tan \theta \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \quad (1.4)$$

En utilisant les coordonnées homogènes, on peut établir la relation générale suivante (Arfaoui, 2009) :

$$\begin{pmatrix} ku_p \\ kv_p \\ k \end{pmatrix} = \begin{pmatrix} f_u & 0 & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \quad (1.5)$$

ou encore :

$$\begin{pmatrix} ku_p \\ kv_p \\ k \end{pmatrix} = M \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (1.6)$$

avec $M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix}$, représente la matrice de projection perspective.

1.3.2 Distorsions

La projection perspective est le modèle linéaire de formation d'images qui ne prend pas en considération les distorsions géométriques induites par l'objectif utilisé.

Lorsqu'on tient compte des paramètres de la distorsion dans le calibrage d'une caméra, le modèle devient non-linéaire.

La modélisation de la distorsion permet alors de décrire l'écart entre les coordonnées observées ou distordues (u_d, v_d) d'un point dans l'image et sa position idéale (u_p, v_p) . Plusieurs modèles mathématiques ont été proposés dans la littérature.

1.3.2.1 Modèle de distorsions radiales et tangentielles

Dans le modèle classique, il y a deux types de distorsions: les distorsions radiales et les distorsions tangentielles. La distorsion radiale est due à l'optique utilisée et au fait que les lentilles présentent des défauts de courbure. Généralement c'est la distorsion la plus importante qui doit être nécessairement corrigée pour obtenir un calibrage précis.

Les corrections des coordonnées dans le plan image sont de la forme :

$$\begin{pmatrix} \Delta u_{rd} \\ \Delta v_{rd} \end{pmatrix} = \begin{pmatrix} u_p \\ v_p \end{pmatrix} (k_1 r^2 + k_2 r^4 + \dots) \quad (1.7)$$

avec (k_1, k_2) : paramètres de la distorsion radiale et $r = \sqrt{(u_p - u_0)^2 + (v_p - v_0)^2}$

La distorsion tangentielle est exprimée par deux composantes : la distorsion de décentrage et la distorsion affine.

La distorsion de décentrage est due au décentrage des lentilles par rapport à l'axe optique.

Les corrections sont sous la forme :

$$\begin{pmatrix} \Delta u_{dc} \\ \Delta v_{dc} \end{pmatrix} = \begin{pmatrix} d_1(r^2 + 2u_p^2) + 2d_2u_pv_p \\ d_2(r^2 + 2v_p^2) + 2d_1u_pv_p \end{pmatrix} \quad (1.8)$$

avec (d_1, d_2) : paramètres de la distorsion de décentrage.

La distorsion affine est la déformation due à la non-orthogonalité des deux axes de l'image et les dimensions du pixel élémentaire, généralement non-carré.

$$\begin{pmatrix} \Delta u_{af} \\ \Delta v_{af} \end{pmatrix} = \begin{pmatrix} -e_1 u_p + e_2 v_p \\ e_1 v_p \end{pmatrix} \quad (1.9)$$

avec (e_1, e_2) : paramètres de la déformation affine.

1.3.2.2 Modèle polynomial

Le modèle polynomial a été proposé afin de modéliser la distorsion pour une caméra fisheye (Basu and Licardie, 1995). En effet, le modèle classique n'est pas capable de modéliser la distorsion pour des caméras grand-angle telles que la caméra fisheye.

Il s'agit alors d'établir la relation suivante entre les coordonnées polaires idéales d'un point (r_p, φ_p) et les coordonnées polaires distordues observées (r_d, φ_d) :

$$\begin{cases} r_d = \sum_{i=1}^n f_i r_p^i \\ \varphi_d = \sum_{i=1}^m g_i \varphi_p^i \end{cases} \quad (1.10)$$

avec (f_i, g_i) les paramètres du modèle polynomial.

1.3.2.3 Modèle logarithmique

Ce modèle vise aussi à établir une relation entre les coordonnées polaires idéales et distordues (Basu and Licardie, 1995). L'expression du modèle est la suivante :

$$\begin{cases} r_d = s \log(1 + \lambda r_p) \\ \varphi_d = \varphi_p \end{cases} \quad (1.11)$$

avec (s, λ) les paramètres du modèle logarithmique.

1.3.2.4 Modèle de division

C'est un modèle approximatif qui propose une correction radiale sous la forme suivante (Fitzgibbon, 2001):

$$r_d = \frac{r_p}{1 + \beta_1 r_p^2 + \beta_2 r_p^4 + \dots} \quad (1.12)$$

avec (β_1, β_2) les paramètres du modèle de division.

1.3.2.5 Modèle rationnel

Ce modèle offre une fonction rationnelle quadratique qui permet d'établir une relation entre les coordonnées idéales et distordues (Ricolfe-Viala and Sanchez-Salmeron, 2010). La fonction est donnée par :

$$f_r = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \end{pmatrix} \begin{pmatrix} u_d^2 \\ u_d v_d \\ v_d^2 \\ u_d \\ v_d \\ 1 \end{pmatrix} \quad (1.13)$$

En utilisant cette fonction, on peut calculer les coordonnées idéales :

$$\begin{cases} u_p = \frac{f_r(1)}{f_r(3)} \\ v_p = \frac{f_r(2)}{f_r(3)} \end{cases} \quad (1.14)$$

1.3.3 Résolution du modèle

Plusieurs méthodes existent dans la littérature permettant le calcul des paramètres de la caméra. Elles sont toutes basées sur la minimisation d'un critère linéaire ou non.

Selon l'équation 1.5, on peut déterminer le modèle linéaire de la projection perspective en fonction des paramètres intrinsèques et extrinsèques de la caméra, soit :

$$\begin{aligned} u_p &= u_0 + \frac{r_{11}X + r_{12}Y + r_{13}Z + t_x}{r_{31}X + r_{32}Y + r_{33}Z + t_z} \\ v_p &= v_0 + \frac{r_{21}X + r_{22}Y + r_{23}Z + t_y}{r_{31}X + r_{32}Y + r_{33}Z + t_z} \end{aligned} \quad (1.15)$$

La technique linéaire résout l'équation 15 en se basant sur la méthode des moindres carrés (Arfaoui, 2009). Son avantage est la simplicité du modèle qui permet un calibrage simple et rapide. L'inconvénient de cette technique est qu'on ne tient pas compte des paramètres de la distorsion dans le calibrage.

Le modèle linéaire est couramment accompagné d'un couplage des facteurs de distorsions avec les paramètres intrinsèques de la caméra. En présence de distorsions, on est face à un modèle non-linéaire.

En tenant compte des composantes de distorsion (Δu , Δv) et des erreurs de mesure dans l'image (e_u , e_v) respectivement selon les coordonnées u et v , le modèle général non linéaire est de la forme :

$$\begin{aligned} u_p + e_u &= u_0 + \frac{r_{11}X + r_{12}Y + r_{13}Z + t_x}{r_{31}X + r_{32}Y + r_{33}Z + t_z} + \Delta u = Q_1(\vartheta) \\ v_p + e_v &= v_0 + \frac{r_{21}X + r_{22}Y + r_{23}Z + t_y}{r_{31}X + r_{32}Y + r_{33}Z + t_z} + \Delta v = Q_2(\vartheta) \end{aligned} \quad (1.16)$$

où :

- ϑ : vecteur de paramètres.

- Q_1 et Q_2 deux fonctions non linéaire de ϑ .

Le problème de calibrage revient alors à déterminer la valeur de ϑ qui minimise la fonction suivante pour n mesures:

$$E = \sum_{i=1}^n (e_{ui} + e_{vi})^2. \quad (1.17)$$

Le modèle non linéaire donne des résultats précis. Pour assurer une convergence rapide, on a besoin d'une bonne estimation des paramètres initiaux qui sont généralement calculés par le modèle linéaire.

1.3.4 Mire de calibrage

Généralement, un système de calibrage de caméras comprend des composants matériels (un objet ou un mire de calibrage et un système d'acquisition d'images) et des composants logiciels (un algorithme pour extraire et identifier les coordonnées images de points de contrôle et un algorithme pour calculer les paramètres de la caméra).

Les méthodes de calibrage dites fortes reposent essentiellement sur le calcul des paramètres intrinsèques et extrinsèques à l'aide d'un objet parfaitement connu (Arfaoui, 2009). Pour les caméras classiques, cet objet est formé d'un ensemble de cibles qui peuvent être des points (nœud d'intersection de droites horizontales et verticales, coins des carreaux du damier, centre de cercles ou d'ellipses) (Schiller et al., 2008, Datta et al., 2009) ou d'autres formes géométriques telles que des droites (Barreto and Araujo, 2005), ou des sphères (Ying and Hu, 2004, Zhang et al., 2007).

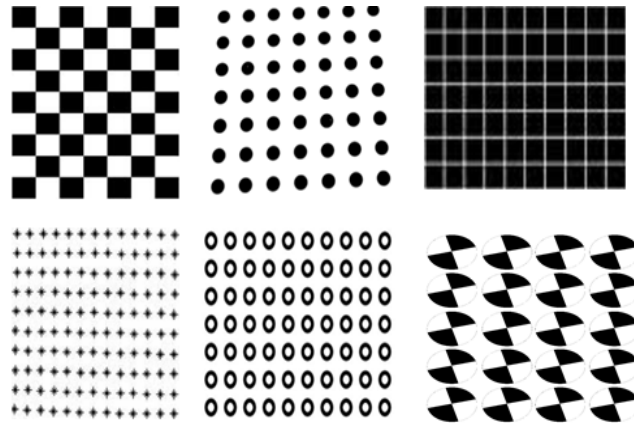


Figure 1.9 Motifs classiques pour calibrer une caméra

Les coordonnées de ces cibles doivent être connues précisément (généralement au dixième de millimètre près) dans un repère associé à la mire.

Dans la littérature, les différents motifs utilisés visent à ce que la localisation des cibles dans le plan image soit précise. Une revue de littérature sur les méthodes les plus utilisées pour détecter les points et les contours a été présentée dans Arfaoui (2009). Ces méthodes utilisent les propriétés des objets de calibration telles que la couleur et la géométrie pour améliorer la précision.

En ce qui concerne les caméras aériennes, la contrainte est de prendre des images d'un réseau de cibles à l'infini dont les coordonnées sont connues précisément en 3 dimensions dans un référentiel lié à l'objet. La distance assez grande par rapport à la caméra et la distorsion contribuent significativement à l'imprécision de localisation de ces cibles dans l'image. L'augmentation de leurs tailles permet de résoudre ce problème. Par contre, l'effort de la construction, l'installation et la mesure sur terrain devient intense.



Figure 1.10 points d'appui pour calibrer des caméras aériennes

Parmi les contraintes du calibrage d'une caméra équipée d'un zoom est la construction d'un prototype permettant sans le déplacer de la calibrer à différentes positions de zoom. En effet lorsque l'on effectue un zoom avant, on augmente la taille des objets à l'image. Par conséquent, le nombre de cibles vues par la caméra et la précision de la correspondance entre la scène et l'image diminuent.

La solution proposée dans la littérature consiste à concevoir une mire de calibrage avec deux motifs différents (carrés et cercles par exemple). Un motif sert à calibrer les niveaux élevés de zoom et l'autre à calibrer les niveaux bas (Pennisi et al., 2013).

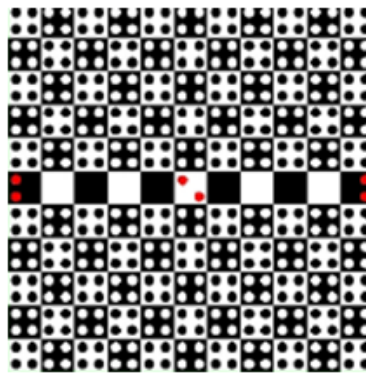


Figure 1.11 Motif spécial pour calibrer une caméra munie d'un zoom

Afin de calibrer les caméras panoramiques, des chercheurs ont proposé d'utiliser simplement les mires de calibrages classiques et d'acquérir plusieurs images afin de couvrir tout le champ de vue (Mei and Rives, 2007, Córdova-Esparza et al., 2014). D'autres chercheurs ont utilisé une mire d'étalonnage particulière, adaptée à la géométrie du capteur (Cauchois et al., 2000). Des nombreuses méthodes proposées sont basées sur la projection

de lignes dans l'image. Le principal avantage est qu'un motif spécial n'est pas nécessaire (Puig et al., 2012).

Problématique et objectifs

La vision panoramique présente des grands avantages particulièrement dans la vidéosurveillance, les systèmes robotisés et la navigation. L'utilisation d'un système de stéréovision mixte ou panoramique ne cesse de croître dans divers applications et domaines. Cependant la précision souhaitée de cette association de caméras est en deçà des attentes. Cela est dû à un calibrage qui n'est pas encore au point. La conception d'un nouveau banc de calibrage permettant de calibrer avec précision des caméras de tous types (classiques, munies d'un zoom, aériennes et panoramiques) apportera une grande contribution à toutes les applications de la vision numérique et de la photogrammétrie.

Le calibrage d'une caméra à focale variable présente deux problématiques. Tout d'abord, cette caméra a une gamme importante de distances focales. Le volume de données nécessaires pour calibrer est très important. En effet, pour chaque configuration focale/mise au foyer, une dizaine d'images est nécessaire pour déterminer les paramètres. Le calibrage de toute la gamme requiert plusieurs images et différentes configurations de l'objectif. Par contre quelques images (et parfois une seule image) permettent de calibrer une caméra à focale fixe.

Un deuxième défi consiste à prendre des images sur une gamme importante de distances focales sans changer l'objet de calibrage. Le nombre de cibles dans un objectif de calibrage diminue lorsque le champ de vision de la caméra diminue (surtout pour les longues focales). Pour les courtes focales, l'objet peut être trop petit pour couvrir toute l'image.

Le décentrage des lentilles par rapport à l'axe optique provoque le déplacement du centre optique en fonction de focale. Le point principal et la focale sont les paramètres intrinsèques à estimer pour chaque configuration de zoom. Les paramètres extrinsèques, quant à eux, identifient l'orientation et la position de la caméra dans un repère spatial fixe.

La stéréovision mixte permet de combiner le champ de vue plus large des caméras grand-angles et la résolution plus importante des caméras perspectives classiques. Alors que la stéréovision panoramique exploite un nombre plus important d'informations par rapport à la stéréovision classique et mixte en utilisant deux caméras grand-angles. La géométrie

épipolaire de telles combinaisons permet d'accéder à des informations qui sont utilisables aussi bien pour la mesure tridimensionnelle d'une scène quelconque que pour la commande d'un système (un robot par exemple). Les modèles géométriques des caméras panoramiques rendent plus complexe la détermination de cette géométrie. L'étude d'une nouvelle méthode plus appropriée et plus rapide permettra de mettre en correspondance plus précisément les pixels des images et d'établir adéquatement la relation entre les modèles géométriques de leurs formations.

La correction de la distorsion dans les images panoramiques rend applicables les algorithmes classiques qui sont très rapides. Par conséquent, un modèle unifié basé sur la projection rectiligne (modèle « pinhole ») qui permet de modéliser de la même façon les caméras panoramiques et perspectives offre la possibilité d'exploiter toutes les techniques de la stéréovision classique.

Pour répondre aux problématiques posées, la présente étude compte deux objectifs généraux. Le premier objectif est de proposer un nouveau banc de calibrage compact et précis permettant de calibrer à la fois des caméras classiques (aériennes, zoom) et panoramiques (fisheye, panomorphe). Le deuxième objectif consiste à proposer un nouveau modèle mathématique qui permet simplifier le calibrage d'une lentille aussi complexe que la lentille panomorphe et la modélisation d'un système stéréoscopique, formé d'un couple mixte ou panoramique. Un tel modèle permettra la reconstruction tridimensionnelle en utilisant les techniques de la stéréovision classique

Les objectifs spécifiques abordés dans les prochains chapitres sont les suivants :

- Définir le modèle mathématique adéquat et optimiser le calibrage pour des caméras classiques en utilisant un banc de calibrage à l'aide des éléments d'optique diffractive (DOEs).
- Proposer une méthode pour calibrer une caméra fisheye en utilisant ce banc de calibrage.
- Proposer une méthode pour calibrer une caméra panomorphe en utilisant ce banc de calibrage.

- Étudier la précision et l'exactitude d'un nouveau modèle générique permettant de modéliser géométriquement, à l'aide d'une projection rectiligne et de splines cubiques composées, un système de stéréovision mixte ou panoramique.

Chapitre 2. Calibrage des caméras classiques à l'aide des éléments de l'optique diffractive

L'objectif de ce chapitre est d'exposer certaines notions théoriques utiles pour introduire la méthode des éléments d'optique diffractive (DOEs), et de présenter le modèle de projection d'un point de la grille (formée par les DOEs) sur l'image. L'opération de calibrage de certaines caméras classiques en utilisant cette méthode et la comparaison avec une autre technique devraient confirmer sa précision et son exactitude.

2.1 Lasers et réseaux de diffraction

2.1.1 Géométrie d'un faisceau laser

Le faisceau gaussien permet de décrire les rayonnements cohérents issus d'un laser. En pratique, un faisceau laser s'écarte de l'axe Z, axe de propagation de la lumière (Figure 2.1). $w(Z)$ est la taille de faisceau qui vérifie l'équation suivante (Alda, 2003):

$$w(z) = w_0 \left(1 + \frac{Z^2}{Z_R^2}\right)^{1/2} \quad (2.1)$$

où :

- w_0 , appelé « waist » ou cercle de gorge, est le rayon le plus faible ou le rayon minimal du faisceau laser.
- Z_R est la longueur de Rayleigh qui vérifie la relation suivante :

$$Z_R = k \frac{w_0^2}{2} = \frac{\pi w_0^2}{\lambda}, \quad (2.2)$$

où k est le vecteur d'onde du faisceau et λ la longueur d'onde)

Le graphique de $w(z)$ est divisé en deux parties. Une partie hyperbolique (zone de Rayleigh) où les faisceaux sont parallèles. Une deuxième partie linéaire qui donne la divergence θ_{div} du faisceau (Weichel, 1990):

$$\theta_{div} = \frac{\lambda \pi}{w_0} \quad (2.3)$$

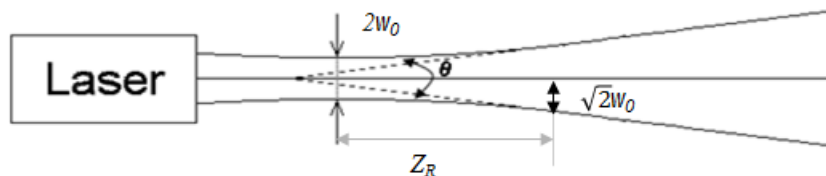


Figure 2.1 Géométrie d'un faisceau laser

(inspirée de Weichel (1990))

2.1.2 Réseaux de diffraction

Un réseau est un dispositif qui varie périodiquement la phase ou l'amplitude d'une onde plane incidente (Figure 2.2). La relation fondamentale de réseau qui modélise les maxima principaux d'intensité lumineuse est donnée par (Palmer et al., 2005):

$$a. (\sin\theta_m - \sin\theta_i) = m\lambda, \quad (2.4)$$

où a est le pas de réseau, m ordre de diffraction, θ_i est l'angle d'incidence, θ_m est l'angle de diffraction pour l'ordre m et λ est la longueur d'onde incidente,

Un réseau de diffraction est caractérisé par sa dispersion angulaire et son pouvoir de résolution. La dispersion angulaire (D_a) est définie comme le taux de variation de l'angle par rapport à la longueur d'onde pour un ordre donné :

$$D_a = \frac{m}{a \cos\theta_m} \quad (2.5)$$

Le pouvoir de résolution est la propriété de distinguer deux longueurs d'onde. Il est défini par le critère de Rayleigh. Si on se place à l'ordre m , il est donné par la relation suivante (Loewen and Popov, 1997):

$$P.R. = m N, \quad (2.6)$$

où N est le nombre de fentes éclairées

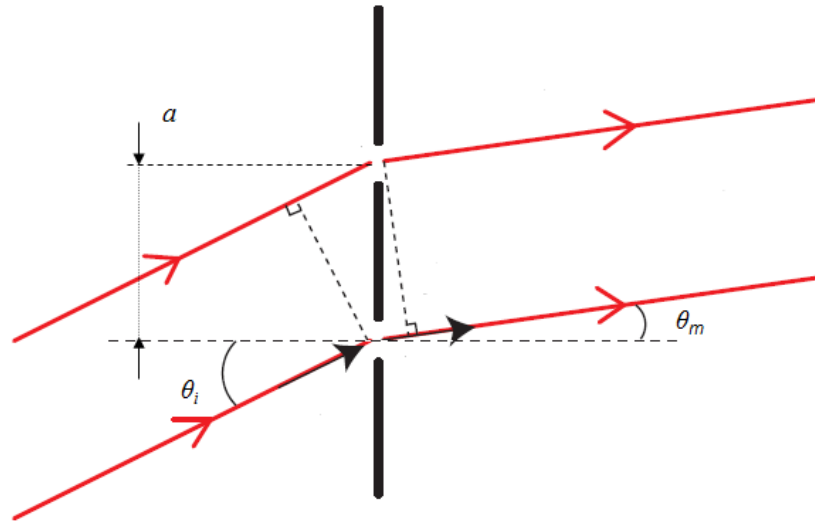


Figure 2.2 Réseau de diffraction

(Loewen and Popov, 1997)

2.1.3 Réseau de Dammann et éléments diffractifs

Les éléments d'optique diffractive (DOEs) sont des réseaux de diffraction utilisés pour diviser un faisceau incident (souvent un faisceau laser) en plusieurs faisceaux répartis de façon prédéterminée (Figure 2.3).

Les réseaux de Dammann sont des DOEs à phase binaire conçus pour séparer un faisceau incident en N faisceaux de puissances égales (Dammann and Görtler, 1971).

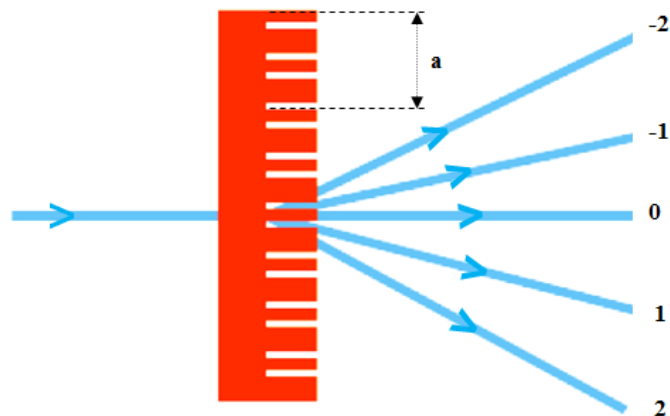


Figure 2.3 Exemple de réseau binaire ou réseau de Dammann

(inspirée de O'Shea et al. (2004))

2.2 Banc de calibrage: DOE setup

Le banc de calibrage réalisé est présenté à la Figure 2.4. Il est composé de six éléments principaux :

- Une source laser HeNe de longueur d'onde $\lambda = 632.8$ nm et de puissance 0.5 mW.
- Un filtre neutre de 2 D.O qui permet d'atténuer l'intensité du faisceau laser.
- Un filtre spatial, constitué d'un objectif de microscope (x20) qui focalise la lumière dans un trou de $25\mu\text{m}$, afin de filtrer la partie non gaussienne et améliorer la qualité du faisceau.
- Une lentille achromatique avec une longueur focale de 400mm utilisée pour collimater le faisceau et éclairer les éléments d'optique diffractive.
- Deux éléments d'optique diffractive (DOEs) mis en cascade avec leurs axes de diffraction parfaitement perpendiculaires. Chaque DOE est un réseau de Dammann (un réseau de phase binaire) conçu pour séparer un faisceau incident en 15 ordres de diffraction de puissances égales. Le pas de réseau est $a = 16.4 \mu\text{m}$ pour les deux DOEs.
- La caméra à calibrer placée à quelques centimètres des DOEs.

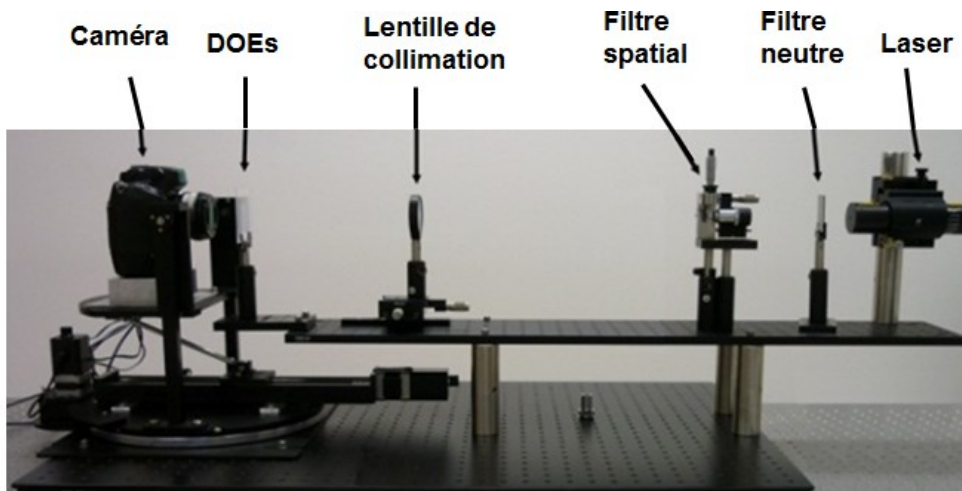


Figure 2.4 Banc de calibrage avec des éléments d'optique diffractive mis en cascade

2.3 Condition générale

Comme la caméra aérienne par exemple dispose d'un très large champ de vue (parfois plus large que 60 degrés pour les capteurs de grand format), les DOEs doit générer de grands angles de sortie. En effet, pour éviter le vignettage, l'ouverture des DOE doit être assez grande pour couvrir la pupille d'entrée de la lentille.

La pupille d'entrée (D_{EP}) de la lentille est à l'intérieur du boîtier d'objectif, par conséquent, les DOEs doivent être surdimensionnés, comme la montre la Figure 2.5.

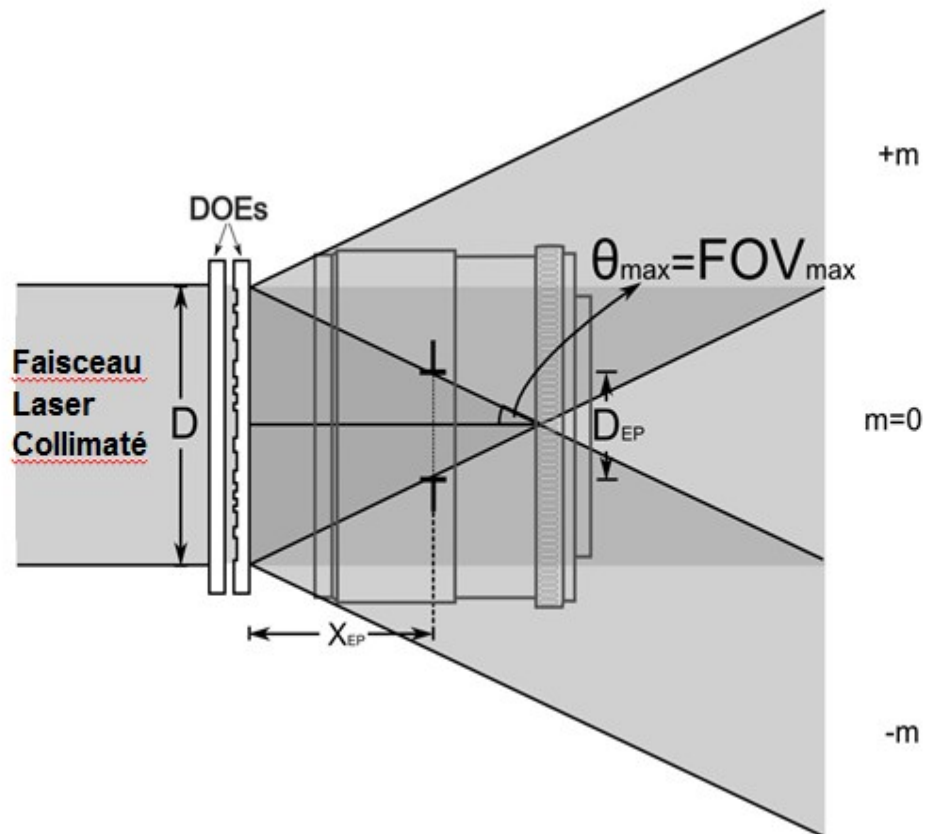


Figure 2.5 Le chevauchement des ordres de diffraction des DOEs et la position de la pupille d'entrée de la caméra à calibrer

(Thibault et al., 2011))

2.4 Modèle mathématique de calibrage

Deux éléments d'optique diffractive (DOEs) mis en cascade avec leurs axes de diffraction perpendiculaires permettent de générer, à partir d'un faisceau laser, une grille de points lumineux (Figure 2.6). Les coordonnées de chacun des points appartenant à cette grille dans un repère espace sont définies, en utilisant les cosinus directeurs et la notation de coordonnées homogènes, par:

$$d = [\lambda f_x \quad \lambda f_y \quad (1 - \lambda^2(f_x^2 + f_y^2))^{1/2} \quad 1]^T \quad (2.7)$$

avec:

λ : longueur d'onde du laser,

(f_x, f_y) : fréquences spatiales des DOEs suivant les deux axes X et Y .

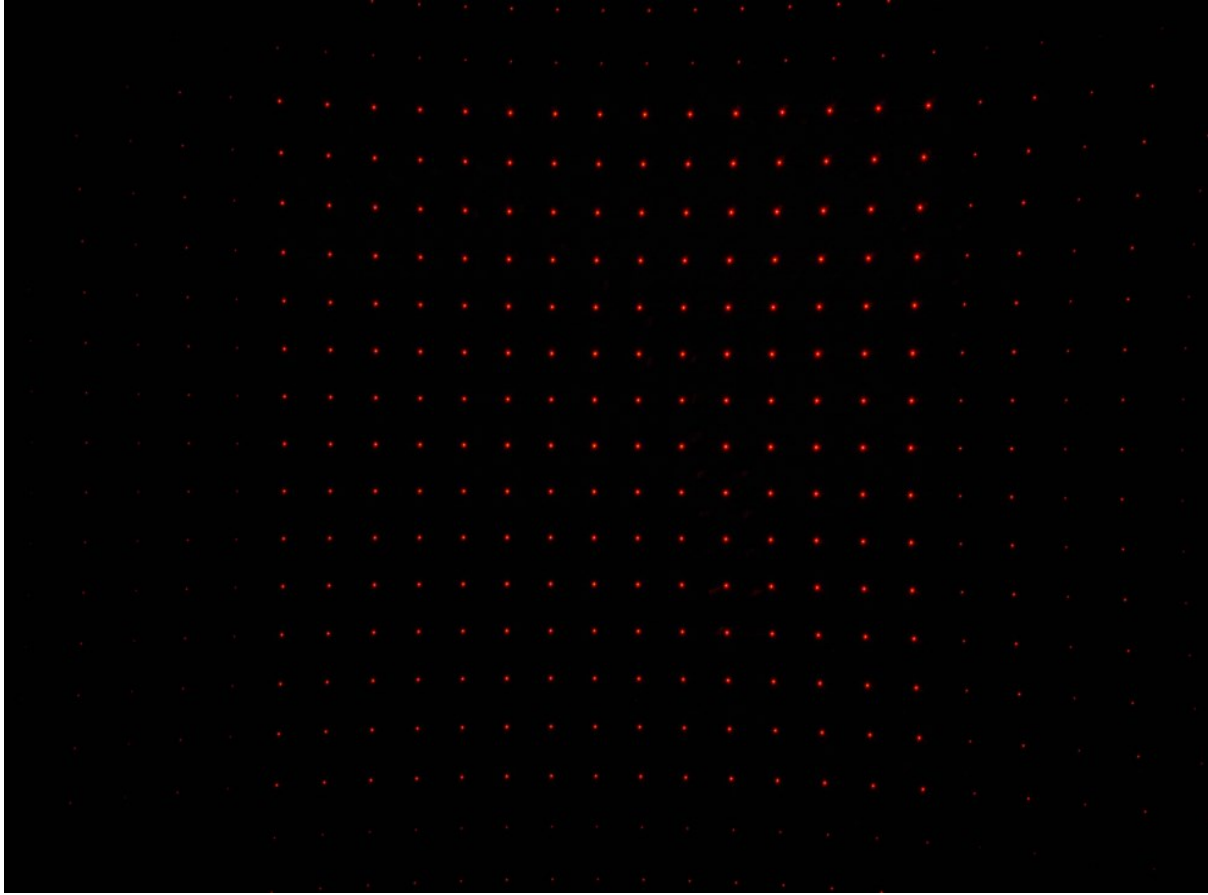


Figure 2.6 Grille formée par les éléments diffractifs sur un capteur CCD

L'image du faisceau laser est renvoyée par la lentille à l'infini. Le vecteur de coordonnées homogènes est donnée par:

$$d' = \begin{bmatrix} \lambda f_x & \lambda f_y & (1 - \lambda^2(f_x^2 + f_y^2))^{\frac{1}{2}} & 0 \end{bmatrix}^T \quad (2.8)$$

En tenant compte de l'angle de l'incidence, le faisceau laser incident s'écrit de la façon suivante :

$$r = [\sin(\beta) \quad -\sin(\alpha) \cos(\beta) \quad \cos(\alpha) \cos(\beta)]^T \quad (2.9)$$

avec (α, β) sont les rotations suivant les deux axes X et Y.

Les directions des faisceaux diffractés sont données ainsi par :

$$d = \begin{bmatrix} \lambda f_x + r(1) & \lambda f_y + r(2) & (1 - (\lambda f_x + r(1))^2 - (\lambda f_y + r(2))^2)^{1/2} & 0 \end{bmatrix}^T \quad (2.10)$$

Comme les axes des deux éléments d'optique diffractive ne sont pas parfaitement perpendiculaires en pratique, on doit tenir compte de l'angle θ_{DOE} entre eux.

Le vecteur de coordonnées spatiales final est décrit par :

$$d = [X \quad Y \quad Z \quad 0]^T \quad (2.11)$$

$$\text{avec : } \begin{cases} X = \lambda f_x + r(1) + (\lambda f_y + r(2)) * \sin(\theta_{DOE}) \\ Y = (\lambda f_y + r(2)) * \cos(\theta_{DOE}) \\ Z = (1 - (X^2 + Y^2))^{1/2} \end{cases}$$

Selon l'équation 1.5, on peut exprimer le modèle de formation de l'image pour une caméra classique par l'expression suivante :

$$\begin{pmatrix} ku_p \\ kv_p \\ k \end{pmatrix} = \begin{pmatrix} f_u & 0 & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 0 \end{pmatrix} \quad (2.12)$$

L'Équation 2.12 est la même que l'équation 1.5 sauf que le facteur d'échelle à gauche du signe d'égalité est égale à 0 (pour signifier que les points sont à l'infini en coordonnées homogènes).

2.5 Détection de points de contrôle dans l'image

Pour détecter les points d'intérêt dans l'image, on a comparé deux algorithmes qui permettent de définir les zones lumineuses (spots) et d'établir leurs caractéristiques (nombre de pixels, leurs intensités, leurs centroïdes).

2.5.1 Algorithme 1: Méthode binaire

Le premier algorithme permet de déterminer les n pixels (u_i, v_i) appartenant à un spot. Ensuite, les coordonnées de son centroïde (C_u, C_v) sont calculées en utilisant les équations suivantes :

$$\begin{cases} C_u = \frac{\sum_i^n u_i}{n} \\ C_v = \frac{\sum_i^n v_i}{n} \end{cases} \quad (2.13)$$

L'algorithme est le suivant :

- Transformer l'image RGB en NG.
- Binariser l'image : Les pixels dépassant une certaine luminosité sont considérés comme du blanc (appartenant aux spots), ceux qui n'atteignent pas le seuil, comme du noir (fond), comme la montre la Figure 2.7.
- Calculer les coordonnées de chacun de spots en utilisant l'équation 2.13.

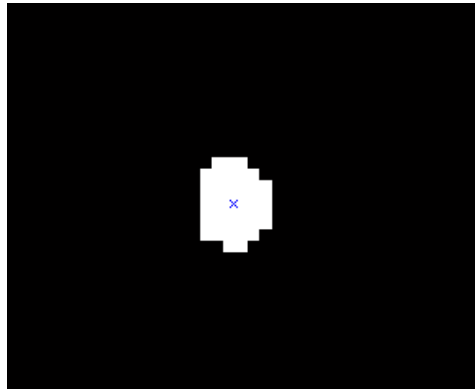


Figure 2.7 Détection d'un centroïde par la méthode binaire

2.5.2 Algorithme 2 : Méthode d'intensité

La deuxième méthode consiste à déterminer les n pixels (u_i, v_i) appartenant à un spot et de déterminer leurs valeurs de gris. En pondérant ces valeurs avec les coordonnées u_i et v_i , on peut calculer le centroïde à l'aide des deux équations suivantes :

$$\begin{cases} C_u = \frac{\sum_i \sum_j F(i, j) * u_i}{\sum_i \sum_j F(i, j)} \\ C_v = \frac{\sum_i \sum_j F(i, j) * v_i}{\sum_i \sum_j F(i, j)} \end{cases} \quad (2.14)$$

où $F(i, j)$ est la fonction intensité du pixel (u_i, v_j) .

L'algorithme est le suivant :

- Transformer l'image RGB en NG.
- Binariser l'image pour déterminer les pixels appartenant aux spots.
- Calculer $F(i, j)$ pour chaque pixel appartenant au spots.
- Calculer les coordonnées de chacun de spots (Figure 2.8) en utilisant l'équation 2.14.

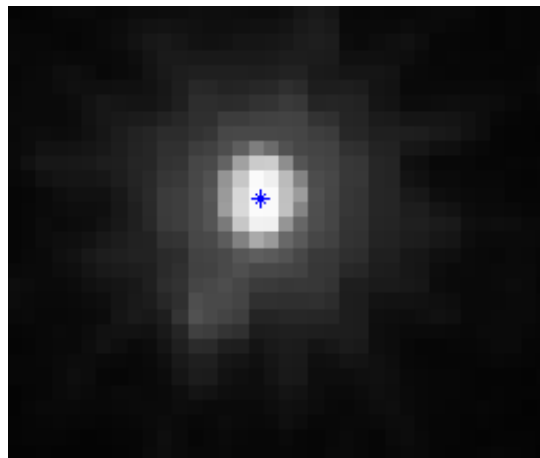


Figure 2.8 Détection d'un centroïde par la méthode d'intensité



2.5.3 Comparaison des deux méthodes

Pour comparer les deux algorithmes, 16 cibles réparties sur toute l'image ont été choisies et leurs coordonnées ont été calculées en utilisant ces deux méthodes pour différentes vitesses d'obturation T (1/8, 1/60, 1/400 s). Ensuite, la différence entre les coordonnées calculées a été estimée.

Tableau 2.1 : Tableau comparatif entre les deux méthodes

Points	T=1/8 s		T=1/60 s		T=1/400 s	
	Δu_i (pixels)	Δv_i (pixels)	Δu_i (pixels)	Δv_i (pixels)	Δu_i (pixels)	Δv_i (pixels)
1	0,002	0,022	0,125	0,052	-0,040	-0,066
2	-0,061	0,028	0,102	-0,017	0,093	-0,175
3	0,046	-0,002	0,068	-0,029	0,118	-0,117
4	0,004	0,013	0,117	-0,066	0,101	-0,074
5	-0,028	0,012	0,009	-0,049	0,030	0,017
6	-0,061	0,009	0,095	-0,028	-0,042	-0,074
7	-0,036	-0,053	0,005	-0,071	0,022	-0,048
8	-0,011	-0,007	0,052	-0,003	0,056	0,001
9	-0,007	-0,039	-0,044	-0,123	0,030	-0,045
10	0	0,021	0,042	-0,049	0,030	-0,174
11	-0,041	-0,097	-0,145	0,009	0,038	-0,015
12	-0,051	0,028	-0,060	0,001	0,019	0,065
13	-0,041	0,027	0,004	-0,013	0,062	-0,001
14	-0,010	-0,003	0,139	0,017	0,176	0,028
15	-0,071	0,059	0,031	0,060	0,078	-0,036
16	0	0	0,064	-0,001	0,040	-0,038
Écart-type (pixels)	0,037	0,036	0,082	0,048	0,073	0,080

On remarque que les résultats de la détermination des points de contrôle par la méthode binaire et la méthode d'intensité sont très proches. Cependant, la méthode d'intensité est jugée la plus précise en comparant les écarts entre les coordonnées calculées lorsqu'on varie la vitesse d'obturation.

2.6 Résultats pour des caméras aériennes

L'étude expérimentale sur le banc de calibrage se scinde en trois parties. Tout d'abord, les résultats de calibrages réels de deux caméras numériques en utilisant la grille 15x15 (grille

principale) sont présentés. Ensuite, l'influence d'une augmentation de points de contrôle en utilisant également certaines cibles secondaires qui apparaissent dans l'image (des points qui correspondent à des ordres de diffraction supérieurs à 15) est étudiée. Enfin, quelques résultats obtenus sont comparés avec ceux calculés en utilisant une mire d'étalonnage.

2.6.1 Conditions de travail

Les principales opérations à réaliser sont :

- l'acquisition de l'image : une fois la caméra est installée (Figure 2.4), l'image est prise dans l'obscurité (Figure 2.6). Ce qui offre un avantage pratique puisqu'il n'y a pas des conditions reliées à l'éclairage ambiant à respecter.
- Calcul des paramètres : l'algorithme de calibrage développé comporte trois fonctions principales. La première fonction permet de détecter les points de contrôle. La deuxième fonction estime les paramètres initiaux en utilisant le modèle linéaire. La troisième fonction calcule les paramètres en se basant sur une optimisation non-linéaire (méthode non-linéaire).

2.6.2 Caméras utilisées

Les caractéristiques des deux caméras qui ont été étalonnées afin de tester le banc de calibrage sont mentionnées dans le tableau ci-dessous.

Tableau 2.2 Caractéristiques des caméras utilisées pour tester le banc de calibrage

Modèle	Distance focale (mm)	Résolution (pixels)	Taille du pixel (mm)
Kiel K2	20 ($f/3.5$)	4288 x 2848	0,0050
Kiel K3	45 ($f/2.8$)	7216 x 5412	0,0068

2.6.3 Calcul des paramètres

Pour calibrer les deux caméras K2 et K3, un ensemble de 225 points de contrôle ont été traités (grille principale 15x15). Nous avons remarqué que des ordres secondaires apparaissent car la fabrication du réseau binaire n'est pas parfaite. Ces points de diffraction

peuvent aussi être utilisés, ce qui augmente le champ pour le calibrage. Les figures suivantes montrent les coordonnées calculées pour les deux caméras en utilisant la méthode d'intensité.

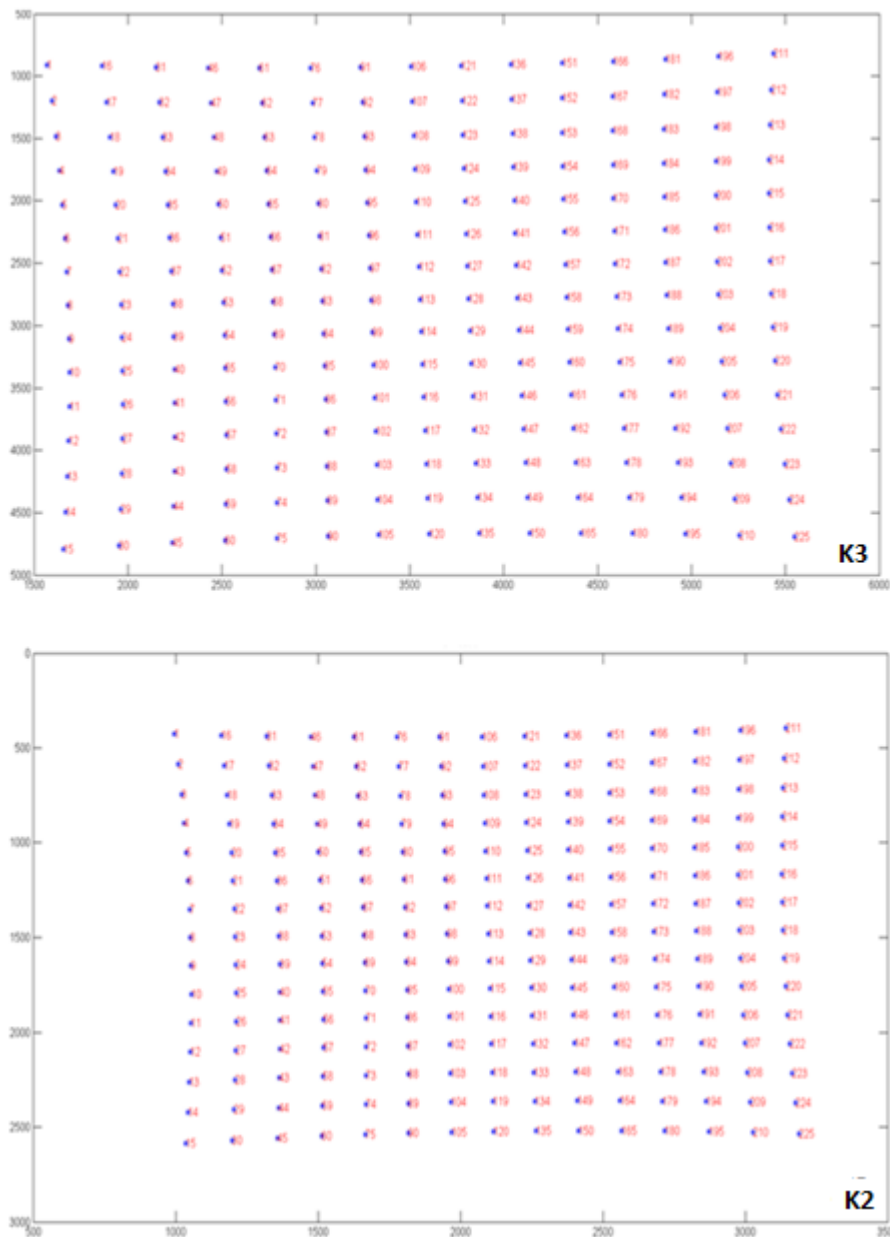


Figure 2.9 Les coordonnées calculées (grille principale) par la méthode d'intensité

Dans un deuxième lieu, la grille principale et les points secondaires (grille hybride) ont été utilisés pour calibrer ces caméras (Figure 2.10).

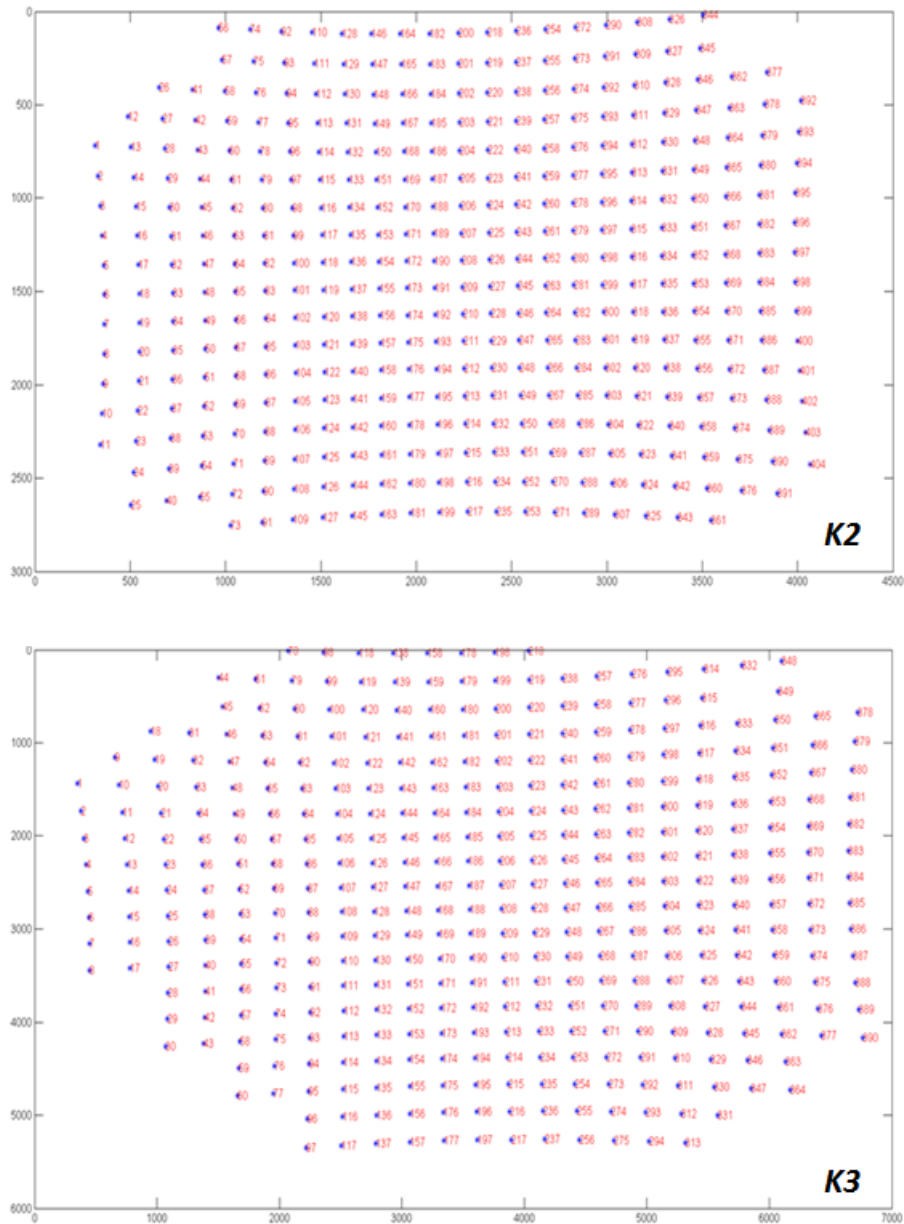


Figure 2.10 Les coordonnées calculées (grille hybride) par la méthode d'intensité

Le but était d'estimer les paramètres de la caméra numérique et des DOEs (équation 2.12). L'algorithme tient compte uniquement de la distorsion radiale (selon les premiers tests, la distorsion tangentielle est négligeable par rapport à la distorsion radiale pour ces caméras) et suppose une modélisation géométrique non-linéaire de la caméra. L'équation 1.16 a été

résolu par optimisation non-linéaire avec une approche de Levenberg-Marquardt (en utilisant la commande lsqnonlin de Matlab).

Les principaux résultats pour les deux caméras sont résumés dans le tableau 2.3.

Tableau 2.3 Paramètres intrinsèques des deux caméras

	Caméra K2		Caméra K3	
Paramètres	Grille principale	Grille hybride	Grille principale	Grille hybride
$f(mm)$	18,68	18,69	45,61	45,61
$u_0 (pixels)$	2158,24	2158,71	3570,36	3570,35
$v_0(pixels)$	1453,17	1455,74	2716,77	2716,56
k_1	$-3,33 \cdot 10^{-04}$	$-3,48 \cdot 10^{-04}$	$-4,05 \cdot 10^{-05}$	$-3,99 \cdot 10^{-05}$
k_2	$7,66 \cdot 10^{-07}$	$8,60 \cdot 10^{-07}$	$2,34 \cdot 10^{-08}$	$2,10 \cdot 10^{-08}$
$EMQ (pixels)$	0,115	0,235	0,191	0,297

Les EMQ (Erreur moyenne quadratique) ont été calculées à partir de la fonction E (équation 1.17) de la façon suivante :

$$RMSE = \sqrt{\frac{E}{n}} \quad (2.15)$$

Une EMQ plus grande a été obtenue en utilisant la grille hybride. Cette différence est peut être expliquée par une incertitude plus grande sur les coordonnées spatiales des points secondaires.

2.7 Comparaison avec un objet de calibrage 3D

Afin de valider l'exactitude de la méthode, les résultats ont été comparés avec une méthode traditionnelle en utilisant une mire de calibrage tridimensionnelle. Cette mire est montée sur un mur dans le laboratoire de métrologie du Département des sciences géomatiques de l'Université Laval. Des cercles de 26 mm de diamètre sont quasiment tous répartis selon un quadrillage régulier de 25 cm. Ce sont leurs centres qui constituent les points de contrôle (Figure 2.11). L'incertitude sur leurs coordonnées est de l'ordre 0,05mm.



Figure 2.11 Mire d'étalonnage de l'Université Laval

Les résultats de ce comparatif pour deux caméras de type K3 sont reportés sur le tableau 2.4.

Tableau 2.4 Tableau comparatif entre les deux techniques de calibrage

Paramètres	Caméra 1		Caméra 2	
	DOEs	Mire	DOEs	Mire
f (mm)	45,61	45,63	45,63	45,63
u_0 (pixels)	3570,35	3569,95	3571,69	3571,99
v_0 (pixels)	2716,56	2716,46	2760,72	2760,2
k_1 (mm ⁻²)	-3,99 10 ⁻⁰⁵	-3,92 10 ⁻⁰⁵	-4,00 10 ⁻⁰⁵	-3,95 10 ⁻⁰⁵
k_2 (mm ⁻⁴)	2,10 10 ⁻⁰⁸	1,64 10 ⁻⁰⁸	1,85 10 ⁻⁰⁸	2,05 10 ⁻⁰⁸
EMQ (pixels)	0,152	0,210	0,161	0,195

Les résultats obtenus par les deux mires de calibrage montrent que les coordonnées spatiales des points de contrôle générés par les DOEs sont exactes. La détermination précise de leurs projections dans l'image permet d'estimer précisément les paramètres d'une caméra classique.

Chapitre 3. Fisheye lens calibration using virtual grid

Dans la première partie de ce chapitre, une technique qui permet de calibrer une caméra fisheye en utilisant notre banc de calibrage est présentée. Ensuite, l'influence du nombre d'images et l'ordre du modèle de projection polynomial sur la qualité des résultats obtenus est examinée. Enfin, les résultats de la comparaison avec une mire de calibrage tridimensionnelle sont analysés.

L'article présenté dans ce chapitre est le suivant :

- ARFAOUI, A. & THIBAUT, S. 2013. Fisheye lens calibration using virtual grid. Applied optics, 52, 2577-2583 (Arfaoui and Thibault, 2013).

Contribution:

Dans cet article, j'ai réalisé la totalité des expérimentations. J'ai effectué l'analyse des résultats et rédigé l'article sous la supervision de Dr. Simon Thibault.

3.1 Résumé

Une technique pour calibrer une caméra fisheye en utilisant des éléments croisés d'optique diffractive est présentée. Cette configuration génère des cibles réelles situées à l'infini, ce qui donne des points de calibrage précis et invariants par translation (modèle de projection réduit). Le calibrage a été réalisé en tournant la caméra autour de deux axes.

Une comparaison de trois modèles de projection et une évaluation de l'influence du nombre d'images sur la qualité des résultats ont été proposées. La comparaison de nos paramètres estimés avec ceux calculés en utilisant un objet 3D montre que notre technique est efficace

3.2 Abstract

We present herein a technique to calibrate fisheye lenses using Cross-Diffractive Optical Elements (CDOEs). The setup generated a robust and accurate virtual calibration grid and the calibration was performed by rotating the camera around two axis. We propose the comparison of three fisheye mathematical models and the evaluation of the number of images in the calibration process. The comparison of our experimental data according to the 3D calibration object results showed that our technique is efficient and reliable.

3.3 Introduction

A fisheye lens is a refractive optical system that provides a wide field of view of about 180 degrees and has a very large apparent depth of field which gives clear pictures without focus adjustment. Therefore, it is used in mobile robots, intelligent vehicles, surveillance and so on. In many such applications which involve accurate quantitative measurements, fisheye camera calibration is a crucial step.

The required accuracy and the nature of the application are the main determinants of the model and the method of fisheye calibration. When lens distortion is not considered, camera calibration consists in the estimation of the perspective projection that represents the relationship between the target coordinates space and their projection in the image (Jain et al., 1995). This projection is based on the fact that the angle of incidence of the ray from an object point is equal to the angle between the ray and the optical axis. However, the

radial distortion component for the fisheye lens is relatively high. Therefore, other projection models were developed and can be divided into three types. The first type is called the radial fisheye model and it describes the conversion from the perspective radial distance of a point in the image to its fisheye distorted image, such as the polynomial fisheye transform (Basu and Licardie, 1993), the division model (Fitzgibbon, 2001). The second type is the fisheye projection and it is based on the principle that the radial distance of an image point depends on the angle of incidence of the ray. The equidistant (Schneider et al., 2009), equisolid (Hughes et al., 2010a), orthographic (Nayar, 1997) and stereographic (Fleck, 1995) models could be briefly mentioned as examples of this. The third type is the field of view (FOV) model (Basu and Licardie, 1995) when the radial distance depends on the field of view of the camera. However, this model is not sufficient to model the complex distortion of the fisheye lens. To this end, it is recommended to use the combination of polynomial and FOV models (Devernay and Faugeras, 2001).

Previous work has shown that fisheye calibration can be performed by using one or several images of a known or an unknown pattern. A calibration room which includes known 3D control points can be established and used to calibrate a fisheye camera with only one image (Schwalbe, 2005). When a known 2D calibration object is used, the camera is fixed and different images can be captured. For each image, the object is placed at different positions and orientations to produce a set of non-coplanar points needed for the non-linear minimization algorithms (Kannala and Brandt, 2004). If the target coordinates in the pattern are unknown, the estimation of camera parameters is called auto-calibration. A set of reliable point pairs in a sequence of images is sufficient to find intrinsic parameters (Kannala and Brandt, 2006, Li and Hartley, 2006, Micusik and Pajdla, 2006). Other kinds of methods used the property that the images of parallel lines converge at two vanishing points, and the line that intersects the two vanishing points also intersects the optical center (Hughes et al., 2010b). Calibration is herein performed by observing two sets of lines in horizontal and vertical directions (Zhu et al., 2009).

In order to get a stable, accurate and compact calibration object, a new setup using cross-diffractive optical elements (CDOE) has been recently proposed and implemented in our laboratory (Thibault et al., 2011). This idea of calibrating a camera from DOEs was introduced in 2008 by Griessbach et al. (Bauer et al., 2008). This laser-based technique

requires only a very limited space which is very efficient to calibrate when an object at the infinity is involved. This is particularly true with long focal length lenses where a very large room is required for calibration. The CDOE also has another advantage over targets on a screen. Target positions must be measured with high precision prior to calibration. Each target will have a residual error associated with the measurement method. In the case of CDOE, the position of each diffractive order is determined by the characteristic of the diffraction gratings. The error on the position of each diffractive order is perfectly determined and only determined by the grating equation. Consequently, the CDOE method will not suffer from random errors due to measurement errors of the calibration targets.

In the case of a fisheye, the short focal length produces a small hyperfocal distance and the calibration will not require an object at the infinity. The hyperfocal distance is a convenient approximation of the infinity based on the depth-of-focus definition. Except for telecentric fisheye, the calibration can be affected over the depth-of-focus. Consequently for very high precision, the calibration must be done using a target at infinity. The CDOE method overcomes this difficulty by imposing real targets located at infinity, which gives precise calibration control points. The calibration problem is then reduced accurately to one of non-translating camera calibration. The CDOE allows different viewing directions simultaneously and provides a very accurate diffraction pattern.

This study aims at exploiting these aforementioned advantages to calibrate fisheye cameras. We propose a technique that requires the diffractive optical virtual grid to be observed by the camera at a few different orientations to cover the wide field of view of the camera. In our previous work (Thibault et al., 2011), we used the CDOE to calibrate a wide angle lens which had a 60 deg full FOV using a single image. This work will use a different approach to calibrate a panoramic lens with a 180 deg FOV. We define in section 3.4 the procedure to cover the entire field of view of the fisheye camera using a diffraction pattern. Section 3.5 presents the mathematical model. Once the mathematical model of the camera calibration using cross-diffractive optical elements was optimized, the calibration procedure was performed in section 3.6. Section 3.7 shows the experimental results and discussion of the calibration approach.

3.4 Setup for fisheye calibration

The calibration pattern consisted of 15×15 equally spaced dots. The camera was mounted in an object to allow the rotation around two axes as shown in Figure 3.1. Detail setup description can be found in Section 2.2.

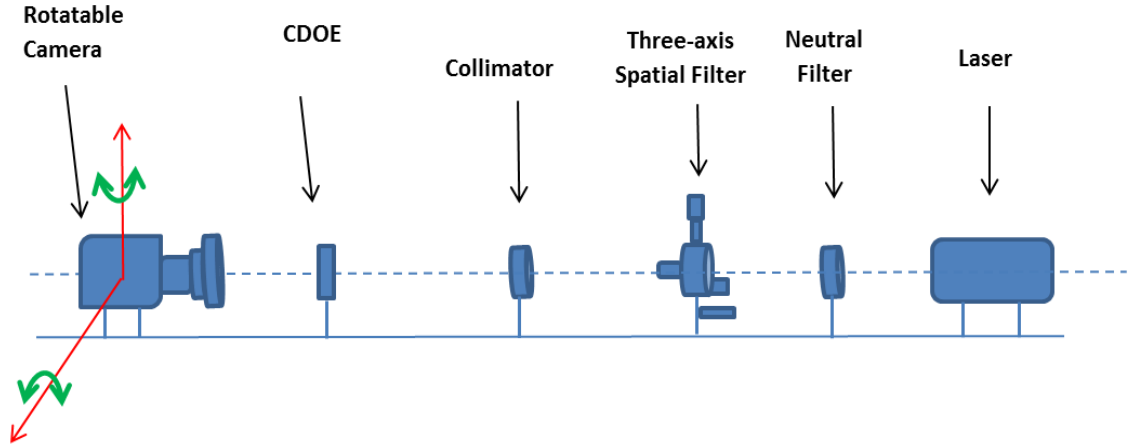


Figure 3.1: Calibration setup with rotatable camera

Since the fisheye lens has a hemispherical view and in order to achieve an accurate calibration, more than one image is needed to fulfill the whole field of view. For each camera rotation angle, one image was taken. Figure 3.2 illustrates the 15×15 grid position in the image for 9 rotation angles (R0 to R8).

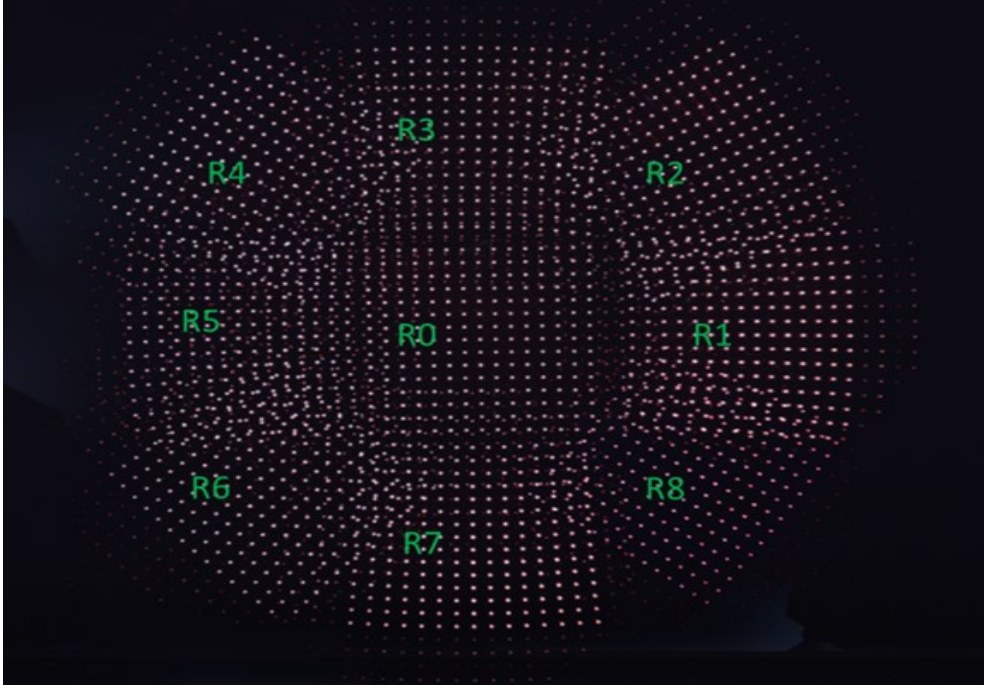


Figure 3.2 : Grid images recorded by fisheye camera in 9 different orientations

3.5 Fisheye calibration with CDOE

3.5.1 Calibration pattern

Following the use of two crossed gratings, Thibault et al. (2011) have shown that the diffractive orders produced by the first grating were not exactly perpendicular to the orders produced by the second grating.

Every spot is a point at infinity and can be presented in homogeneous coordinates:

$$P = [X, Y, Z, 0]^T, \quad (3.1)$$

where:

$$\begin{cases} X = \lambda f_x + r_x + (\lambda f_y + r_y) \cdot \sin(\alpha), \\ Y = (\lambda f_y + r_y) \cdot \cos(\alpha), \\ Z = \left(1 - (X^2 + Y^2)\right)^{1/2}. \end{cases}$$

f_x and f_y are the grating's spatial frequencies, λ is the wavelength of the laser beam, r_x and r_y are errors caused by the non-perpendicular incidence of the laser beam to the DOE plane in

x and y directions with respect to the optical axis (z) and α is the clocking error between the two gratings (in rotation around the z axis).

3.5.2 Fisheye projection

Fisheye lenses do not obey the pinhole geometry, they instead, and usually follow an equidistant projection as:

$$r = f\theta, \quad (3.2)$$

where θ is the angle between the incident ray and the optical axis, r is the radial or the polar distance between the image point and the principal point and f is the focal length.

Occasionally, other projections were used where the sine or the tangent of the incident angle is related to the radial distance in the image by a constant (equisolid, orthographic and stereographic models). Therefore we consider projections in the general form as following:

$$r = k_1\theta + k_2\theta^3 + k_3\theta^5 + \dots, \quad (3.3)$$

$k_1, k_2, k_3, \dots$ are the distortion coefficients. The order of the polynomial can be determined experimentally.

3.5.3 Projection procedures

The fisheye calibration is the determination of the mapping model between points in a world coordinate system and their corresponding elements in an image coordinate system. It is characterized by four transformations (Figure 3.3):

- The beam directions P are transformed into camera coordinates system with the relative rotation R .

$$P' = \begin{pmatrix} R_{3 \times 3} & 0 \\ 0 & 1 \end{pmatrix} \cdot P \quad (3.4)$$

- P' is projected to a point p on a unit sphere centered at the viewpoint O .

$$p = [\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta]^T \quad (3.5)$$

- By assuming that the pixel coordinate system is orthogonal, the relation between image coordinates $[x, y]$ in millimeters and image coordinates $[u, v]$ in pixels is given by an affine transformation as:

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \alpha_u & 0 \\ 0 & \alpha_v \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}, \quad (3.6)$$

where $[u_0, v_0]^T$ are the pixel coordinates of the principle point, (α_u, α_v) are the scale factors of the image axes u and v in pixels.

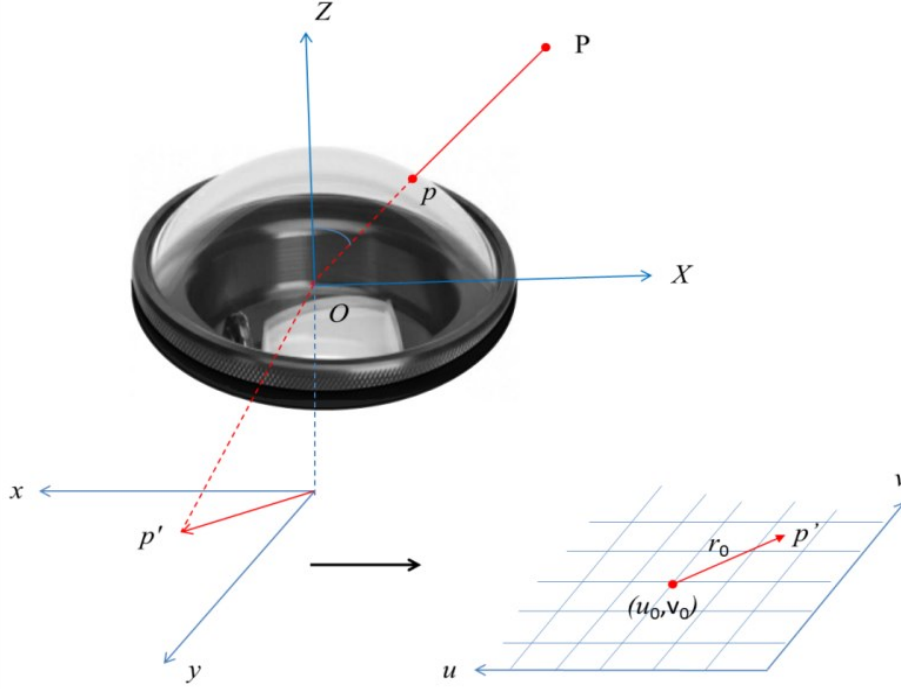


Figure 3.3 Fisheye projection

The goal of fisheye camera calibration is to compute intrinsic and extrinsic parameters. The intrinsic parameters described the internal camera's geometric and optical characteristics (focal length, scales factor and principle point, distortion coefficients). The extrinsic parameters include the position and orientation of the camera (rotation matrix).

3.6 Calibration Algorithm

As mentioned before, the fisheye calibration is the estimation of the extrinsic parameters that describe the orientation of the camera frame relative to the CDOE coordinate system (matrix R), as well as the intrinsic parameters $(u_0, v_0, k_1, k_2, k_3...)$. The process is solved by taking N images of the CDOE pattern in different orientations. In each image, M control points (15×15 spots) were observed. We used the centroid algorithm to achieve image data

extraction. The unknown parameters were estimated iteratively with the non-linear least squares adjustment method. We have optimized the following function:

$$E_{tot} = \sum_{k=1}^N \sum_{j=1}^M (r_{cjk} - r_{ojk})^2, \quad (3.7)$$

where:

$r_c = [(u-u_0)^2 + (v-v_0)^2]^{1/2}$ is the radial distance computed by the model.

r_o is the radial distance observed or measured in the image.

The accuracy evaluation was based on the root mean square error (*RMSE*) calculated from the residuals. Each residual being the difference between observed and computed pixel coordinates. The following diagram (Figure 3.4) illustrates the steps for fisheye camera calibration.

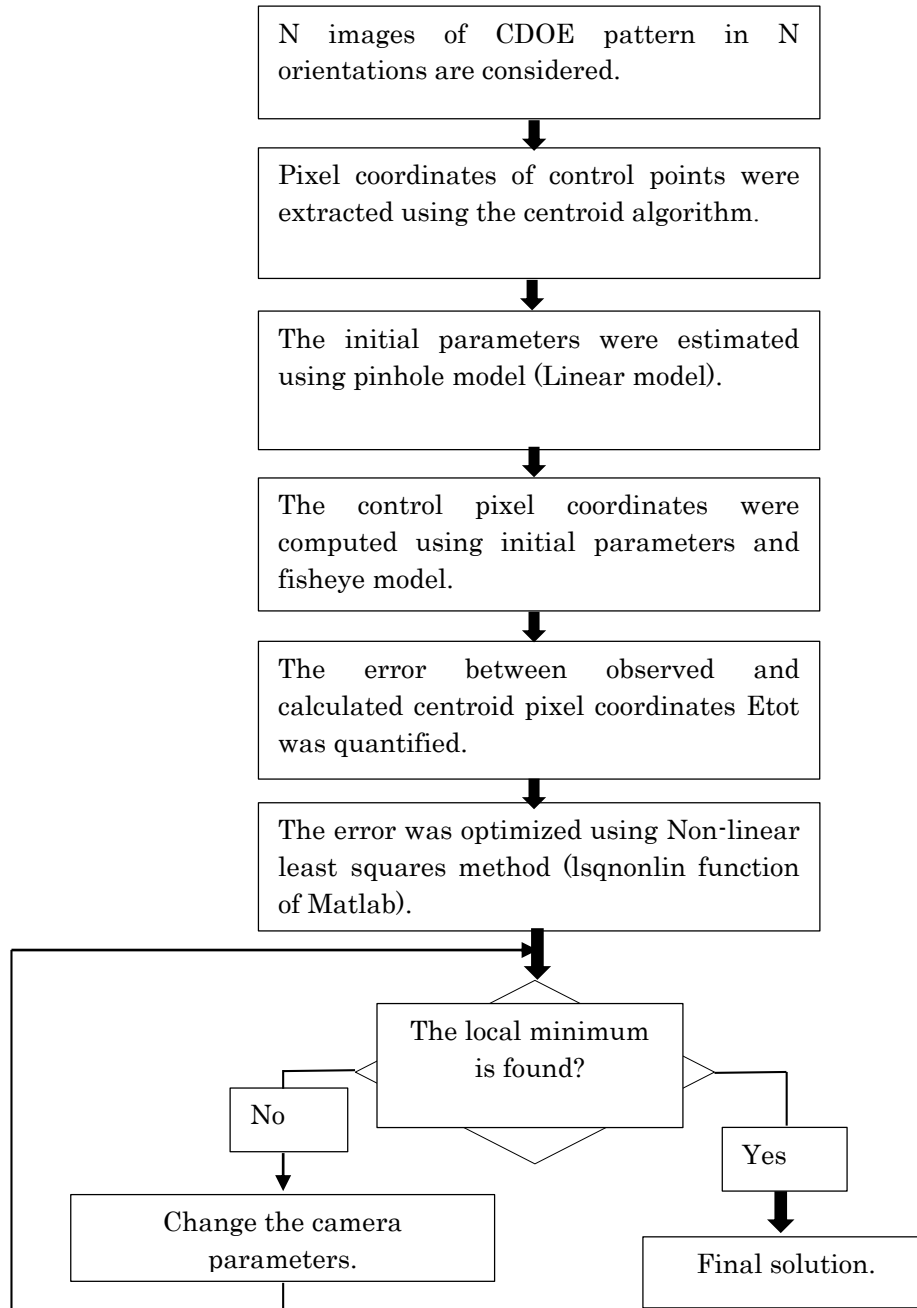


Figure 3.4 Diagram of calibration algorithm

3.7 Results and Discussion

The calibration algorithm was applied to a fish-eye lens with an effective focal length of 1,87 mm ($f/2.8$) and a 180 degree angle of view. The dimension of its image is 1944 x 2592 pixels. The pixel size is 0,0022 mm.

Since the distribution of control points has a strong effect on the calibration accuracy, nine images were taken in varying orientations to fulfill the entire field of view of the fisheye camera (Figure 3.2).

A set of 15×15 points per image gave $225 \times N$ pairs (r_c, r_o) (Eq. (10)). The image coordinates were calculated with subpixel accuracy using a centroid algorithm in gray scale mode. Figure 3.5 shows the computed pixel coordinates for all of the 9 images taken.

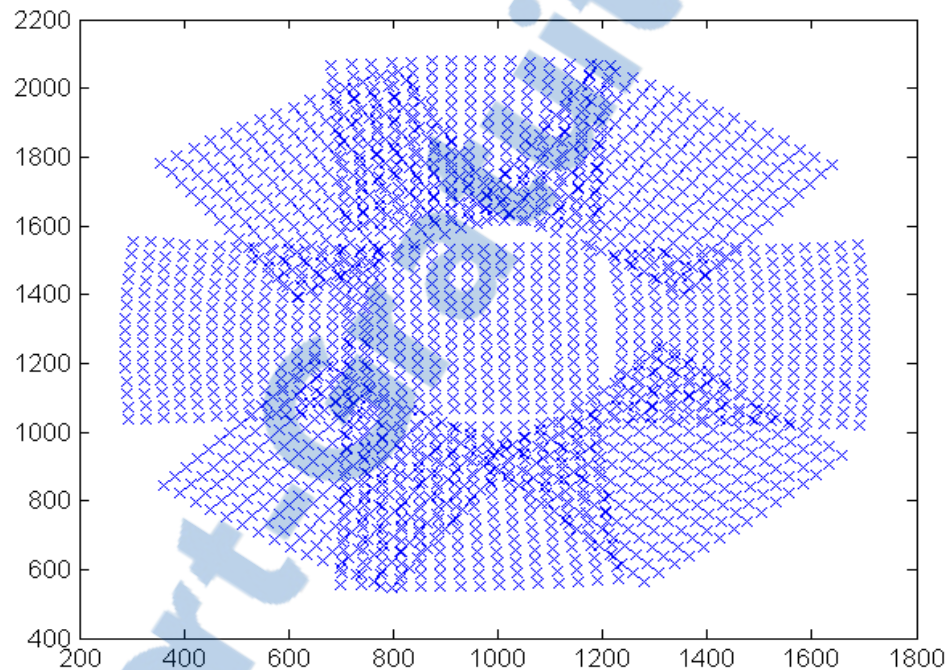


Figure 3.5 Pixel coordinates for all images computed by centroid algorithm

The effect of the number of images on the calibration process was studied by comparing four cases:

- Case 1: calibration using four images (R2,R4,R6,R8)
- Case 2: five images (R0, R2,R4,R6,R8)
- Case 3: eight images (R1,R2,R3,R4,R5,R6,R7,R8)
- Case 4: nine images (R0,R1,R2,R3,R4,R5,R6,R7,R8)

The accuracy of the general model and the impact of the polynomial degree on the RMSE were examined using:

- First order polynomial (equidistant model):

$$r = k_1\theta \quad (3.8)$$

- Third-order polynomial:

$$r = k_1\theta + k_2\theta^3 \quad (3.9)$$

- Fifth-order polynomial :

$$r = k_1\theta + k_2\theta^3 + k_3\theta^5 \quad (3.10)$$

The intrinsic parameters and the corresponding RMSE have been computed for the four image combinations and for the different models. Tables 1, 2 and 3 illustrate the results in pixels.

Table 3.1 Calibration results using the equidistant model

	4 images	5 images	8 images	9 images
u_o	1336,04	1340,40	1340,06	1340,47
v_o	925,85	924,61	924,54	924,38
k_1	847,83	858,76	843,67	843,67
$RMSE$	0,180	0,181	0,183	0,187

Table 3.2 Calibration results using the third-order model

	4 images	5 images	8 images	9 images
u_o	1336,68	1338,80	1339,90	1340,32
v_o	925,33	924,10	923,95	924,22
k_1	857,68	865,17	862,79	862,63
k_2	-10,14	-10,26	-12,42	-14,44
$RMSE$	0,128	0,149	0,155	0,162

Table 3.3 Calibration results using the fifth-order model

	4 images	5 images	8 images	9 images
u_o	1338,29	1339,69	1340,04	1340,11
v_o	923,94	923,95	924,58	924,22
k_1	858,98	864,23	862,74	862,37
k_2	-10,84	-11,20	-13,69	-13,67
k_3	-0,40	-0,41	0,79	0,78
$RMSE$	0.137	0.138	0.147	0.149

Interestingly and as shown in Tables 3.1, 3.2 and 3.3, the results using 8 and 9 images were very close. It can be noted also that the fifth order model returns the lowest error (Table 3.3) as expected. However, the accuracy improvement of this model is not very high compared to the third-order and the equidistant models. The residuals for each control point using 8 images and the equidistant model are shown in Figure 3.6. The figure shows the residual in pixels (Y-axis) for all the 225 points (X-axis) and for different camera orientations (R0 to R8).

Tables 3.1, 3.2 and 3.3 illustrate the results obtained as a function of the number of images used in the computation. Many other combinations of images are studied and the results showed that the stability is ensured with a large number of well distributed control points (cases 3 and 4 for example).

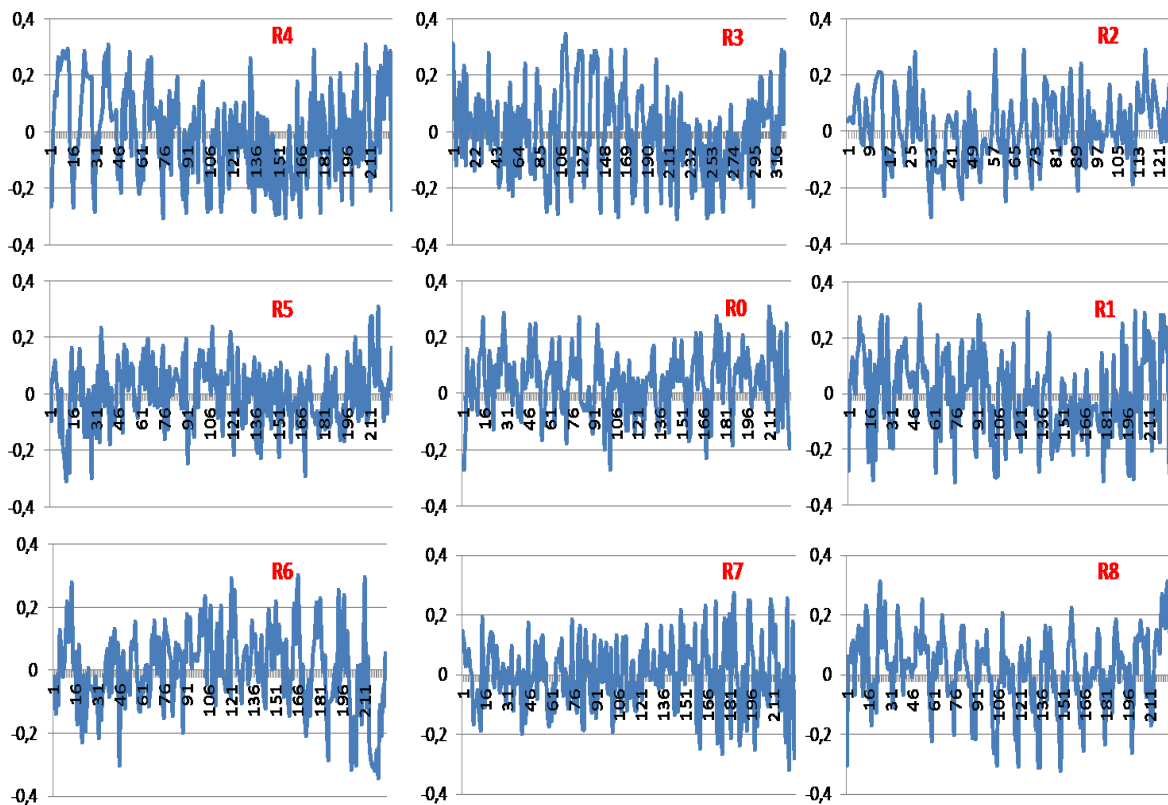


Figure 3.6 Remaining residuals for all images after calibration (in pixels)



The results from the virtual grid results were compared to a 3D calibration object (the focus of the lens is the same for both measurements). This calibration object consisted of 339 targets regularly spaced on perpendicular walls as shown in Figure 3.7. Their geometry in 3D space is known with very good precision (between 0.04 and 0.08 mm). For the determination of the target centres giving the control points values, the ellipse method was used (Arfaoui and Plante, 2011). We applied the calibration algorithm (Figure 3.4) to this set of targets.

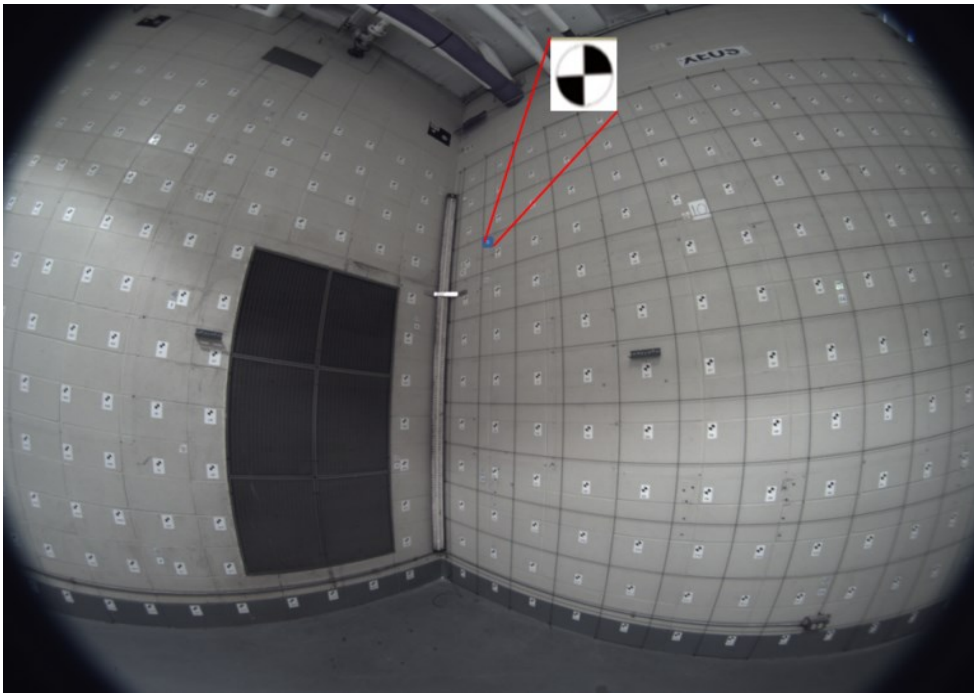


Figure 3.7 Targets of the 3D calibration object

The results of both methods were compared. Table 3.4 shows the calibration parameters. Using the same algorithm but with different calibration objects, we observed that the CDOE method always returns the lowest residual errors as shown in Table 3.5. It seems that we have a different RMSE of 0.08 pixels between the two methods. This error is quite small but we have it in all cases. The difference is not significant and cannot be used to clearly state that the CDOE method is better. However, it clearly shows that the CDOE can be used with success to calibrate a fisheye camera with results that are as good as a standard 3D object method.

Table 3.4 Calibration comparison

	u_o	v_o	k_1	k_2	k_3
<i>CDOE</i>	1340,04	924,58	862,74	-13,69	0,79
<i>3D Object</i>	1339,89	924,01	862,33	-12,95	0,48

Table 3.5 RMSE for fisheye camera calibration using CDOE and 3D object

	<i>First order</i>	<i>Third order</i>	<i>Fifth order</i>
<i>RMSE using CDOE</i>	<i>0,183</i>	<i>0,155</i>	<i>0,147</i>
<i>RMSE using Object</i>	<i>0,263</i>	<i>0,242</i>	<i>0,239</i>

The RMSE of our technique is comparable to the results reported in Schneider et al. (2009) and Kannala and Brandt (2006).

Moreover, to calibrate a camera like a fisheye with only one image, we must use a calibration setup with appropriate dimensions and shape, that is sometimes less convenient. However, most studies have used several observations with different orientations to estimate the camera parameters and they showed that the results are more accurate. The CDOE produces 225 points per frame and allows different viewing directions simultaneously. We have shown that only 8 images, with this large number of control points (225x8), produced a completely and accurately calibrated fisheye camera.

The calibration accuracy depends on the way the control points are measured. Several location methods have been proposed. They were sometimes complex and have provided subpixel accuracy. The DOE herein produces light spots on a black background. The control point coordinates are easily estimated with subpixel accuracy using a centroid algorithm in gray scale mode. The accuracy can be improved in the future using other techniques of pixel coordinate measurements like adaptive thresholding and dynamic windowing methods (Vyas et al., 2009).

3.8 Conclusion

In this study, we have described a method using a compact calibration object for fisheye lens calibration. The proposed setup is based on two commercially available diffractive

optical elements in a crossed configuration which generates a robust and accurate virtual grid suitable for camera calibration. The calibration procedure was successfully applied to real images acquired by rotating the camera. A comparison of three models and an evaluation of the effect of the number of images on the calibration process were done using this technique. The experimental results and the comparison with a 3D calibration object showed that the virtual grid method is efficient and reliable.

The principal limiting factor of this technique is that only one image is not enough to calibrate a fisheye camera. The required accuracy and the nature of the application are used to choose the number of images required for the calibration procedure. Consequently, this technique can be adopted to calibrate different lens types with any FOV (wide-angle and omnidirectional cameras) using images taken by rotating the camera.

The size of the CDOE can also limit the size of the fisheye that can be calibrated using our technique. As discussed in the previous paper (Thibault et al., 2011), the technique will work as long as the entrance pupil of the fisheye can be illuminated by the CDOE (see Figure 2.5). Even for large format fisheye, the entrance pupil is small compared to the front lens diameter. Moreover, the entrance pupil position (from the first lens surface) is also close to the front surface. With our CDOE, we can illuminate a 12 mm entrance pupil as far as 65 mm of the front optical surface of the fisheye. This corresponds to a fisheye ($f/2.8$) of 34 mm focal length. Consequently, our technique can also be used to calibrate larger format fisheye.

In future work, the virtual object grids of dots produced by the CDOE will be used to calibrate zoom cameras. We also intend to test the reliability of this technique for stereo system calibration.

Chapitre 4. Calibration of zoom-lens with virtual optical pattern

Ce chapitre aborde le calibrage des caméras munies d'un zoom. Il décrit les nouveaux éléments d'optique diffractive conçus dans LRIO (Laboratoire de recherche en Ingénierie Optique). Ces DOEs permettraient de calibrer avec précision une caméra à focale variable et d'étudier la variation de ses paramètres intrinsèques en fonction de zoom.

L'article présenté dans ce chapitre est le suivant :

- ARFAOUI, A. THIBAUT, S & DÉSAULNIERS, P. 2015. Calibration of zoom-lens with virtual optical pattern. Optical engineering. 54(5), 054107-054107 (Arfaoui et al., 2015)

Contribution:

Dans cet article, Pierre Désaulniers a pris les images selon mes consignes. Il a également rédigé la description du banc de calibrage expérimental. J'ai réalisé le reste des expérimentations. J'ai effectué l'analyse des résultats et rédigé l'article sous la supervision de Dr. Simon Thibault.

4.1 Résumé

La précision des mires de calibrage et leur fabrication sont parmi les enjeux importants dans le calibrage d'une caméra à focale variable. Cet article présente une nouvelle approche pour résoudre ces problèmes en utilisant deux éléments croisés d'optique diffractive qui ont été conçus spécialement dans LRIO (Laboratoire de recherche en Ingénierie Optique). La grille de calibrage générée fournit assez de points de contrôle pour couvrir le champ de vue (FOV) de la caméra sur toute la plage de zoom. Un modèle mathématique a été proposé et a permis de modéliser les paramètres intrinsèques de la caméra pour toute position de zoom. La comparaison avec un autre objet de calibrage montre que les résultats sont conformes.

4.2 Abstract

The accuracy of calibration patterns, their fabrication and the setup are some of the important challenges in zoom-lens camera calibration process. In this paper, we address these problems by using a cross-diffractive optical element which generates a virtual, dense and robust calibration grid. We show that a 33X33 calibration grid provides enough control points to fill the entire field of view (FOV) for a 10-20X zoom lenses. We show that the calibration of a zoom-camera at infinity can be done using this method. A polynomial function has been used to model the variation of the intrinsic calibration parameters over the zoom range. The obtained calibration model has also been validated using a well-known target pattern.

4.3 Introduction

Camera calibration is a process whereby the geometric camera parameters are determined. A zoom-lens is an optical system in which the axial displacement of lens groups produces a change in the effective focal length and image magnification. In the zoom-lens camera calibration procedure, the objective is to estimate the intrinsic parameters expressed in terms of the adjustable lens parameters. Generally, there are two steps involved in the operation. First, the parameters of the image formation model are computed with non-linear optimization techniques. Second, an adjustable model relating the camera parameters to the zoom position is estimated (Willson, 1994).

A variety of zoom camera calibration techniques have been proposed to compute intrinsic parameters over the entire zoom range using 3D or 2D objects. Many polynomial functions are used to yield lookup tables relating the calibration parameters to the zoom settings (Willson and Shafer, 1994, Li and Lavest, 1996). Linear or non-linear variation functions are also applied to estimate the relationship between principal distance, principal point, and radial distortion in a fixed model and a nonlinear model (Fraser and Al-Ajlouni, 2006). The self-calibration technique differs from previous methods as no calibration object is needed. Computing of the camera's intrinsic parameters is directly made from multiple images without the need for calibration objects. For a given zoom setting, the calibration parameters are estimated by analyzing inter-image homographies computed from the matches in the images of the same scene. Interpolation algorithms are then used to determine how the parameters of the fixed model vary with zoom settings (Seo and Hong, 1998, Agapito et al., 2001). Likewise, a neuronal network method has been proposed to calibrate a zoom-lens camera. One of the initial problems of this technique is the determination of the network architecture. The multi-layer perceptron is proposed to model the nonlinear problem of camera calibration. The weight matrix of the output layer represents the intrinsic parameters (Ahmed and Farag, 2002).

The calibration of lenses with variable parameters poses several challenges. Indeed, calibration for a large zoom range is inconvenient, sometimes not feasible and time consuming. Consequently the calibration is performed at repeated or defined intervals within the zoom range. A sufficient number of well-distributed targets are required to perform accurate calibration for each zoom level. Particularly, at large effective focal lengths for imaging at the infinity, the target must be posed several meters away from the zoom camera which required long calibration laboratory space or range. Furthermore, as the magnification of the zoom system increases, the image size of the targets becomes larger and thus the number of control points within the FOV can become lower than the minimum number of control points needed to carry out the computation of parameters.

Given the constraints related to the calibration setup of cameras with a zoom lens, it would be desirable to have a more compact calibration setup with sufficient control points to be used over a wide zoom range. Recently, it was proposed a new calibration setup based on

the use of a cross-diffractive optical element, CDOE (Thibault et al., 2011). Using CDOE can be very beneficial for zoom calibration because the diffracted virtual sources are points located at infinity in a compact setup. Moreover, the object to be imaged is invariant to translation which facilitates the calibration for each zoom setting. It avoids the need to use bundle adjustments or similar optimization routines between zoom settings. The CDOE setup is stable, compact and provides a very accurate diffraction pattern. It also produces laser light spots on a black background. The control-point coordinates are easily evaluated with sub-pixel accuracy using a centroid algorithm in gray-scale mode. However, the 15X15 dots CDOE proposes previously did not provide a sufficient number of control points for a large zoom range.

In this paper, we show that a 33X33 CDOE provides enough control points to fill the entire field of view (FOV) for a 10-20X zoom lenses. We show that the calibration of a zoom-camera at infinity can be done efficiently using the CDOE method. Our calibration algorithm uses non-linear optimization to compute intrinsic parameters at a particular zoom setting. The relationship between the camera parameters and the zoom settings is then estimated using polynomial functions.

The following section describes the CDOE, fabrication, characterization, and calibration setup. Section 4.5 and 4.6 recall briefly the fixed and adjustable models both adapted to CDOE calibration setup. Section 4.7 presents our calibration algorithm. Finally, experimental results and discussion are presented in section 4.8.

4.4 Cross-diffractive optical element setup

4.4.1 Grating Fabrication

The crossed gratings used in this experiment are one-dimensional Dammann gratings that generate 33 evenly distributed spots. The grating period of these elements is 99.5 microns. The gratings are etched on an S1-UV grade fused silica substrate (Esco Optics Q120063). The specifications of these window substrates are presented in Table 4.1.

Table 4.1 Substrate specifications

Diameter	<i>50, 8 +/- 0,125 mm</i>
Thickness	<i>1,59 +/- 0,125 mm</i>
Surface Accuracy	<i>~5 wave per inch</i>
Surface Quality	<i>60-40, scratch-dig or better</i>
Parallelism	<i><15'</i>
Edges	<i>Fine ground and beveled</i>

The etching process is straightforward. The photoresist (Microposit S1805 positive photoresist from Rohm and Haas Electronic Materials LLC) is manually deposited on the substrate and then spun using an RC8 spin coater from Karl Suss. The photoresist is then baked using a hot plate from VWR. Exposition of the photoresist is done with a UV lamp at 365 nm and a chrome photomask which is positioned using a mask aligner (MJB3 from Karl Suss). Development of the photoresist is manually made with a Microposit MF-319 developer solution from Rohm and Haas Electronic Materials LLC on a spin coater (model WS-400B-6NPP/LITE from Laurell Technologies Corporation). Finally, the exposed fused silica is engraved using the reactive ion etching (RIE) method (Unaxis model 790). The etching depth of the gratings is 628 nm.

4.4.2 Grating Characterization

The produced gratings were characterized using a 632.8 nm laser (Thorlabs HRR005). The 33 spots generated by the Dammann grating are projected onto a screen at a distance of 3.016 m behind the measured grating. The angle of each diffractive order and its respective power are measured. The grating period (Δ) is recalculated for each diffractive order by using the diffraction grating equation $\Delta \sin(\theta_i) = m\lambda$, where m is the diffraction order and λ is the wavelength of the laser). The grating efficiency is evaluated to be 74.29% by measuring and adding the fractional power of each generated spot and comparing this value to the total power output of the laser (6.240 mW). These results are presented in Figure 4.1.

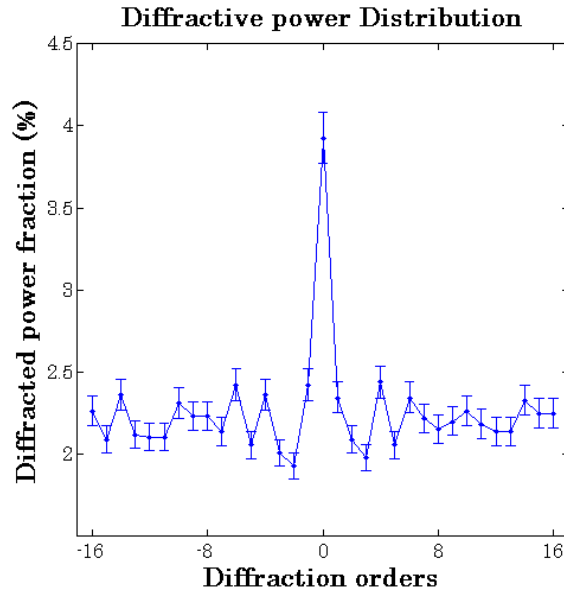


Figure 4.1 Diffraction power distribution of the 33 diffraction orders. The zero-order power is less than the double of the remaining average power fraction

4.4.3 Experimental Setup

In order to generate a 33x33 dot matrix, two one-dimensional Dammann gratings (33 dots) were mounted on a Thorlabs FH2D dual filter holder, with a spacing of 2.54 mm between them. One of the gratings is clocked at 90 degrees around the optical axis of the other. The clocking error between the two gratings and the incident angle deviation from normal of the laser beam are considered into the calculation of the homogenous coordinates of the targets (Thibault et al., 2011). These gratings are placed as close as possible from the entrance pupil of the PTZ camera in order to maximize the light input for each diffractive order. To get a good quality collimated beam impinging on the gratings, the laser light is passed through a three-axis spatial filter having a 25 micron pinhole and then collimated using a precisely positioned 400 mm focal length achromatic doublet collimating lens. The position of the collimating lens is adjusted using a precision linear stage and the collimation quality is assessed by a Shear-plate collimation tester (Melles Griot 09 SPM 001). This setup generates a 2 inches (50.8mm) collimated beam that is split into 33x33 identical beams. These collimated beams are then focused onto the camera sensor by the zoom lens as 33X33 dots. Consequently, the dots are at the infinity for the entire zoom range. The principal components of the setup are summarized in table 4.2. The setup is shown in Figure 4.2.

Table 4.2 Principal setup components illustrated at Figure 4.2

Component		Model
A	PTZ camera	Samsung SNP-5200 (focal= 4,45-89 mm (20x), 1280(H) x 1024(V) pixels, 1/3" 1,3M PS CMOS sensor)
B	Dammann gratings (2x)	2 X (33 X 1) beams, 99.5 microns grating period
C	Collimating lens	Thorlabs AC508-400-A1 (f=400 mm)
D	Three-Axis Spatial Filter	Newport's Model 900 with 900PH-25 High Energy Pinhole
E	Neutral filters	Thorlabs NE220B.
F	Laser	Thorlabs HRR005 HeNe Laser, 632,8 nm, 0,5 mW

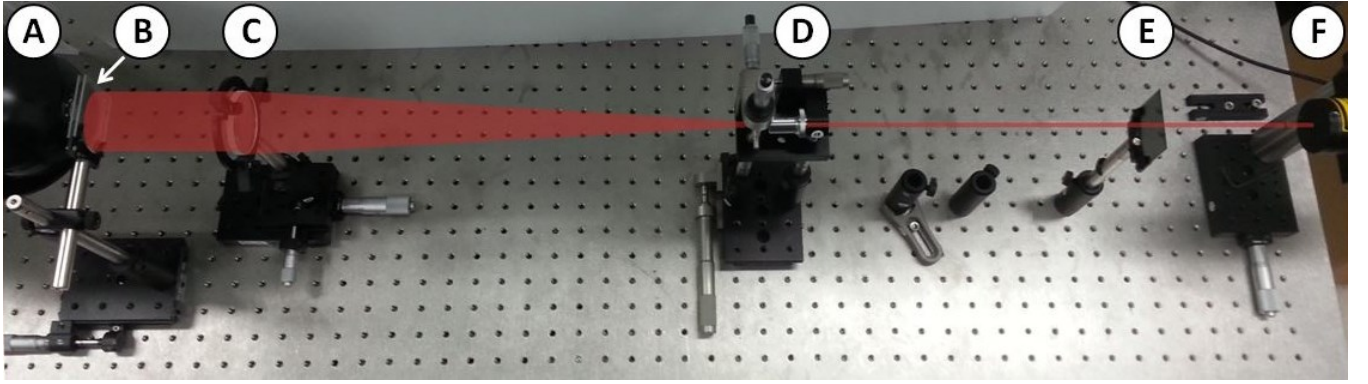


Figure 4.2 Setup of the PTZ camera calibration using crossed diffractive elements

4.5 Fixed camera model using CDOE

In Thibault et al. (2011), a grid diffraction pattern is generated by illuminating the crossed gratings using a collimated laser beam. The relationship between the coordinates of the spot in object space (X, Y, Z) and its image projection (u, v) is given by the linear model as follows:

$$\begin{pmatrix} s(u - u_0) \\ s(v - v_0) \\ s \end{pmatrix} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (4.1)$$

or:

$$\begin{pmatrix} s(u - u_0) \\ s(v - v_0) \\ s \end{pmatrix} = \begin{pmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{pmatrix} R \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \quad (4.2)$$

where R is the rotation matrix between pattern and camera coordinate systems, f is the focal length and (u_0, v_0) are the pixel coordinates of the principal point of the optical system.

A more accurate model must take into account the displacement of the image coordinates caused by distortion. For simplicity, only the radial distortion is considered but azimuthal distortion can also be used. In our case, we determined that the radial distortion is sufficient. The image coordinates can be approximated using the following expression:

$$\begin{pmatrix} \delta_u \\ \delta_v \end{pmatrix} = \begin{pmatrix} u - u_0 \\ v - v_0 \end{pmatrix} (k_1 r^2 + k_2 r^4 + \dots), \quad (4.3)$$

where (δ_u, δ_v) are the displacement due to distortion, (k_1, k_2) : distortion coefficients and r is the radial distance defined as:

$$r = \sqrt{(u - u_0)^2 + (v - v_0)^2}. \quad (4.4)$$

The relationship between the distorted image (u_d, v_d) and the undistorted image (u, v) is expressed as:

$$\begin{pmatrix} e_u \\ e_v \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} \delta_u \\ \delta_v \end{pmatrix} - \begin{pmatrix} u_d \\ v_d \end{pmatrix}, \quad (4.5)$$

where (e_u, e_v) are the residual errors of (u, v) .

We minimized the standard error function to compute the parameters in equations 4.1-4.3 (intrinsic and extrinsic parameters of the camera, clocking error and errors caused by the non-perpendicular incidence of the laser beam to the DOE). The error function is given by:

$$F = \sum_{i=1}^N e_{u_i}^2 + e_{v_i}^2, \quad (4.6)$$

where N is the number of control points.

4.6 Mathematical model for zoom-lens camera

We used a zoom-lens camera standard model which is mostly based on perspective projection (Willson, 1994). The adjustable model represents the variation of the internal

and the external parameters expressed in terms of the zoom setting S_j . The value of S_j is used rather than the zoom values (express in X) because it corresponds to motor units from the controller of the PTZ lens, the values range from 1 to 9999. For each zoom position j , the calibration parameters must change to maintain the proper calibration. For any given zoom position j , it is required to estimate the following parameters matrix:

$$M_j = \begin{pmatrix} R(S_j) \\ [u_0(S_j), v_0(S_j)] \\ f(S_j) \\ [k_1(S_j), k_2(S_j), \dots] \end{pmatrix}, \quad (4.7)$$

where all the parameters are function of S_j .

For the adjustable model, the unknown parameters can be computed for an optimal number (L) of zoom steps ($j = 1..L$) using fixed model and non-linear optimization. In this case, the error function expressed in equation 4.6 is generalized to include L zoom positions:

$$F_L = \sum_{j=1}^L (\sum_{i=1}^N e_{u_i}^2 + e_{v_i}^2)_j, \quad (4.8)$$

where L is the optimal number of zoom positions and N is the number of control points.

4.7 Calibration algorithm for zoom-lens camera

The variable parameters of a zoom-lens camera can be obtained by applying the following steps:

1. L images of the CDOE pattern in L different zoom positions are recorded.
2. The parameters of the fixed model are computed for the L zoom position. For any given zoom setting S_j , the solution is given by minimizing the sum of the errors between the observed and the calculated image coordinates of N control points (equation 4.6). The estimation of the camera parameters requires applying a non-linear calibration technique (Arfaoui and Thibault, 2012).
3. Given L data for each calibration parameters, interpolation algorithms are used to determine how the parameters of the fixed model vary with zoom settings.

Consequently, generic functions relating the internal parameters to the zoom settings are estimated.

4. Global error optimization (equation 4.8) is carried out in order to optimize the adjustable model for L zoom position and to refine interpolation functions.
5. To test the generality of the adjustable camera model and the effectiveness of our technique, the internal parameters are computed for P zoom position using an independent pattern of calibration (being different from that used in step 1). The accuracy evaluation is based on the root-mean-square error ($RMSE$) calculated from the residuals. Each residual is the difference between values computed using patterns 1 and 2

$$RMSE = \sqrt{\frac{\sum_{i=1}^P (m_{1,i} - m_{2,i})^2}{P}}, \quad (4.9)$$

where $m_{1,i}$ is the parameter computed using pattern 1 and $m_{2,i}$ is its equivalent computed parameter using pattern 2.

4.8 Results and discussion

4.8.1 Step 1: CDOE pattern and image recording

Figure 4.3 illustrates the number of control points in an image recorded by the camera. The remotely operated PTZ (pan, tilt, zoom) camera used has a 20x optical zoom with a focal range of 4.45-89 mm (see table 4.2). The dimension of its sensor is 1280×1024 pixels. The camera API is used to control, via an Ethernet port, the zoom-lens by segmenting its phases from 1 to 9999 ($1 < S_j < 9999$).

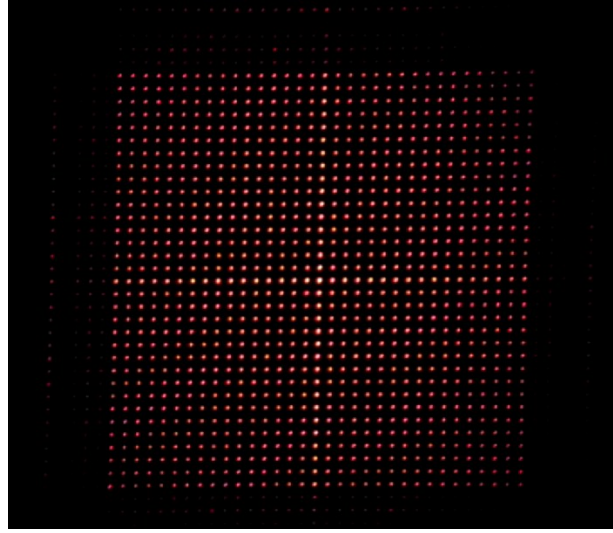


Figure 4.3 CDOE orders calibration grid (33x33)

To perform the calibration at repeated intervals within the entire range of zoom settings, images were recorded at $L = 20$ different zoom positions (Figure 4.4, top).

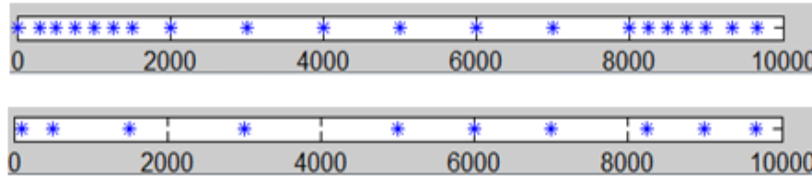


Figure 4.4 Zoom settings (S_i) used to record calibration data for $L=20$ (top). Zoom settings (S_j) required to calibrate the zoom lens for $L=10$ (bottom).

4.8.2 Step 2: Camera calibration with $L = 20$

For each image, pixel coordinates of control points were extracted with subpixel accuracy using a centroid algorithm in gray-scale mode.

The camera parameters are computed for each zoom position using the fixed model and non-linear estimation method. While a more complex distortion model with radial and tangential components could have been used with this camera, experimental investigations have shown that the first coefficient of radial component (k_1) is sufficient to model the zoom-lens distortion. Figure 4.5 shows the values for the internal parameters plotted against the zoom setting S_j .

4.8.3 Step 3: Fitting the models of internal parameters

After an extensive investigation in experimental testing, the polynomial model was found to provide an excellent representation of the internal characteristics of the camera. Focal length f , principal point coordinates (u_0, v_0) and radial component k_1 are polynomial functions of the zoom setting S_j . Each parameter is characterised by a suitable function as follows:

$$\begin{cases} f = a_1 S^3 + a_2 S^2 + a_3 S + a_4 & \text{for } S \leq 6000 \\ f = b_1 S^3 + b_2 S^2 + b_3 S + b_4 & \text{for } S > 6000 \end{cases} \quad (4.10)$$

$$\begin{cases} u_0 = c_1 S^2 + c_2 S + c_3 \\ v_0 = d_1 S^2 + d_2 S + d_3 \end{cases}, \quad (4.11)$$

et

$$k_1 = e_1 S^4 + e_2 S^3 + e_3 S^2 + e_4 S + e_5. \quad (4.12)$$

Many combinations of images were tested. Assuming that each residual is the difference between the observed value of the parameter and the value suggested by the model, the *RMSE* of residuals showed that 10 zoom configurations, as shown in Figure 4.4, are required to fit the polynomial functions to the computed data. Figure 4.6 shows a plot of the camera parameters as a function of the zoom setting with fitted curves of the polynomial functions.

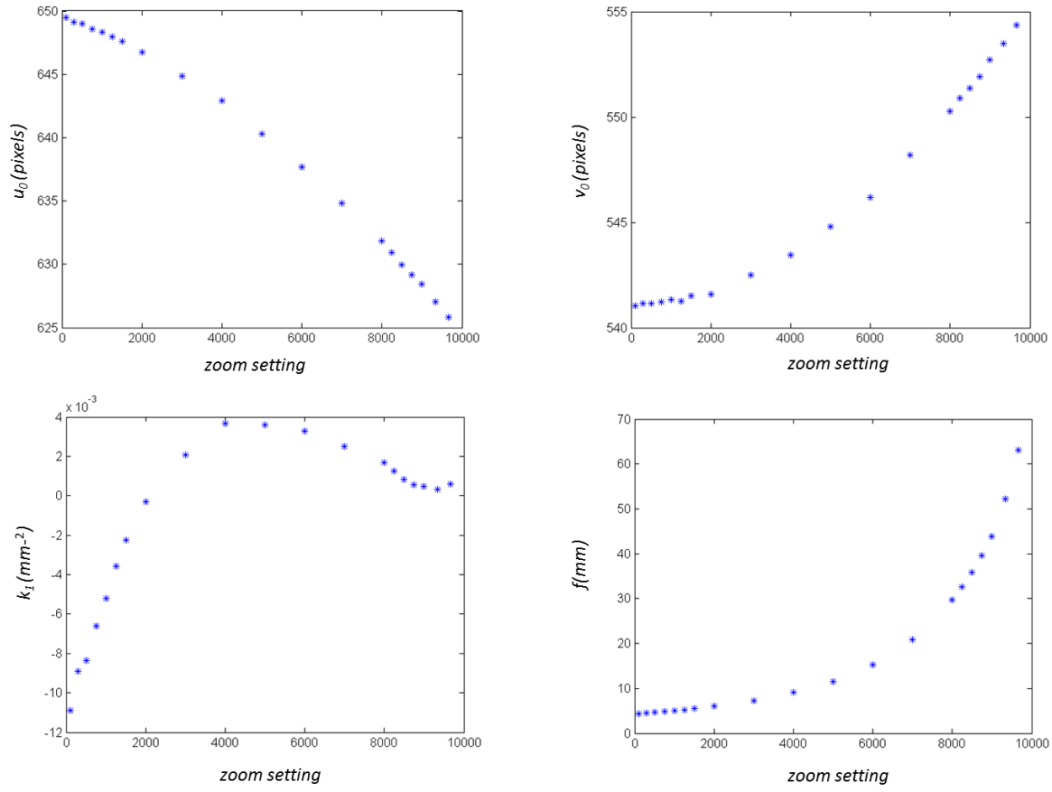


Figure 4.5 Internal camera parameters versus zoom setting S .

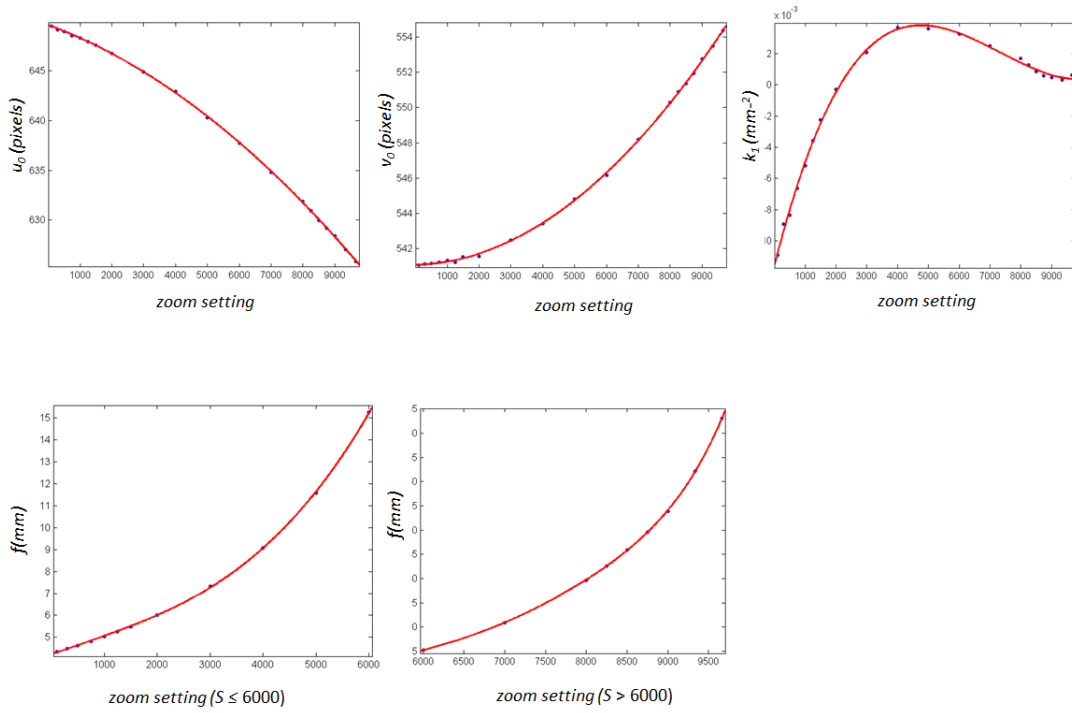


Figure 4.6 Polynomial fit of the internal characteristics of the camera

4.8.4 Step 4: Global error optimization of the adjustable model

After obtaining the initial estimate of adjustable model (polynomial functions) in the previous step, the function F (equation 4.8) is minimized to compute the final parameter estimates. The root-mean-square error ($RMSE_{tot}$) is calculated from residuals on all measurement points as follows:

$$RMSE_{tot} = \sqrt{\frac{\sum_{j=1}^L (\sum_{i=1}^N e_{u_i}^2 + e_{v_i}^2)_j}{L*N}}. \quad (4.13)$$

The model has a good convergence with an $RMSE$ equal to 0.143 pixels. Table 4.3 illustrates the coefficients of the polynomial functions that predict the variation of internal parameters with zoom settings.

Table 4.3 Coefficients of polynomial functions

a₁	3,75E-11	c₃	649,610
a₂	-6,43E-08	d₁	1,38E-07
a₃	0,0008	d₂	3,76E-05
a₄	4,206	d₃	541,11
b₁	3,77E-9	e₁	-4,24 E-19
b₂	-9,21E-05	e₂	6,38 E-14
b₃	0,761	e₃	-1,25E-09
b₄	-2098,01	e₄	7,83E-06
c₁	-1,32E-07	e₅	-0,0116
c₂	-0,0011		

4.8.5 Step 5: Evaluation of the generality and repeatability of the calibration model

To test the model and the effectiveness of our technique, another virtual calibration pattern (Thibault et al., 2011), consisting of 15×15 targets, was used to calibrate the camera over multiple zoom settings. These tests were done at different moment and the extrinsic parameters of the zoom camera were different between both tests.

The $RMSE$ s of the calibration for both patterns are calculated using equation 4.9. Table 4.4 shows the calibration parameters using the two patterns and the $RMSE$ for 10 zoom positions. The calibration values for both patterns are surprisingly similar. Pattern 1 is a 33×33 and Pattern 2 is only half of that (15×15). However, the calibration parameters have

similar values. We think that the higher number of points will help considerably for lenses with higher distortion. Indeed, the relative error for the k_1 parameter is higher than the other parameters. For example, $S_j = 9000$, the relative error is about 14% and less than 4% for $S_j = 1$. The value of k_1 for high zoom position (longer focal length) is closed to zero leading to a larger error as it is small number.

Table 4.4 Comparison of calibration parameters using two different patterns

S_j	f (mm)		u_0 (pixels)		v_0 (pixels)		k_1 (mm ⁻²)	
	Pattern	Pattern	Pattern	Pattern	Pattern	Pattern	Pattern	Pattern
	1	2	1	2	1	2	1	2
1	4,265	4,291	649,60	649,80	541,10	541,12	-0,0109	-0,0105
400	4,545	4,539	649,24	649,43	541,12	541,15	-0,0107	-0,0105
1250	5,248	5,244	648,29	648,47	541,28	541,31	-0,0097	-0,0109
2000	6,009	6,003	646,72	646,86	541,73	541,79	-0,0088	-0,0077
3000	7,306	7,305	644,89	644,97	542,46	542,57	0,0020	0,0018
5000	11,575	11,587	640,42	640,32	544,75	545,03	0,0036	0,0034
6000	15,252	15,289	637,79	637,57	546,31	546,70	0,0035	0,0029
7000	20,890	20,948	634,90	634,54	548,35	548,67	0,0032	0,0028
8250	35,650	35,691	630,70	630,34	551,83	551,55	0,0012	0,0015
9000	43,899	43,842	628,31	628,61	553,75	553,50	0,0007	0,0006
RMSE	0,0325		0,234		0,223		5,96x10-04	

4.9 Conclusion

The use of a compact calibration setup using an accurate CDOE has been proposed to address challenges of zoom-lens camera calibration. The most important challenge lies in producing a sufficient number of accurately positioned targets over a wide range of magnification in limited space.

The calibration setup was successfully applied to images acquired at different zoom positions over the entire camera zoom range. An adjustable nonlinear model has also been used to predict the camera parameter variations according to each zoom position. Furthermore, the effect of the number of zoom positions required for the calibration process has been evaluated. The experiment results and the comparison with a reference pattern showed that the setup is effective and reliable. We also show that a high number of targets is necessary to properly calibrate the distortion specially at long focal length (high zoom).

Finally, we show that the CDOE technique is suitable to calibrate zoom lens in infinite conjugate in a very compact setup. Further works include the calibration of zoom in finite conjugate using CDOE.

Chapitre 5. A new mathematical model for hybrid and panoramic stereovision systems: panoramic to rectilinear conversion model

Ce chapitre présente un modèle générique de calibrage pour un système de stéréovision hybride ou panoramique. Il permet de modéliser un couple de caméras avec la projection perspective et les splines cubiques composées et n'exige pas d'hypothèses ou des configurations particulières de caméras.

L'article présenté dans ce chapitre est le suivant :

- ARFAOUI, A. & THIBAUT, S. 2015. Mathematical model for hybrid and panoramic stereovision systems: panoramic to rectilinear conversion model. Applied Optics, 54(21), 6534-6542 (Arfaoui and Thibault, 2015)

Contribution:

Dans cet article, j'ai établi le nouveau modèle théorique. J'ai réalisé la totalité des expérimentations. J'ai effectué l'analyse des résultats et rédigé l'article sous la supervision de Dr. Simon Thibault.

5.1 Résumé

Pour simplifier la modélisation d'un système de stéréovision hybride (composé d'une caméra classique et une caméra panoramique) ou de stéréovision panoramique (composé de deux caméras panoramiques), on propose un modèle permettant d'unifier les modèles de caméras classique et panoramique. La projection s'effectue en deux étapes. La première étape consiste à projeter un point de l'espace sur le plan image en utilisant le modèle rectiligne (une image parfaite est obtenue). La seconde étape consiste à transformer l'image parfaite en une image distordue en utilisant les splines cubiques composées.

Cette approche permet de simplifier le calcul de la géométrie épipolaire liant les deux images et par la suite la reconstruction tridimensionnelle en utilisant les algorithmes classiques.

5.2 Abstract

This paper introduces a new generic model for a stereovision system composed either of a perspective camera and a panoramic camera, or of two panoramic cameras. We present a new simple model in order to establish matching images and to formulate 3D reconstruction. Indeed, we show that composed cubic splines functions are suitable for conventional and wide-angle cameras and model accurately tangential and radial distortions. Matching is greatly simplified using this technique and classical methods can be used to evaluate the relationship between image pairs and to achieve 3D reconstruction. Such a technique allows to model complex forms of distortions as is the case for panomorph lenses.

5.3 Introduction

Stereo vision is a technique for 3D reconstruction of a scene observed from two or several viewpoints. It is a challenging research problem with a wide array of applications including photogrammetry, robot navigation, virtual reality, security and industrial automation (Barrois and Wöhler, 2013, Sabattini et al., 2012).

A number of stereo vision systems have been proposed in the literature and can be classified in two methods. The first method uses one or more cameras which are configured

to produce 3D measurements (Nedevschi et al., 2004, Cucchiara et al., 2007). The second method is performed by combining one camera with an optical system such as a projector (Li et al., 2014) or multiple combined mirrors (Gluckman et al., 1998, Jang et al., 2005).

The conventional method uses one (McBride et al., 2005) or more (Wan and Zhou, 2008) perspective cameras to capture stereo images, extract the matching targets (stereo correspondence) between them and compute the estimated 3D world coordinates through stereo triangulation. Many applications, such as robot navigation, require a large field-of-view (FOV). Since the FOV of a perspective camera is small, two solutions were proposed to overcome this problem. The first technique was performed using a hybrid stereovision system. It was based on the combination of wide angle and classical cameras such as catadioptric and perspective (Sturm, 2002) or multiple central camera types (Caron and Eynard, 2011). The second solution used to enhance FOV was the panoramic stereo imaging system. Indeed, several techniques that combined two identical or different wide FOV cameras have been discussed in the literature (Svoboda and Pajdla, 2000, Nishimoto and Yamaguchi, 2007). In most of these cases, each proposed system had its own applications and benefits.

An essential step in the stereovision technique is the camera calibration procedure. It is a rigorous step in which the extrinsic and intrinsic parameters are to be estimated for each camera. When the most frequently used model to calibrate classical cameras is the rectilinear model, various solutions have been offered to accurately model panoramic lens. As an example, models which were proposed for fisheye lenses are mainly classified into several types: equisolid, orthographic, stereographic and polynomial (Arfaoui and Thibault, 2013).

Another step in the 3D reconstruction procedure is to solve the stereo correspondence problem. It is often required to produce mappings between images of the scene. Classical stereovision techniques used the epipolar constraints to reduce the stereo matching region from the whole image to a straight line (Xu and Zhang, 1996). When an alternative model other than perspective projection is used to model cameras, the correspondence problem

becomes more complex, mainly for a hybrid or a panoramic stereovision system (Abraham and Förstner, 2005).

This study aims to contribute to this field by proposing a compact algorithm to calibrate and rectify either a hybrid stereo system that comprises a classical camera and a panoramic camera or a panoramic stereo system. Since panoramic cameras do not satisfy the same geometric constraints as perspective cameras, the algorithm uses the rectilinear model improved by composed cubic splines functions to calibrate the cameras. Indeed, we show that the parameters of these functions are able to control more complex forms of distortions as is the case for wide angle cameras. Using this improve model to model the radial and tangential distortions the panoramic camera can be seen as a classical camera, which simplifies the epipolar geometry and 3D reconstruction. As a result, the epipolar geometry for a hybrid or panoramic system can be exploited similarly as in the case of standard perspective cameras. In other words, the epipolar conics (complex geometries) for a hybrid or panoramic stereovision system are replaced by epipolar lines. The algorithm requires the application of a linear method to get an initial solution (Arfaoui and Plante, 2011). Then, composed cubic splines functions and camera parameters are computed using an iterative algorithm in order to obtain a true rectilinear model for the each camera. This method makes it possible to do the direct mapping of the reference image (classical image) to an area in the target image (panoramic image), and to obtain 3D reconstruction using classical techniques.

The development of this new technique has been motivated by the increased potential of panoramic cameras for providing accurate measurements and by the more complex geometry that these cameras involve. Indeed, the panormorph camera for example was conceived to use distortion as a design parameter in order to provide a variable pixel/angle ratio and to increase the resolution in a particular FOV (Thibault, 2007). Moreover, panoramic lenses become more and more popular on the market for consumer product. Combining classical and panoramic cameras is something that we will see in the near future.

The paper is organized as follows: Section 5.4 presents the mathematical models used to calibrate and to test stereovision systems. Section 5.5 illustrates the calibration procedure

and the test method. Section 5.6 reports experimental results of three actual cases of systems. Finally conclusions are drawn in Section 5.7.

5.4 Mathematical formulation

5.4.1 Classical model

The rectilinear (pinhole) model is based on collinearity principle and it describes the mapping from 3D point $P(X,Y,Z)$ in the scene to a 2D point $p(u_p, v_p)$ in the image. The equation of the model is given by:

$$\begin{pmatrix} ku_p \\ kv_p \\ k \end{pmatrix} = \begin{pmatrix} f_u & 0 & u_0 & 0 \\ 0 & f_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (5.1)$$

where (u_0, v_0) are the pixel coordinates of the principal point, (f_v, f_u) the focal length on the v and u axis, R and T are rotation and translation between scene and camera coordinate systems,

Or by:

$$\begin{cases} u_p = \frac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \\ v_p = \frac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \end{cases} \quad (5.2)$$

M , the projection matrix, describes the intrinsic $(u_0, v_0, f_u$ and $f_v)$ and extrinsic parameters (R and T) of the camera and has the following form (Arfaoui and Plante, 2011) :

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix}$$

Using θ , the angle between the incident ray and the optical axis, and φ , the angle between the radial direction of the point and the u direction axis of the image, the rectilinear model can be written as:

$$\begin{aligned} u_p &= f_u \tan\theta \cos\varphi \\ v_p &= f_v \tan\theta \sin\varphi \end{aligned} \quad (5.3)$$

where f_u and f_v are respectively the focal length measured on the u and v axis.

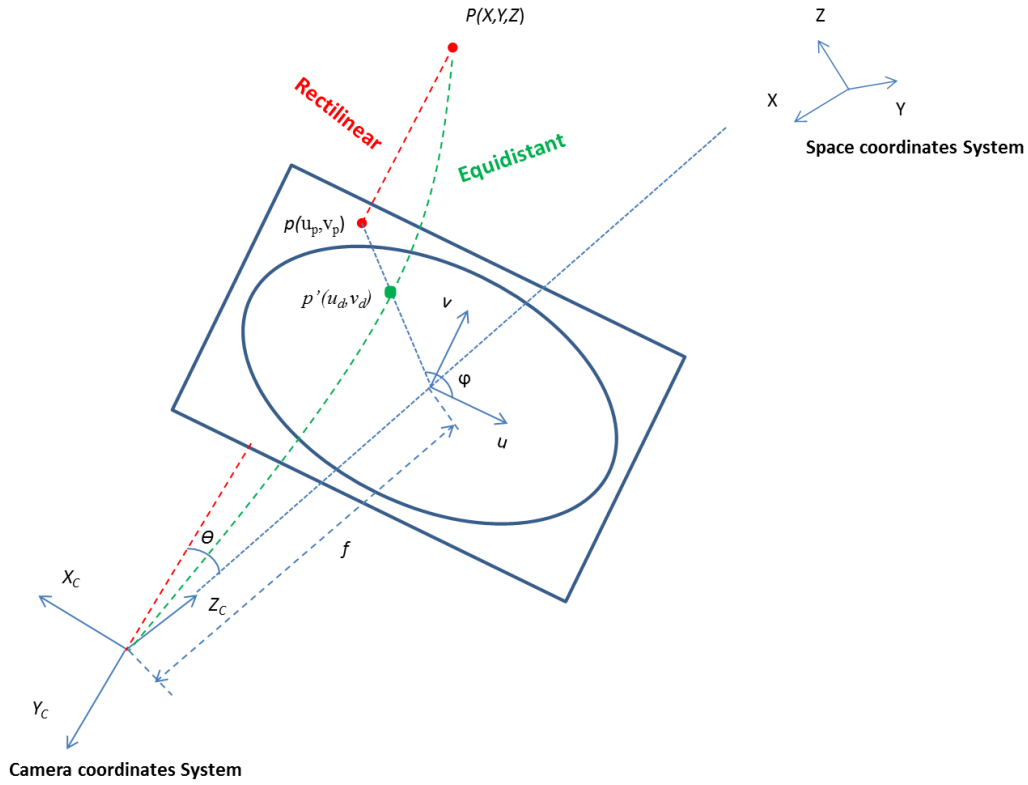


Figure 5.1 Rectilinear and equidistant projection

The prementioned linear model describes the ideal projection and a better suited model must take into account the displacement caused by radial and tangential distortions. Therefore, the nonlinear model is used to describe the deviation from the perspective projection caused by radial and tangential distortion. It has the following form in general:

$$\begin{cases} u_p = u_d - \Delta u \\ v_p = v_d - \Delta v \end{cases} \quad (5.4)$$

where:

- (u_p, v_p) : the undistorted or ideal image coordinates,
- (u_d, v_d) : the distorted or true image coordinates
- $(\Delta u, \Delta v)$: the relative corrections to the undistorted coordinates, giving the true image coordinates in the u and v directions.

5.4.2 Fisheye and Panomorph camera models

Even if our approach is to apply the rectilinear model to our panoramic camera, it must be noted that fisheye and panomorph cameras are naturally designed to obey other kinds of projections. There are several fisheye projections such as equidistant projection (Figure 1) where the radial distance on the image is directly proportional to the angle of the incident ray (Arfaoui and Thibault, 2013).

The equidistant projection function is given by the equation:

$$\begin{cases} u_d = f_u \theta \cos \varphi \\ v_d = f_v \theta \sin \varphi \end{cases} \quad (5.5)$$

In the case of a panomorph camera, the variation of the resolution within the FOV calls for a more complex projection function as described by the equation (6). The projection function can be written as:

$$\begin{cases} u_d = f_1(\theta) \theta \cos \varphi \\ v_d = f_2(\theta) \theta \sin \varphi \end{cases} \quad (5.6)$$

where f_1 et f_2 are functions of θ .

As we can see in equations 5.5 and 5.6, both fisheye and panomorph lenses required a specific projection model which make them difficult to use as stereovision system. This is why we propose to use the rectilinear model improved by composed cubic splines functions to calibrate those two panoramic cameras. Of course, the rectilinear model cannot be used for angle of 90 degrees. Consequently we limit the angle to 85 degrees in this paper.

5.4.3 Composed cubic Splines functions (CCS functions)

A Composed Cubic Splines function models a theoretical form (surfaces and curves) using a combination of linear terms (Plante, 1999). The interpolation formula gives an accurate result both in terms of parameters and errors compared to other method of interpolation. An additional advantage comes from the fact that the nodes can be placed anywhere on the surface. It was used to generate data fitting and to model distortion for classical camera (Arfaoui and Plante, 2011, Plante, 1999). It provides a more realistic model of distortion which is represented as a composed surface. In the present application, it maps a distorted

image into a perfect image that obeys the perspective model. With composed cubic splines functions, the processing challenges of a hybrid stereovision system or a panoramic stereovision system are turned into a simple classical stereovision problem. For the classical/panoramic image pair, the relationship between distorted coordinates (u_d, v_d) and ideal pinhole coordinates (u_p, v_p) is described by the following equations:

$$\begin{cases} u_p = u_d - \sum_{i=1}^N a_i \sqrt{(u_d - u_{ki})^2 + (v_d - v_{ki})^2}^3 \\ v_p = v_d - \sum_{i=1}^N b_i \sqrt{(u_d - u_{ki})^2 + (v_d - v_{ki})^2}^3 \end{cases} \quad (5.7)$$

where:

- N : number of control points
- a_i : spline cubic parameters in the u direction of the i^{th} control point.
- b_i : spline cubic parameters in the v direction of the i^{th} control point.
- (u_{ki}, v_{ki}) : distorted (observed) coordinates of the i^{th} control point.

The continuous functions model the local transformations in the u and v directions for a classical or panoramic image in order to eliminate the effect of lens distortion and to transform a distorted image into a true pinhole image.

The a_i and b_i parameters are determined by minimizing the residuals of N control points between their observed values (u_{ki}, v_{ki}) and the corresponding (u_{di}, v_{di}) computed using equation 5.7. (u_{di}, v_{di}) are solution of equation 5.7 for a given set of a_i and b_i . The following equation has to be minimized:

$$S_1 = \sum_{i=1}^N (u_{ki} - u_{di})^2 + (v_{ki} - v_{di})^2 \quad (5.8)$$

5.4.4 Epipolar geometry and Fundamental matrix

The epipolar geometry is the geometry of corresponding rays in a couple of stereo images. The fundamental matrix (denoted F) describes this correspondence in a 3x3 matrix. For a

given point p_1 in the first image, its epipolar geometry l_p in the second image satisfies the following relation:

$$l_p = Fp_1 \quad (5.9)$$

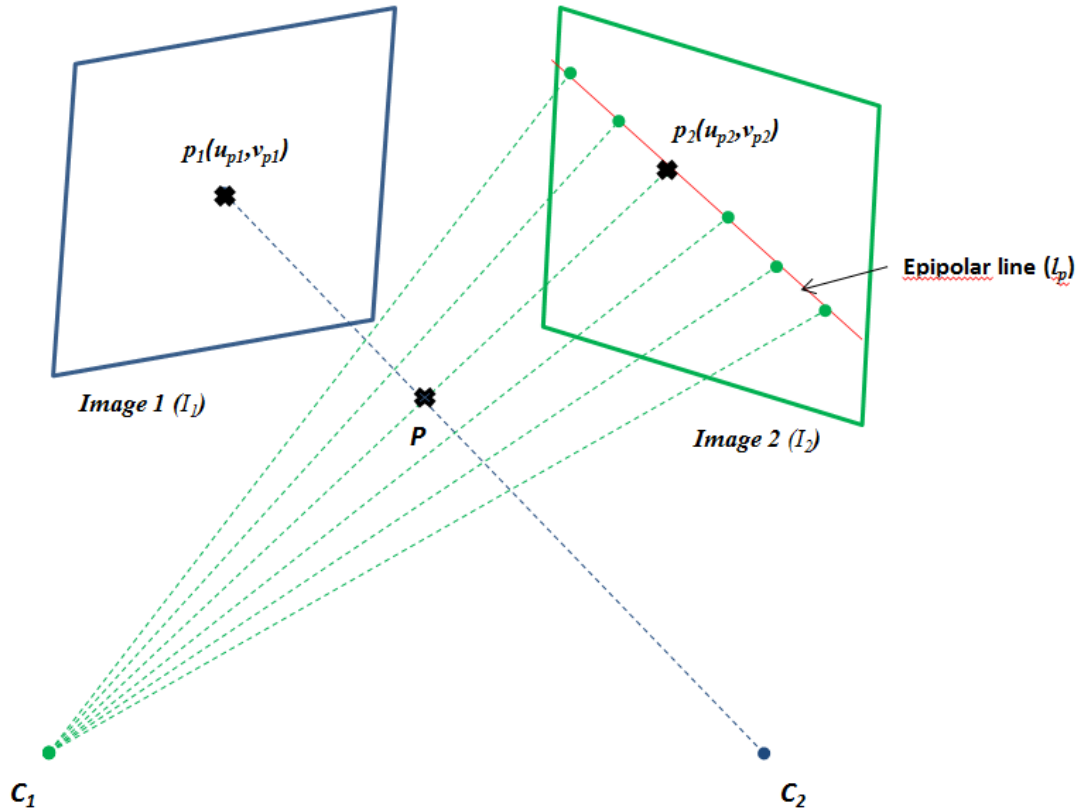


Figure 5.2 Epipolar geometry

If p_2 is the corresponding point of p_1 in the second image, it must belong to l_p and the epipolar constraint is written as:

$$p_2^T F p_1 = 0 \quad (5.10)$$

If the intrinsic parameters of the cameras are known, the fundamental matrix is reduced to an essential matrix E (Luong et al., 1993). The defining matrix is given by:

$$q_2^T E q_1 = 0 \quad (5.11)$$

where:

- $q_1 = K_1^{-1} p_1$
- $q_2 = K_2^{-1} p_2$
- K_1 et K_2 is the matrix of intrinsic parameters respectively of the camera 1 et camera 2

Appendix A presents a method to estimate the best fundamental matrix from a set of corresponding points in Image 1 and 2.

5.4.5 Retroprojection and triangulation

If the couple (p_1, p_2) are images of the same point P respectively in Image 1 and Image 2, the triangulation process is to find the intersection of the two rays corresponding to these images in space. If projection matrix (M_1, M_2) are known, we have the following relation:

$$\begin{pmatrix} u_1 M_1^{3T} - M_1^{1T} \\ v_1 M_1^{3T} - M_1^{2T} \\ u_2 M_2^{3T} - M_1^{1T} \\ v_2 M_2^{3T} - M_2^{2T} \end{pmatrix} P = 0 \quad (5.12)$$

where M_1^{iT} and M_2^{iT} are respectively the i^{th} row of M_1 et M_2 .

The resolution method of the triangulation process is provided in Appendix B.

5.5 Stereovision calibration algorithm

In the present work, the camera calibration is the estimation of parameters that describe the rectilinear projection (Equation 5.1) as well as the composed cubic splines parameters. The process is solved by taking N images of the CDOE pattern in different orientations for panoramic camera to cover the wide FOV (Arfaoui and Thibault, 2013) and an image for classical cameras. Figure 5.3 illustrates an example of 15x15 grid position in the panomorph image for 9 rotation angles.

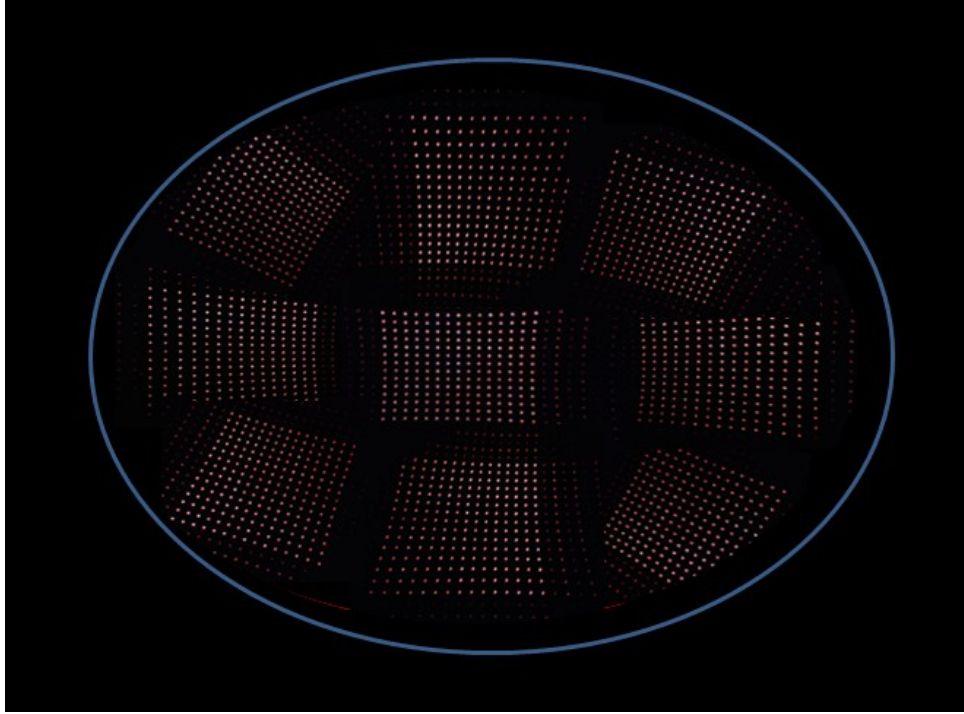


Figure 5.3 Grid images recorded by fisheye camera in 9 different orientations

The validation procedure carries out using 3D targets of Circular Features. It is divided into four major steps. First, given the coordinates of a point in each image captured by the stereo system (perspective/panoramic or panoramic/panoramic), a composed cubic splines function will determine the coordinates of the same point in the corrected image. Second, the intensity value should be assigned to its new position. Third, the epipolar geometry technique is used to compute corresponding points in the stereo pair. Finally, the triangulation process determines the 3D space location for each target.

The diagram in Figure 4 illustrates the steps for stereo system calibration

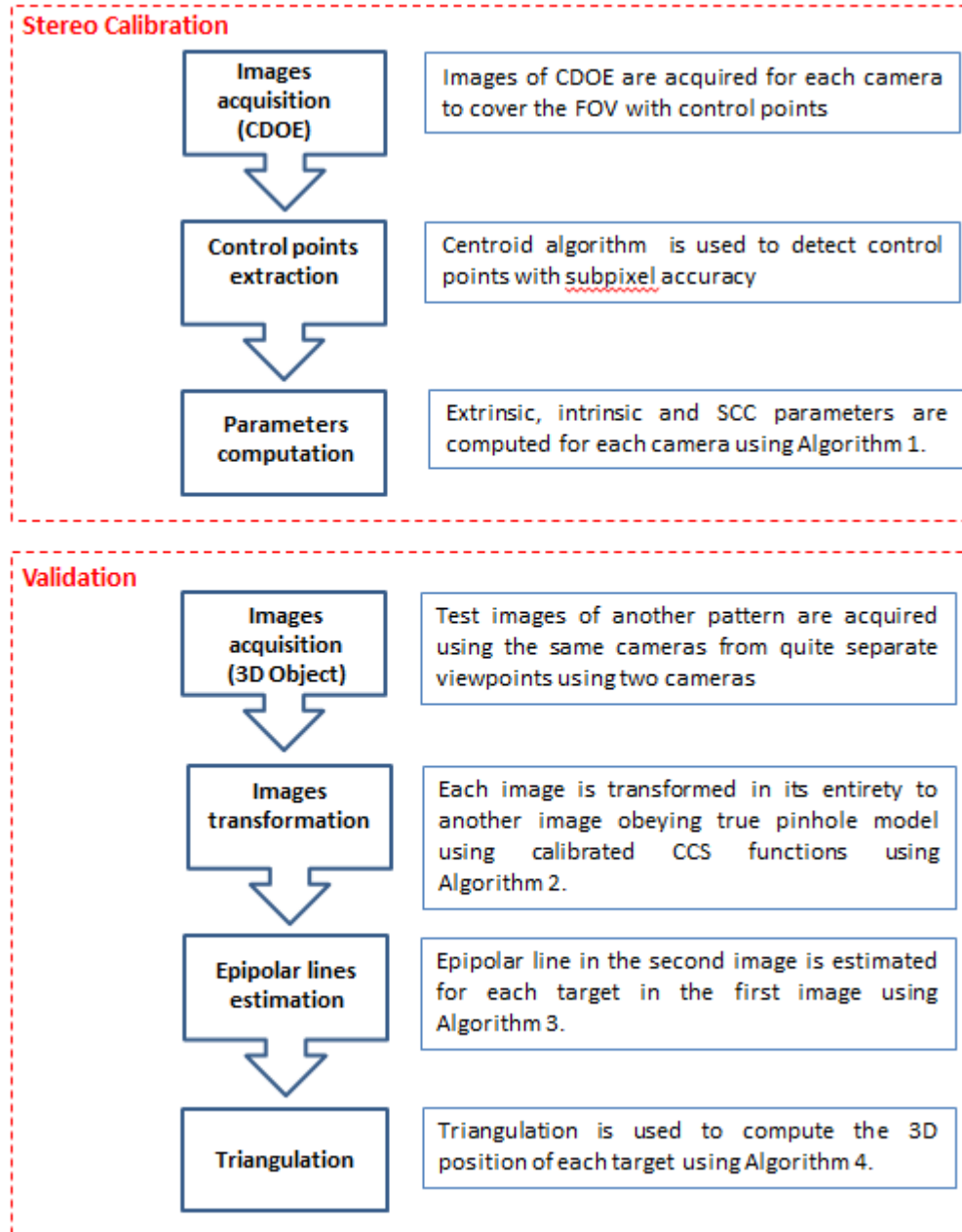


Figure 5.4 Diagram of stereovision calibration

The pseudo codes of calibration and verification are illustrated in Appendix C.

5.6 Experimental results and discussion

The calibration algorithm was applied to following cameras:

- AVT Guppy F-503 CMOS sensor (1/2,5 inch, 2592X1944) with a fisheye lens (SenKo model HFE1916C) with an effective focal length of $1,87\text{ mm}$,

- AVT Guppy F-503 CMOS sensor (1/2,5 inch, 2592X1944) with the panomorph lens (Immervision model IMV1-1/3) with anamorphic ratio of 1,3 and focal length of 1.15 mm on the long axis and 0.9 mm on the short axis,
- Samsung SNP-5200 PTZ (pan, tilt, zoom) camera (1/3 inch 1329x1049) that has a 20x optical zoom with a focal range of 4,45-89 mm. For this study, it is used with a fixed focal length ($f=6$ mm).

Calibration procedure was performed using virtual grid. It consisted of 15×15 equally spaced diffracted dots used by Thibault et al. (2011). For fisheye and panormorph cameras, nine images were taken in varying orientations to fulfill the entire FOV (Figure 5.3). For PTZ camera, an image of CDOE pattern was taken at focal length of $f=6$ mm.

The image coordinates were calculated with subpixel accuracy using a centroid algorithm in gray-scale mode (Thibault et al., 2011).

Intrinsic, extrinsic and CCS parameters for the three cameras were determined using Algorithm 1 (Appendix C). Table 5.1 illustrates the computation results of intrinsic and extrinsic parameters. For every control point, two CCS parameters are estimated along the u and v axis (a_i and b_i in equation 5.7).

Table 5.1 Intrinsic parameters for the three cameras

	<i>Panomorph</i>	<i>Fisheye</i>	<i>PTZ</i>
<i>Control points number: N</i>	2025	2025	1089
f_u (mm)	1,12	1,67	6,00
f_v (mm)	0,88	1,71	6,01
u_0 (pixels)	1249,85	1340,11	646,72
v_0 (pixels)	1017,83	924,22	541,73

To verify the calibration performance and to check the accuracy of 3D reconstruction using a stereo pair, a 3D grid consisting in circular targets regularly spaced on perpendicular walls was used. The images were taken from different points of view in such way that there is a common area in the FOV of cameras (at a distance of about 3.5m from the wall). Algorithm 1 (Appendix C) was applied a second time with known intrinsic and CCS parameters (computed in the calibration step), to estimate the new extrinsic parameters (R and T). Then, Algorithm 2 (Appendix C) was executed to convert a distorted image into a

rectilinear image by adjusting the position of each pixel. Figure 5.5 illustrates the results for the fisheye and panomorph images.

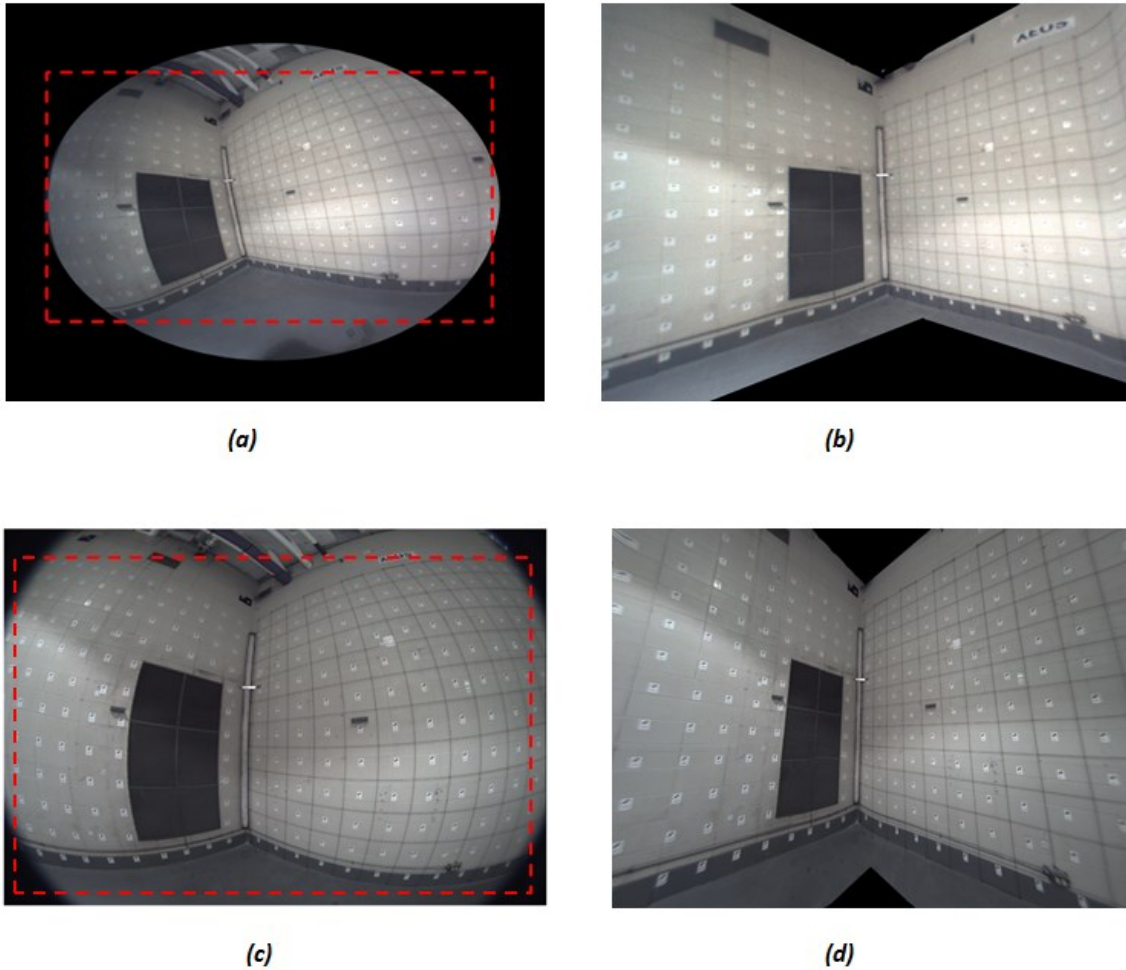


Figure 5.5 Stereo pairs rectifications

(a) original panomorph image, (b) rectified panomorph image, (c) original fisheye image, (d) rectified fisheye image

Generally, the projections of 3D vertical, horizontal, and diagonal lines in the image are curves because of the lens distortions. As shown in Figure 5, in rectilinear images, lines in the space are displayed as straight lines.

Using the undistorted image, it is easy to search correspondence pixels by applying Algorithm 3 to N_1 control points. Figure 5.6 shows five points in the first image and their epipolar line in the second image for PTZ/Fisheye, PTZ/Panomorph and Fisheye/Panomorph stereo pairs.

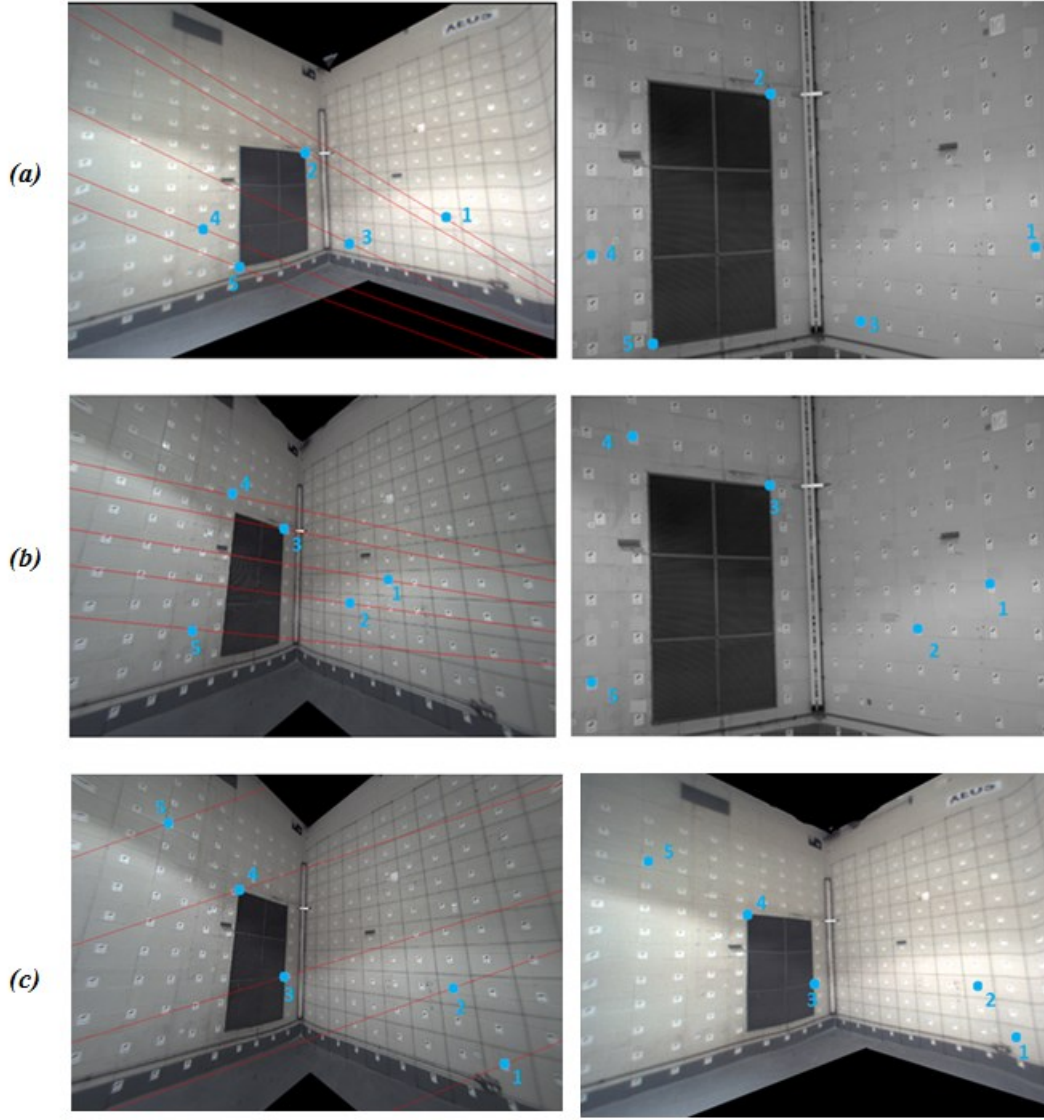


Figure 5.6 Geometry epipolar estimation.

(a) Panomorph/PTZ, (b) Fisheye/ PTZ, (c) Fisheye/Panomorph

The evaluation of the epipolar geometry computation is based on the root mean square error ($RMSE_{epi}$) calculated from the residuals for every pair as follows:

$$RMSE_{epi} = \sqrt{\frac{\sum_{i=1}^{N_1} d(p_2^i, Fp_1^i)^2 + d(p_1^i, F^T p_2^i)^2}{N_1}} \quad (5.13)$$

The results for the three pairs are summarized in Table 5.2 (baseline was about 1,5m for reference , at mean distance of about 3,5 from the targets).

Table 5.2 RMSE for the epipolar geometry computation

	N_I	<i>RMSE (pixels)</i>
<i>PTZ/Fisheye</i>	50	0,245
<i>PTZ/Panomorph</i>	50	0,251
<i>Fisheye/Panomorph</i>	50	0,290

When the intrinsic and the extrinsic parameters of cameras were known and the matched pair of each point was estimated, the triangulation step was applied to find the 3D position (Algorithm 4 in Appendix C).

The triangulation accuracy depends on the accuracy of camera calibrations and the way the matched points are measured in the image (matched error). The function f_{trian} (equation B.1) was minimized to compute the final 3D position. The root-mean-square error was calculated from space coordinate residuals on all measurement points N_{tri} as follows:

$$\begin{cases} RMSE_X = \sqrt{\frac{\sum_{i=1}^{N_{tri}} (X_{io} - X_{im})^2}{N_{tri}}} \\ RMSE_Y = \sqrt{\frac{\sum_{i=1}^{N_{tri}} (Y_{io} - Y_{im})^2}{N_{tri}}} \\ RMSE_Z = \sqrt{\frac{\sum_{i=1}^{N_{tri}} (Z_{io} - Z_{im})^2}{N_{tri}}} \end{cases} \quad (5.14)$$

where (X_{io}, Y_{io}, Z_{io}) are the space coordinates observed for a point P_i and (X_{im}, Y_{im}, Z_{im}) the space coordinates computed by triangulation.

Table 5.3 illustrates the results in X, Y and Z directions for the three pairs.

Table 5.3 RMSE in X, Y and Z directions for the three pairs.

	N_{tri}	<i>RMSE_X (mm)</i>	<i>RMSE_Y (mm)</i>	<i>RMSE_Z (mm)</i>
<i>PTZ/Fisheye</i>	109	0,62	0,71	0,88
<i>PTZ/Panomorph</i>	109	0,55	0,59	0,91
<i>Fisheye/Panomorph</i>	174	0,78	0,89	0,75

The maximum error deviation of estimated coordinates along the 3D Object is less than 1mm. The *RMSE* evaluate the accuracy of this technique and shows that the CCS functions precisely model high and complex component of distortion such as the case for panomorph and fisheye cameras.

5.7 Conclusions

In this paper, a new calibration procedure has been presented for either a mixed stereovision system composed of a perspective camera and a panoramic camera or a panoramic stereovision system. This simplify approach has been developed to address new possibilities and challenges in the field of panoramic stereovision.

Perspective/fisheye, perspective/panomorph and fisheye/panomorph stereovision pairs have been used to investigate the efficacy of this technique. The procedure is divided into three major steps. First, the calibration procedure was performed using but not limited to CDOE setup. Then, the epipolar geometry relating the pair of stereo images was computed. Finally, 3D positions of real targets were computed using triangulation. The experimental results showed that the rectilinear projection improved by composed cubic splines functions form an accurate model for both classical and panoramic cameras. The parameters of composed splines cubic functions control the more complex model of new panoramic lenses such as is the case for the panomorph camera.

In future work, we will perform some tests using planar calibration patterns to examine their influence on calibration and 3D reconstruction accuracy.

Discussion générale, conclusions et perspectives de recherche.

Dans cette thèse, nos contributions à la recherche dans le domaine de la photogrammétrie et de la vision numérique ont été présentées, en particulier aux deux problèmes fondamentaux dans la reconstruction tridimensionnelle: le calibrage des caméras et la stéréovision.

Pour le calibrage géométrique de caméras, un nouveau banc de calibrage à l'aide des éléments d'optique diffractive (DOEs) a été présenté, qui permet de calibrer plusieurs types de caméras (classique, munie d'un zoom, panoramique). Un modèle mathématique a été présenté, qui permet de modéliser la relation entre la grille produite par les DOEs et sa projection sur l'image en fonction de paramètres intrinsèques et extrinsèques de la caméra et des caractéristiques géométriques de notre banc.

Concernant la stéréovision, un nouveau modèle générique permettant de modéliser un couple de caméras avec la projection perspective et les splines cubiques composées a été présenté. Cette modélisation permet d'utiliser les techniques classiques de la mise en correspondance stéréoscopique et de la reconstruction 3D à partir des propriétés des points homologues.

Le premier chapitre a abordé les notions théoriques et a présenté les différentes techniques du calibrage et de la stéréovision proposées dans la littérature. La dernière partie de ce chapitre a exposé les problématiques ainsi que les objectifs spécifiques visés par la présente recherche.

Le deuxième chapitre a été consacré à décrire la nouvelle approche pour le calibrage de caméras. Le banc et le modèle géométrique relatif au processus de saisie d'une image de la grille formée ont été introduits. Dans ce même chapitre, une validation expérimentale de la méthode a été présentée. Les résultats expérimentaux montrent tout d'abord la précision de la méthode avec un RMSE inférieure à 0.2 pixels. Ensuite, en comparant avec les résultats obtenus en utilisant un autre objet de calibrage, l'exactitude des mesures obtenues par la nouvelle technique a été approuvée.

Dans le troisième chapitre, une nouvelle technique pour calibrer une caméra fisheye a été proposée. En effet, le nouveau banc de calibrage impose des cibles réelles situées à l'infini, ce qui donne des points de calibrage précis et invariants par translation (modèle de projection réduit).

Huit images prises en faisant tourner la caméra autour de deux axes sont suffisantes pour calibrer une caméra fisheye. La comparaison avec les paramètres obtenus en utilisant une mire d'étalonnage a confirmé la qualité des résultats. Nous pouvons ainsi considérer que cette méthode constitue une technique adéquate et précise de calibrage pour les caméras panoramiques.

Le quatrième chapitre a présenté le banc de calibrage avec des DOEs plus large qui permettent de générer une grille de 33x33 points et qui sont conçus dans LRIO (Laboratoire de recherche en Ingénierie Optique). Ce banc permet de surmonter les difficultés liées au calibrage des caméras équipées d'un zoom. En effet, il permet de garder un nombre suffisant de cibles vues par la caméra lorsqu'on modifie la position de zoom. Les coordonnées de ces cibles sont mesurées précisément quel que soit la distance focale. Un modèle mathématique a été proposé qui permet de prédire les paramètres de la caméra à partir de dix positions optimisées de zoom. Les RMSEs qui sont obtenues pour l'ensemble des images prises (pour différentes positions de zoom) montrent que notre méthode est robuste.

Un nouveau modèle générique a été introduit dans le dernier chapitre. Il permet de modéliser géométriquement, à l'aide de la projection rectiligne et les splines cubiques composées, les systèmes de stéréovision mixte ou panoramique et de simplifier le calibrage d'une caméra aussi complexe que la lentille panomorphe. Par conséquence, la projection panoramique est convertie en une projection rectiligne et le problème de stéréovision mixte ou panoramique est transformé en un problème de stéréovision classique. La validation expérimentale en utilisant des caméras classiques, fisheye et panomorphe a donné des résultats précis et a confirmé la précision du nouveau modèle.

Ce travail de recherche ouvre de vastes perspectives expérimentales et théoriques. Nous décrivons quelques perspectives par rapport au deux thèmes abordés dans la thèse.

- Calibrage de caméras:

Certains aspects de la méthode proposée peuvent être améliorés. En effet, deux grilles différentes ont été utilisées pour calibrer différentes caméras dans ce travail. Une nouvelle méthode d'optimisation du banc de calibrage pourrait donner les caractéristiques théoriques parfaites des DOEs à utiliser en fonction de la focale, le diamètre de la pupille d'entrée et la taille du pixel de la caméra à calibrer, ainsi que la limite de certain choix.

Une étude de l'influence de la couleur du laser sur la qualité des résultats serait intéressante.

Deux algorithmes ont été proposés pour détecter les points d'intérêts dans l'image. La méthode d'intensité est assez précise mais pourrait être améliorée en utilisant des techniques plus adaptés aux éléments diffractifs comme celle proposée par Vas et al. (2009).

Bien qu'elle soit suffisamment rapide, il serait possible d'accélérer la méthode du calcul lorsqu'on utilise plus qu'une image (cas de fisheye par exemple). En effet, les paramètres extrinsèques initiaux sont estimés actuellement en utilisant les données de chaque images à part alors qu'on pourrait les déduire directement de la position initiale en considérant que la caméra est en rotation pure autour d'un certain axe.

Concernant des perspectives plus globales, la procédure de calibrage est actuellement semi-automatique. En effet, la sélection des points de contrôle est faite manuellement. Afin de tendre vers un protocole de calibrage encore plus souple, nous pouvons adapter la sélection automatique des cibles en fonction de leurs tailles et de leurs intensités.

- Stéréovision :

Le nouveau modèle générique proposé est une percée importante surtout avec la très grande variété de caméras et leurs modèles de projection assez complexe parfois. Des expérimentations seraient à réaliser avec d'autres systèmes stéréoscopiques formés par d'autres types de caméras, comme par exemple les caméras catadioptriques, afin de pouvoir le généraliser pour tous les types de caméra.

Il serait intéressant de pousser l'investigation un peu plus loin en effectuant des expériences de reconstruction tridimensionnelle sur des objets de forme connue.

Un algorithme de conversion permettant de transformer une image panoramique en une image classique pourrait être développé.

Appendix A

A method for determining fundamental matrix

Using a set of corresponding points (p_1^i, p_2^i) , equation 1.10 can be written as:

$$\begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ u_{p_2}^i u_{p_1}^i & u_{p_2}^i v_{p_1}^i & u_{p_2}^i & v_{p_2}^i u_{p_1}^i & v_{p_2}^i v_{p_1}^i & v_{p_2}^i & u_{p_1}^i & v_{p_1}^i & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \begin{pmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{pmatrix} = 0, \quad (\text{A.1})$$

where $\begin{pmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix} = F$.

The equation being of the form $Af = 0$, the initial solution is given using Singular Value Decomposition (SVD) by minimizing the following expression:

$$\|Af\|^2 \quad (\text{A.2})$$

The final solution can be computed by minimizing the following nonlinear criterion [20]:

$$\sum_i d(p_2^i, Fp_1^i)^2 + d(p_1^i, F^T p_2^i)^2 \quad (\text{A.3})$$

where:

- $d(p_2^i, Fp_1^i)^2$, the Euclidean distance from the point to its equivalent epipolar line in the second image.
- $d(p_1^i, F^T p_2^i)^2$, the Euclidean distance from the point to its equivalent epipolar line in the first image.

Appendix B

Resolution of the triangulation

The equation 5.12 being of the form $BP = 0$, the Singular Value Decomposition is used to find the linear solution (Henrichsen, 2000).

The desired solution is the two ray's intersection in the space. However, in reality, the rays do not intersect because of uncertainty in the computing of some calibration parameters and the stereo correspondence. That is why the real position of P is estimated by minimizing the following function (Henrichsen, 2000):

$$f_{trian} = d(p_1, \hat{p}_1)^2 + d(p_2, \hat{p}_2)^2 \quad (\text{B.1})$$

where:

- p_1 and $\hat{p}_1$ are respectively the observed and the computed (using M_1) image coordinates of P in the *Image 1*.
- p_2 and $\hat{p}_2$ are respectively the observed and the computed (using M_2) image coordinates of P in the *Image 2*.
- $d(i, j)$: Euclidean distance between points i and j .

Appendix C

Pseudo codes of calibration and verification

Algorithm 1: Computing the pinhole image (calibration)

**** N control points used ****

For each control point of the input image

Obtain the value of pixel position using centroid algorithm (Thibault et al., 2011)

EndFor

Calculate the initial parameters of the projection matrix M using linear method (Arfaoui, 2009)

Initialize spline cubic parameters to zero.

Optimize the nonlinear criterion (equation 5.8) using Levenberg–Marquardt algorithm to find final spline cubic parameters

Algorithm 2: Transformation of the image

**** Transform distorted image to a true pinhole image ****

For each pixel (u_p, v_p) of output image (pinhole image)

Determine the corresponding location (u_d, v_d) in input image (equation 5.7)

Apply bilinear interpolation to obtain intensity values $I(u_d, v_d)$ at non-integer pixels in the input image

Assign this intensity value to pixel (u_p, v_p)

EndFor

Algorithm 3 : Computing the fundamental (or essential) matrix

For each image of the stereo pair

Execute Algorithm 2

EndFor

**** N_l control points equal or above 8 ****

For each control point of the common FOV of both cameras

Obtain the value of pixel position from pinhole image 1

Obtain the value of pixel position from pinhole image 2

EndFor

Construct homogeneous system $Af=0$ (equation A.1)

**** Calculate F that minimizes the algebraic error (equation A.2) ****

$A \leftarrow UDV^T$ **** singular decomposition of A ****

$F \leftarrow V(9)$ **** F is the last column of V ****

****Find final solution ****

Optimize the nonlinear criterion (equation A.3) using Levenberg–Marquardt algorithm

Algorithm 4: Triangulation

```
**  $N_2$  control points to compute the extrinsic parameters **  
  For each image  
    Optimize the nonlinear criterion (Equation 5.8) to find extrinsic parameters by fixing intrinsic and  
    CCS parameters and using Levenberg–Marquardt algorithm.  
    Construct fundamental matrix.  
  EndFor  
  
  For each target of the common FOV of both cameras  
    Obtain the value of pixel position from pinhole image 1  
    Use fundamental matrix to compute the corresponding epipolar line  
    Obtain the value of pixel position from pinhole image 2 ** It must belongs to epipolar line*  
  EndFor  
  Construct homogeneous system  $BP=0$  (equation 5.12)  
** Estimate initial value of  $P$  **  
   $B \leftarrow UDV^T$    ** singular decomposition of  $B$  **  
   $P \leftarrow V(9)$    **  $X$  is the last column of  $V$  **  
** Compute the final value of  $P$  **  
  Optimize the nonlinear criterion (equation B.1) using Levenberg–Marquardt algorithm
```

Publications

Ce travail de thèse a donné lieu aux publications suivantes :

- THIBAUT, S., ARFAOUI, A., & DÉSAULNIERS, P. (2011). Cross-diffractive optical elements for wide angle geometric camera calibration. *Optics letters*, 36(24), 4770-4772.
- ARFAOUI, A., & THIBAUT, S. (2012). Calibration of aerial and panoramic camera using cross-diffractive optical elements. In *Instrumentation and Measurement Technology Conference (I2MTC), 2012 IEEE International* (pp. 1089-1093). IEEE.
- Arfaoui, A., & Thibault, S. (2013). Fisheye lens calibration using virtual grid. *Applied optics*, 52(12), 2577-2583.
- ARFAOUI, A. THIBAUT, S & DÉSAULNIERS, P. 2015. Calibration of zoom-lens with virtual optical pattern. *Optical engineering*. (Article accepté).

En soumission :

- ARFAOUI, A. & THIBAUT, S. 2015. A new mathematical model for hybrid and panoramic stereovision. *Journal of Electronic Imaging*.

Bibliographie

- ABRAHAM, S. & FÖRSTNER, W. 2005. Fish-eye-stereo calibration and epipolar rectification. *ISPRS Journal of photogrammetry and remote sensing*, 59, 278-288.
- AGAPITO, L., HAYMAN, E. & REID, I. 2001. Self-calibration of rotating and zooming cameras. *International Journal of Computer Vision*, 45, 107-127.
- AHMAD, S., BORECZKY, J. & FOOTE, J. 2006. Automatic video system using multiple cameras. Google Patents.
- AHMED, M. & FARAG, A. 2002. A neural approach to zoom-lens camera calibration from data with outliers. *Image and Vision Computing*, 20, 619-630.
- AHMED, M. T. & FARAG, A. A. A neural optimization framework for zoom lens camera calibration. *Computer Vision and Pattern Recognition*, 2000. Proceedings. IEEE Conference on, 2000. IEEE, 403-409.
- ALDA, J. 2003. Laser and Gaussian beam propagation and transformation. *Encyclopedia of Optical Engineering*, 999-1013.
- ALIAKBARPOUR, H., TAHRI, O. & ARAUJO, H. 2014. Visual servoing of mobile robots using non-central catadioptric cameras. *Robotics and Autonomous Systems*.
- ALVAREZ, L., GÓMEZ, L. & HENRÍQUEZ, P. 2012. Zoom dependent lens distortion mathematical models. *Journal of Mathematical Imaging and Vision*, 44, 480-490.
- AQI, Y., JIANFENG, Y., JIANZHONG, C., QING, Z., ZHI, Z., DENGSHAN, W. & HAO, W. 2011. Optical system design of space fisheye camera. *Acta Optica Sinica*, 31, 1022004.
- ARFAOUI, A. 2009. *Etalonnage de caméras numériques à l'aide de splines cubiques composées*. Université Laval.
- ARFAOUI, A. & PLANTE, F. 2011. Camera calibration using composed cubic splines. *Geomatica*, 65, 189-197.
- ARFAOUI, A. & THIBAUT, S. Calibration of aerial and panoramic camera using cross-diffractive optical elements. *Instrumentation and Measurement Technology Conference (I2MTC)*, 2012 IEEE International, 2012. IEEE, 1089-1093.
- ARFAOUI, A. & THIBAUT, S. 2013. Fisheye lens calibration using virtual grid. *Applied optics*, 52, 2577-2583.
- ARFAOUI, A. & THIBAUT, S. 2015. Mathematical model for hybrid and panoramic stereovision systems: panoramic to rectilinear conversion model. *Applied Optics*, 54, 6534-6542.
- ARFAOUI, A., THIBAUT, S. & DESAULNIERS, P. 2015. Calibration of zoom lens with virtual optical pattern. *Optical Engineering*, 54, 054107-054107.
- ATIENZA, R. & ZELINSKY, A. A practical zoom camera calibration technique: an application on active vision for human-robot interaction. *Proc. Australian Conference on Robotics and Automation*, 2001. 85-90.
- BAKER, S. & NAYAR, S. K. A theory of catadioptric image formation. *Computer Vision*, 1998. Sixth International Conference on, 1998. IEEE, 35-42.
- BAKER, S. & NAYAR, S. K. 1999. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35, 175-196.
- BARRETO, J. P. & ARAUJO, H. 2005. Geometric properties of central catadioptric line images and their application in calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27, 1327-1333.
- BARROIS, B. & WÖHLER, C. 2013. 3D Pose Estimation of Vehicles Using Stereo Camera. *Transportation Technologies for Sustainability*. Springer.
- BASU, A. & LICARDIE, S. Modeling fish-eye lenses. *Intelligent Robots and Systems' 93, IROS'93. Proceedings of the 1993 IEEE/RSJ International Conference on*, 1993. IEEE, 1822-1828.
- BASU, A. & LICARDIE, S. 1995. Alternative models for fish-eye lenses. *Pattern Recognition Letters*, 16, 433-441.
- BAUER, M., GRIEBBACH, D., HERMERSCHMIDT, A., KRÜGER, S., SCHEELE, M. & SCHISCHMANOW, A. 2008. Geometrical camera calibration with diffractive optical elements. *Optics express*, 16, 20241-20248.

- BEARDSLEY, P. A., ZISSERMAN, A. & MURRAY, D. W. 1994. Navigation using affine structure from motion. *Computer Vision—ECCV'94*. Springer.
- BEARDSLEY, P. A., ZISSERMAN, A. & MURRAY, D. W. 1997. Sequential updating of projective and affine structure from motion. *International Journal of Computer Vision*, 23, 235-259.
- BENNEWITZ, M., STACHNISS, C., BURGARD, W. & BEHNKE, S. Metric localization with scale-invariant visual features using a single perspective camera. European Robotics Symposium 2006, 2006. Springer, 195-209.
- BÖRLIN, N. & IGASTO, C. 2009. 3D measurements of buildings and environment for harbor simulators. Technical Report 19, Umeå University, Department of Computing Science.
- CAO, X., XIAO, J., FOROOSH, H. & SHAH, M. 2006. Self-calibration from turn-table sequences in presence of zoom and focus. *Computer Vision and Image Understanding*, 102, 227-237.
- CARON, G. & EYNARD, D. Multiple camera types simultaneous stereo calibration. Robotics and Automation (ICRA), 2011 IEEE International Conference on, 2011. IEEE, 2933-2938.
- CAUCHOIS, C., BRASSART, E., DELAHOCHÉ, L. & DELHOMMELLE, T. Reconstruction with the calibrated cyclop sensor. Intelligent Robots and Systems, 2000.(IROS 2000). Proceedings. 2000 IEEE/RSJ International Conference on, 2000. IEEE, 1493-1498.
- COLLINS, R. T. & TSIN, Y. Calibration of an outdoor active camera system. Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., 1999. IEEE.
- CÓRDOVA-ESPARZA, D.-M., GONZALEZ-BARBOSA, J.-J., HURTADO-RAMOS, J.-B. & ORNELAS-RODRIGUEZ, F.-J. 2014. A panoramic 3D reconstruction system based on the projection of patterns. *Int J Adv Robot Syst*, 11, 55.
- CUCCHIARA, R., PRATI, A. & VEZZANI, R. 2007. A multi-camera vision system for fall detection and alarm generation. *Expert Systems*, 24, 334-345.
- DAMMANN, H. & GÖRTLER, K. 1971. High-efficiency in-line multiple imaging by means of multiple phase holograms. *Optics communications*, 3, 312-315.
- DATTA, A., KIM, J.-S. & KANADE, T. Accurate camera calibration using iterative refinement of control points. Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on, 2009. IEEE, 1201-1208.
- DAVISON, A. J. Real-time simultaneous localisation and mapping with a single camera. Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on, 2003. IEEE, 1403-1410.
- DENKER, K. & UMLAUF, G. 2011. Accurate real-time multi-camera stereo-matching on the gpu for 3d reconstruction.
- DEVERNAY, F. & FAUGERAS, O. 2001. Straight lines have to be straight. *Machine vision and applications*, 13, 14-24.
- DING, Y., YU, J. & STURM, P. Multiperspective stereo matching and volumetric reconstruction. Computer Vision, 2009 IEEE 12th International Conference on, 2009. IEEE, 1827-1834.
- DU, B. & ZHU, H. Estimating fisheye camera parameters using one single image of 3D pattern. Electric Information and Control Engineering (ICEICE), 2011 International Conference on, 2011. IEEE, 367-370.
- DUNNE, A. K., MALLON, J. & WHELAN, P. F. 2010. Efficient generic calibration method for general cameras with single centre of projection. *Computer Vision and Image Understanding*, 114, 220-233.
- DWYER, P. H. & SOUTHAM, G. R. 2004. Dual camera mount for stereo imaging. Google Patents.
- EL RAHMAN SHABAYEK, A., FOFI, D. & MOREL, O. A novel non-central catadioptric auto-calibration approach. Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on, 2009. IEEE, 2228-2235.
- ENCISO, R., VIÉVILLE, T. & FAUGERAS, O. 1993. Approximation du changement de focale et de mise au point par une transformation affine à trois paramètres.
- EYNARD, D., VASSEUR, P., DEMONCEAUX, C. & FRÉMONT, V. UAV altitude estimation by mixed stereoscopic vision. Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, 2010. IEEE, 646-651.

- FAYMAN, J. A., SUDARSKY, O., RIVLIN, E. & RUDZSKY, M. 2001. Zoom tracking and its applications. *Machine Vision and Applications*, 13, 25-37.
- FITZGIBBON, A. W. Simultaneous linear estimation of multiple view geometry and lens distortion. *Computer Vision and Pattern Recognition*, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 2001. IEEE, I-125-I-132 vol. 1.
- FLECK, M. M. 1995. Perspective projection: the wrong imaging model. *Department of Computer Science, University of Iowa*, 1-27.
- FRASER, C. & AL-AJLOUNI, S. 2006. Zoom-dependent camera calibration in digital close-range photogrammetry. *Photogrammetric Engineering & Remote Sensing*, 72, 1017-1026.
- FUJIKI, J., TORII, A. & AKAHO, S. 2007. Epipolar geometry via rectification of spherical images. *Computer Vision/Computer Graphics Collaboration Techniques*. Springer.
- GANDHI, T. & TRIVEDI, M. M. 2006. Reconfigurable omnidirectional camera array calibration with a linear moving object. *Image and Vision Computing*, 24, 935-948.
- GAO, C. & AHUJA, N. Single camera stereo using planar parallel plate. *Pattern Recognition*, 2004. ICPR 2004. Proceedings of the 17th International Conference on, 2004. IEEE, 108-111.
- GEYER, C. & DANIILIDIS, K. 2000. A unifying theory for central panoramic systems and practical implications. *Computer Vision—ECCV 2000*. Springer.
- GLUCKMAN, J. & NAYAR, S. K. A real-time catadioptric stereo system using planar mirrors. *Proceedings of the 1998 DARPA Image Understanding Workshop*, 1998. Citeseer, 309-313.
- GLUCKMAN, J., NAYAR, S. K. & THORESZ, K. J. Real-time omnidirectional and panoramic stereo. *Proc. of Image Understanding Workshop*, 1998. Citeseer, 299-303.
- GRAMMATIKOPOULOS, L., KARRAS, G. & PETSA, E. 2007. An automatic approach for camera calibration from vanishing points. *ISPRS journal of photogrammetry and remote sensing*, 62, 64-76.
- HARTLEY, R. & ZISSERMAN, A. 2003. *Multiple view geometry in computer vision*, Cambridge university press.
- HARTLEY, R. I. & STURM, P. 1997. Triangulation. *Computer vision and image understanding*, 68, 146-157.
- HE, B. & LI, Y. F. 2008. Camera calibration from vanishing points in a vision system. *Optics & Laser Technology*, 40, 555-561.
- HENRICHSSEN, A. 2000. *3 D reconstruction and camera calibration from 2 D images*. University of Cape Town.
- HILL, M. C. & TIEDEMAN, C. R. 2006. *Effective groundwater model calibration: with analysis of data, sensitivities, predictions, and uncertainty*, John Wiley & Sons.
- HUGHES, C., DENNY, P., GLAVIN, M. & JONES, E. 2010a. Equidistant fish-eye calibration and rectification by vanishing point extraction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 32, 2289-2296.
- HUGHES, C., DENNY, P., JONES, E. & GLAVIN, M. 2010b. Accuracy of fish-eye lens models. *Applied optics*, 49, 3338-3347.
- JAIN, R., KASTURI, R. & SCHUNCK, B. G. 1995. *Machine vision*, McGraw-Hill New York.
- JANG, G., KIM, S. & KWEON, I. Single camera catadioptric stereo system. *Proc. of Workshop on Omnidirectional Vision, Camera Networks and Nonclassical cameras (OMNIVIS2005)*, 2005.
- JUN, J. & KIM, C. Robust camera calibration using neural network. *TENCON 99. Proceedings of the IEEE Region 10 Conference*, 1999. IEEE, 694-697.
- KANNALA, J. & BRANDT, S. A generic camera calibration method for fish-eye lenses. *Pattern Recognition*, 2004. ICPR 2004. Proceedings of the 17th International Conference on, 2004. IEEE, 10-13.
- KANNALA, J. & BRANDT, S. S. 2006. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28, 1335-1340.
- KUMLER, J. J. & BAUER, M. L. Fish-eye lens designs and their relative performance. 2000. 360-369.
- LÄBE, T. & FÖRSTNER, W. Geometric stability of low-cost digital consumer cameras. *Proceedings of the 20th ISPRS Congress, Istanbul, Turkey*, 2004. 528-535.

- LAUER, M., SCHÖNBEIN, M., LANGE, S. & WELKER, S. 2011. 3D-objecttracking with a mixed omnidirectional stereo camera system. *Mechatronics*, 21, 390-398.
- LAVEST, J.-M., RIVES, G. & DHOME, M. 1992. Utilisation d'un objectifàfocalevariable en vision monoculaire en vue de la reconstruction 3D. *Revue Traitement du Signal*, 9, 491-506.
- LEITE, D., BERNARDINO, R. & GASPAR, J. 2011. Auto-calibration of pan-tilt-zoom cameras: Estimating intrinsic and radial distortion parameters.
- LHULLIER, M. 2011. A generic error model and its application to automatic 3d modeling of scenes using a catadioptric camera. *International Journal of Computer Vision*, 91, 175-199.
- LI, B., KARPINSKY, N. & ZHANG, S. 2014. Novel calibration method for structured-light system with an out-of-focus projector. *Applied Optics*, 53, 3415-3426.
- LI, H. & HARTLEY, R. 2006. Plane-based calibration and auto-calibration of a fish-eye camera. *Computer Vision-ACCV 2006*. Springer.
- LI, M. & LAVEST, J.-M. 1996. Some aspects of zoom lens camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18, 1105-1110.
- LIAO, H. & CHO, Y. 2008. A new calibration method and its application for the cooperation of wide-angle and pan-tilt-zoom cameras. *Information Technology Journal*, 7, 1096-1105.
- LOEWEN, E. G. & POPOV, E. 1997. *Diffraction gratings and applications*, CRC Press.
- LUONG, Q.-T., DERICHE, R., FAUGERAS, O. & PAPADOPOULOU, T. 1993. On determining the fundamental matrix: Analysis of different methods and experimental results.
- MCBRIDE, J., SNORRASON, M., GOODSSELL, T., EATON, R. & STEVENS, M. R. Single camera stereo for mobile robot surveillance. *Computer Vision and Pattern Recognition-Workshops*, 2005. CVPR Workshops. IEEE Computer Society Conference on, 2005. IEEE, 128-128.
- MEI, C. & RIVES, P. Single view point omnidirectional camera calibration from planar grids. *Robotics and Automation*, 2007 IEEE International Conference on, 2007. IEEE, 3945-3950.
- MICUSIK, B. & PAJDLA, T. Autocalibration & 3D reconstruction with non-central catadioptric cameras. *Computer Vision and Pattern Recognition*, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, 2004. IEEE, I-58-I-65 Vol. 1.
- MICUSIK, B. & PAJDLA, T. 2006. Structure from motion with wide circular field of view cameras. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28, 1135-1149.
- MIRISOLA, L. G., DIAS, J. & DE ALMEIDA, A. T. Trajectory recovery and 3d mapping from rotation-compensated imagery for an airship. *Intelligent Robots and Systems*, 2007. IROS 2007. IEEE/RSJ International Conference on, 2007. IEEE, 1908-1913.
- MOREAU, J., AMBELLOUIS, S. & RUICHEK, Y. 3D reconstruction of urban environments based on fisheye stereovision. *Signal Image Technology and Internet Based Systems (SITIS)*, 2012 Eighth International Conference on, 2012. IEEE, 36-41.
- MOURAGNON, E., LHULLIER, M., DHOME, M., DEKEYSER, F. & SAYD, P. 2009. Generic and real-time structure from motion using local bundle adjustment. *Image and Vision Computing*, 27, 1178-1193.
- NAYAR, S. K. Catadioptric omnidirectional camera. *Computer Vision and Pattern Recognition*, 1997. Proceedings., 1997 IEEE Computer Society Conference on, 1997. IEEE, 482-488.
- NEDEVSCI, S., DANESCU, R., FRENTIU, D., MARITA, T., ONIGA, F., POCOL, C., SCHMIDT, R. & GRAF, T. High accuracy stereo vision system for far distance obstacle detection. *IEEE Intelligent Vehicles Symposium*, 2004. 292-297.
- NISHIMOTO, T. & YAMAGUCHI, J. I. Three dimensional measurement using fisheye stereo vision. *SICE*, 2007 Annual Conference, 2007. IEEE, 2008-2012.
- O'SHEA, D. C., SULESKI, T. J., KATHMAN, A. D. & PRATHER, D. W. 2004. *Diffraction optics: design, fabrication, and test*, Spie Press Bellingham, WA.
- OSKIPER, T., ZHU, Z., SAMARASEKERA, S. & KUMAR, R. Visual odometry system using multiple stereo cameras and inertial measurement unit. *Computer Vision and Pattern Recognition*, 2007. CVPR'07. IEEE Conference on, 2007. IEEE, 1-8.
- PALMER, C. A., LOEWEN, E. G. & THERMO, R. 2005. *Diffraction grating handbook*, Newport Corporation Springfield, OH.

- PARK, S.-W. & HONG, K.-S. 2001. Practical ways to calculate camera lens distortion for real-time camera calibration. *Pattern Recognition*, 34, 1199-1206.
- PELEG, S., BEN-EZRA, M. & PRITCH, Y. 2001. Omnistereo: Panoramic stereo imaging. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23, 279-290.
- PENNISI, A., BLOISI, D., GAZ, C., IOCCHI, L. & NARDI, D. 2013. Novel Patterns and Methods for Zooming Camera Calibration.
- PLANTE, F. 1999. *Rétroingénierie de surfaces complexes continues à l'aide des splines cubiques composées et de la décomposition triangulaire*. thèse de doctorat, Laval.
- PUIG, L., BERMÚDEZ, J., STURM, P. & GUERRERO, J. J. 2012. Calibration of omnidirectional cameras in practice: A comparison of methods. *Computer Vision and Image Understanding*, 116, 120-137.
- RAMALINGAM, S., STURM, P. & LODHA, S. K. Towards complete generic camera calibration. *Computer Vision and Pattern Recognition*, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005. IEEE, 1093-1098.
- REMONDINO, F. & FRASER, C. 2006. Digital camera calibration methods: considerations and comparisons. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36, 266-272.
- RICOLFE-VIALA, C. & SANCHEZ-SALMERON, A.-J. 2010. Lens distortion models evaluation. *Applied optics*, 49, 5914-5928.
- ROBERTI, F., VASSALLO, R. F., TOIBERO, J. M., SORIA, C. & CARELLI, R. 2008. Geometry of a hybrid stereo vision system for robotics applications. *V Jornadas Argentinas de Robotica, Universidad Nacional del Sur, Bahía Blanca, Argentina*, 12-14.
- SABATTINI, L., LEVRATTI, A., VENTURI, F., AMPLO, E., FANTUZZI, C. & SECCHI, C. Experimental comparison of 3d vision sensors for mobile robot localization for industrial application: Stereo-camera and rgb-d sensor. *Control Automation Robotics & Vision (ICARCV)*, 2012 12th International Conference on, 2012. IEEE, 823-828.
- SCHILLER, I., BEDER, C. & KOCH, R. 2008. Calibration of a PMD-camera using a planar calibration pattern together with a multi-camera setup. *The international archives of the photogrammetry, remote sensing and spatial information sciences*, 37, 297-302.
- SCHINDLER, G., KRISHNAMURTHY, P., LUBLINERMAN, R., LIU, Y. & DELLAERT, F. Detecting and matching repeated patterns for automatic geo-tagging in urban environments. *Computer Vision and Pattern Recognition*, 2008. CVPR 2008. IEEE Conference on, 2008. IEEE, 1-7.
- SCHNEIDER, D., SCHWALBE, E. & MAAS, H.-G. 2009. Validation of geometric models for fisheye lenses. *ISPRS Journal of Photogrammetry and Remote Sensing*, 64, 259-266.
- SCHUSTER, K. H. 2011. Catadioptric objective. Google Patents.
- SCHWALBE, E. Geometric modelling and calibration of fisheye lens camera systems. *Proc. 2nd Panoramic Photogrammetry Workshop, Int. Archives of Photogrammetry and Remote Sensing*, 2005. W8.
- SEO, Y. & HONG, K.-S. Auto-Calibration of a Rotating and Zooming Camera. *MVA*, 1998. Citeseer, 274-277.
- SINHA, S. N. & POLLEFEYS, M. 2006. Pan-tilt-zoom camera calibration and high-resolution mosaic generation. *Computer Vision and Image Understanding*, 103, 170-183.
- SMITH, W. J. 1966. *Modern optical engineering*, Tata McGraw-Hill Education.
- STURM, P. Mixing catadioptric and perspective cameras. *Omnidirectional Vision*, 2002. Proceedings. Third Workshop on, 2002. IEEE, 37-44.
- STURM, P. & RAMALINGAM, S. 2003. A generic calibration concept: Theory and algorithms.
- STURM, P. & RAMALINGAM, S. 2004. A generic concept for camera calibration. *Computer Vision-ECCV 2004*. Springer.
- STURM, P. F. & MAYBANK, S. J. On plane-based camera calibration: A general algorithm, singularities, applications. *Computer Vision and Pattern Recognition*, 1999. IEEE Computer Society Conference on., 1999. IEEE.
- SUNDARESAN, A. & CHELLAPPA, R. Markerless motion capture using multiple cameras. *Computer Vision for Interactive and Intelligent Environment*, 2005, 2005. IEEE, 15-26.

- SVOBODA, T. & PAJDLA, T. Panoramic cameras for 3D computation. Proceedings of the Czech Pattern Recognition Workshop, 2000. 63-70.
- SVOBODA, T. & PAJDLA, T. 2002. Epipolar geometry for central catadioptric cameras. *International Journal of Computer Vision*, 49, 23-37.
- SZWOCH, G., DALKA, P., CZYŻEWSKI, A. & FARNAND, S. 2013. Spatial Calibration of a Dual PTZ-Fixed Camera System for Tracking Moving Objects in Video. *Journal of Imaging Science and Technology*, 57, 20507-1-20507-10.
- THIBAUT, S. Enhanced surveillance system based on panomorph panoramic lenses. Defense and Security Symposium, 2007. International Society for Optics and Photonics, 65400E-65400E-8.
- THIBAUT, S., ARFAOUI, A. & DESAULNIERS, P. 2011. Cross-diffractive optical elements for wide angle geometric camera calibration. *Optics letters*, 36, 4770-4772.
- THIBAUT, S. 360° Vision System: Opportunities in Transportation. Proceedings of SPIE, the International Society for Optical Engineering, 2007. Society of Photo-Optical Instrumentation Engineers, 67580H. 1-67580H. 9.
- TSAI, R. Y. 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *Robotics and Automation, IEEE Journal of*, 3, 323-344.
- VYAS, A., ROOPASHREE, M. & PRASAD, B. R. 2009. Centroid detection by Gaussian pattern matching in adaptive optics. *arXiv preprint arXiv:0910.3386*.
- WAN, D. & ZHOU, J. 2008. Stereo vision using two PTZ cameras. *Computer Vision and Image Understanding*, 112, 184-194.
- WAXMAN, A. M. & SINHA, S. S. 1986. Dynamic stereo: passive ranging to moving objects from relative image flows. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 406-412.
- WEICHEL, H. 1990. *Laser beam propagation in the atmosphere*, SPIE press.
- WILLSON, R. G. Modeling and calibration of automated zoom lenses. Photonics for Industrial Applications, 1994. International Society for Optics and Photonics, 170-186.
- WILLSON, R. G. & SHAFER, S. A. Perspective projection camera model for zoom lenses. Optical 3D Measurement Techniques II: Applications in Inspection, Quality Control, and Robotics, 1994. International Society for Optics and Photonics, 149-158.
- WOO, D.-M. & PARK, D.-C. Implicit camera calibration using multilayer perceptron type neural network. Intelligent Information and Database Systems, 2009. ACIIDS 2009. First Asian Conference on, 2009. IEEE, 313-317.
- XIANG, Z., DAI, X. & GONG, X. 2013. Noncentral catadioptric camera calibration using a generalized unified model. *Optics letters*, 38, 1367-1369.
- XU, G. & ZHANG, Z. 1996. *Epipolar geometry in stereo, motion and object recognition: a unified approach*, Springer.
- YING, X. & HU, Z. 2004. Catadioptric camera calibration using geometric invariants. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26, 1260-1271.
- ZHANG, B. L., ZHANG, S. J., DING, W. Q. & SHI, H. S. 2014. Fisheye Lens Distortion Calibration Based on the Lens Characteristic Curves. *Applied Mechanics and Materials*, 519, 634-637.
- ZHANG, H., WONG, K.-Y. & ZHANG, G. 2007. Camera calibration from images of spheres. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29, 499-502.
- ZHANG, L., DU, X., ZHU, Y. & LIU, J. Central catadioptric camera calibration with single image. Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on, 2009. IEEE, 1253-1256.
- ZHANG, Y., ZHAO, L. & HU, W. 2013. A Survey of Catadioptric Omnidirectional Camera Calibration. *International Journal of Information Technology and Computer Science (IJITCS)*, 5, 13.
- ZHANG, Z. 1998. Determining the epipolar geometry and its uncertainty: A review. *International journal of computer vision*, 27, 161-195.
- ZHU, H., YANG, J. & LIU, Z. Fisheye camera calibration with two pairs of vanishing points. Information Technology and Computer Science, 2009. ITCS 2009. International Conference on, 2009. IEEE, 321-324.

