



**UNIVERSITÉ IBN ZOHR
FACULTÉ DES SCIENCES
DÉPARTEMENT D'INFORMATIQUE**

Master Spécialisé
Systèmes Informatiques Distribués & Big Data

Calibrage de la caméra Omnidirectionnelle

Présenté par :

EL-ADARISSI Abdelaziz
EN-NAHEL Aissam

Encadré par :

Professeur ALIBOUCH Brahim

Année universitaire 2022/2023

Remerciements

Nous remercions tout d'abord Monsieur Brahim ALIBOUCH, d'une part, de m'avoir proposé ce sujet riche en informations et sa richesse scientifique théorique et appliquée.

Votre encouragement et vos commentaires nous ont permis d'avancer avec assurance et d'atteindre notre objectif final. Votre détermination à aider les étudiants à réussir a également été une source d'inspiration pour nous tous.

Nous sommes profondément reconnaissants pour l'opportunité de travailler avec vous sur ce projet et pour tout ce que vous avez fait pour nous aider à progresser dans nos études. Nous espérons que notre travail montre notre reconnaissance pour tout ce que vous avez fait pour nous.

Table des matières

| | |
|--|----------|
| Introduction Générale | 1 |
| Étalonnage d'une caméra | 3 |
| 0.1 Introduction | 3 |
| 0.2 Modélisation d'une caméra | 3 |
| Vision omnidirectionnelle | 7 |
| 0.1 Introduction | 7 |
| 0.2 Capteurs omnidirectionnels | 7 |
| 0.2.1 Camera polydioptrique | 7 |
| 0.2.2 Camera fisheye | 8 |
| 0.2.3 Capteurs catadioptriques | 8 |
| 0.2.4 Contrainte du PVU | 11 |
| 0.2.5 Modèle d'équivalence sphérique | 12 |
| 0.3 Le modèle de caméra omnidirectionnelle | 13 |
| 0.3.1 Explication partielle du modèle de caméra omnidirectionnelle | 14 |
| 0.4 Étalonnage de caméra | 17 |
| 0.4.1 Résolution des paramètres intrinsèques et extrinsèques | 17 |
| 0.4.2 Détection du centre de l'image | 19 |
| 0.4.3 Affinement non-linéaire | 21 |

Table des figures

| | | |
|----|--|----|
| 1 | Caméra perspective et image perspective | 2 |
| 2 | Caméra paracatadioptrique et image paracatadioptrique | 2 |
| 1 | Projection d'un point 3D sur une image avec un modèle sténopé de caméra . . . | 3 |
| 2 | La première transformation | 4 |
| 3 | La deuxième transformation | 5 |
| 1 | Caméra Sphericam et Caméra Gear 360 Samsung | 8 |
| 2 | Caméra fisheye et exemple d'image acquise | 8 |
| 3 | Miroirs catadioptriques | 9 |
| 4 | Exemple de caméras centrales : Projection perspective et projection catadiop- trique à travers un miroir hyperboliquer | 10 |
| 5 | Exemple de caméras non centrales | 10 |
| 6 | La géométrie utilisée pour dériver la classe des miroirs satisfaisant la contrainte du point de vue unique (adapté de [Baker 1999]). | 11 |
| 7 | Modèle sphérique unifié | 12 |
| 8 | Système de coordonnées dans le cas catadioptrique | 13 |
| 9 | (b) Plan du capteur, en coordonnées métriques (c) Plan de l'image de la caméra, exprimé en coordonnées de pixels. | 14 |
| 10 | Lorsque la position du centre est correcte, les points 3D du damier se projettent correctement (ronds verts) sur les points de calibration (croix rouges) (a). Au contraire, lorsque la position du centre est incorrecte, les points ne se projettent pas sur les points de calibration réels (b). | 20 |
| 11 | Une image omnidirectionnelle utilisée pour l'étalonnage avec un échiquier utilisé comme motif de calibration. Les points rouges identifient les emplacements de centre candidats pris lors de la première étape de l'algorithme. À chaque étape, les points candidats occupent une région de plus en plus petite autour du point de convergence final. | 21 |

Introduction Générale

Une caméra omnidirectionnelle est un système de vision fournissant une vue panoramique à 360 degrés de la scène. Un champ de vision amélioré peut être obtenu soit en utilisant des systèmes catadioptriques, qui combinent des miroirs et des caméras conventionnelles, soit en utilisant des objectifs fisheye purement dioptriques. Les caméras omnidirectionnelles peuvent être classées en deux classes, centrales et non centrales, en fonction de la propriété de point de vue unique ou non (Baker et Nayar, 1998). Comme indiqué dans (Svoboda et T. Pajdla, 1997), il est hautement souhaitable que de tels systèmes d'imagerie aient un point de vue unique. Lorsque cette propriété est vérifiée, il existe un seul centre de projection, c'est-à-dire que chaque pixel dans les images capturées mesure l'irradiance de la lumière passant par le même point de vue dans une direction particulière. La raison pour laquelle un seul point de vue est si désirable est qu'il permet à l'utilisateur de générer des images de perspective géométriquement correctes à partir des images capturées par une caméra omnidirectionnelle.

La calibration de la caméra omnidirectionnelle est le processus consistant à déterminer les paramètres intrinsèques et extrinsèques d'une caméra capable de capturer des images dans toutes les directions (360 degrés) autour de son axe optique. Ce type de caméra est également appelé caméra panoramique ou sphérique. Les paramètres intrinsèques comprennent la longueur focale, le point principal et les coefficients de distorsion de la caméra, tandis que les paramètres extrinsèques comprennent la position et l'orientation de la caméra par rapport à la scène capturée.

Le but de la calibration d'une caméra omnidirectionnelle est de déterminer les paramètres qui permettent de corriger les distorsions géométriques et radiométriques des images capturées par la caméra. Les distorsions géométriques se produisent en raison des propriétés optiques de la caméra et peuvent inclure des effets tels que la distorsion radiale et tangentielle, qui déforment les formes et les positions des objets dans l'image. Les distorsions radiométriques se produisent en raison des propriétés du capteur d'image de la caméra et peuvent inclure des effets tels que le bruit, la saturation des couleurs, etc.

La calibration d'une caméra omnidirectionnelle peut être un processus complexe en raison de la complexité de la géométrie de la lentille omnidirectionnelle et des distorsions radiales importantes qui en résultent. Cependant, la calibration est essentielle pour obtenir des images précises et cohérentes dans les applications de vision par ordinateur et de traitement d'images.

Introduction Générale

qui utilisent des caméras omnidirectionnelles.

Dans ce travail, nous nous sommes intéressés à différents types de caméras : les caméras perspectives (Fig.1), les caméras catadioptriques à point de vue unique (Fig.2). Les caméras perspectives permettent d'obtenir une photographie de la scène proche de la perception humaine (Fig.1). Les caméras catadioptriques aboutissent à une image à 360° de la scène au prix de fortes distorsions (Fig.2). Dans la première partie on va s'intéresser à la calibration des caméras perspectives et la deuxième partie on passera à la calibration des caméras omnidirectionnelles.

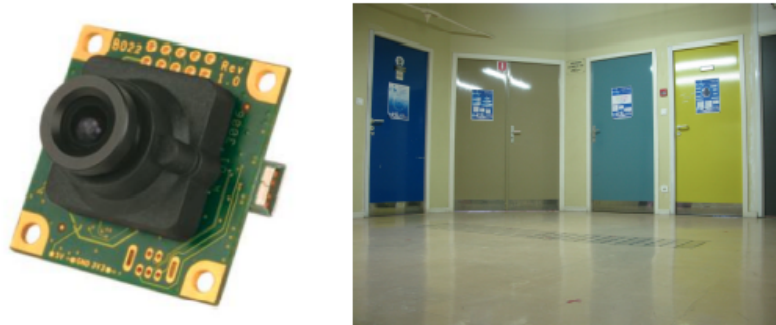


Figure 1 – Caméra perspective et image perspective

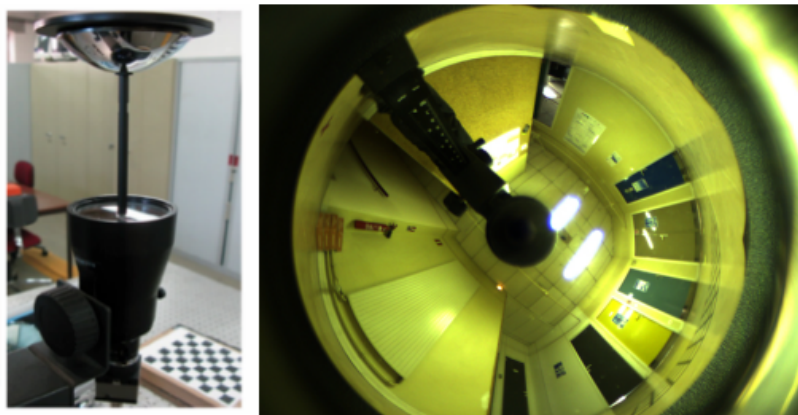


Figure 2 – Caméra paracatadioptrique et image paracatadioptrique

Étalonnage d'une caméra

0.1 Introduction

L'opération de calibration de caméra revient à modéliser le processus de formation des images, c'est-à-dire trouver la relation entre les coordonnées spatiales d'un point de l'espace avec le point associé dans l'image prise par la caméra. Le processus d'estimation des paramètres du modèle de caméra, consistant en un jeu de paramètres intrinsèques et un jeu de paramètres extrinsèques, est connu sous le nom d'étalonnage.

Dans la suite, on présentera tout d'abord le modèle de projection (mathématiquement) décrivant une caméra.

0.2 Modélisation d'une caméra

Le modèle géométrique d'une caméra, représentée en Figure 1, est basé sur trois transformations élémentaires. Leur combinaison forme ce qu'on appelle le modèle sténopé ou modèle pin-hole, il s'agit d'une modélisation simple et linéaire du processus de formation des images au sein d'une caméra.

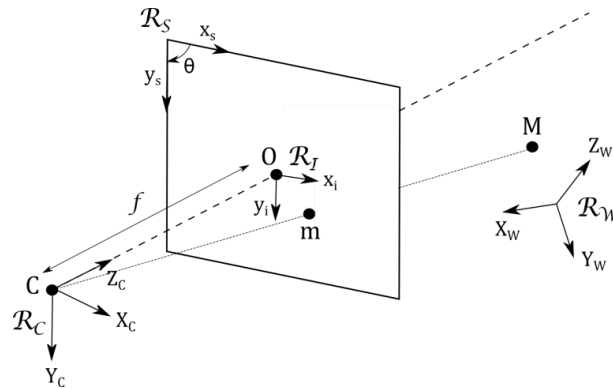


Figure 1 – Projection d'un point 3D sur une image avec un modèle sténopé de caméra

La position d'un point 3D M de la scène, définie dans le repère monde R_W par ses coordonnées

$$\begin{bmatrix} X_w & Y_w & Z_w \end{bmatrix}^T \quad (0.1)$$

Étalonnage d'une caméra

est exprimée dans le repère local R_c attaché à la caméra C avec les coordonnées locales.

$$\begin{bmatrix} X_c & Y_c & Z_c \end{bmatrix}^T \quad (0.2)$$

Ce changement de repère constitue la première transformation et ne dépend que de trois rotations et de trois translations. Cette première transformation entre les coordonnées homogènes monde

$$X_w = \begin{bmatrix} X_w & Y_w & Z_w & 1 \end{bmatrix}^T \quad (0.3)$$

et les coordonnées homogènes locales de caméra

$$X_c = \begin{bmatrix} X_c & Y_c & Z_c & 1 \end{bmatrix}^T \quad (0.4)$$

est représentée par une matrice 4×4 notée T . Elle peut être décomposée en une matrice de rotation R (paramétrable par exemple par trois angles de rotation) et un vecteur de translation t (définie par trois composantes).

$$\begin{Bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{Bmatrix} = [T] \begin{Bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{Bmatrix} = \begin{bmatrix} [R] & (t) \\ \{0_{1 \times 3}\} & 1 \end{bmatrix} \begin{Bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{Bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{Bmatrix}$$

Figure 2 – La première transformation

Ces six paramètres (trois angles, trois translations) sont appelés paramètres extrinsèques et définissent donc le positionnement de la caméra dans l'espace 3D.

La deuxième transformation correspond à la projection du point 3D M sur le plan image auquel est attaché le repère local RI en un point 2D m . Elle fait intervenir uniquement la focale f de la caméra et un facteur d'échelle s selon :

$$s \cdot \begin{Bmatrix} x_i \\ y_i \\ 1 \end{Bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{Bmatrix} X_C \\ Y_C \\ Z_C \\ 1 \end{Bmatrix}$$

Figure 3 – La deuxième transformation

La dernière transformation exprime les coordonnées 2D homogènes du point projeté m dans le repère du capteur R_s .

$$x_s = \begin{bmatrix} x_s & y_s & 1 \end{bmatrix}^T \quad (0.5)$$

La combinaison des deux précédentes transformations peut être décrite à travers une matrice de projection 3×4 K qui s'écrit :

$$K = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (0.6)$$

Cette matrice de projection est donc définie par quatre paramètres intrinsèques à la caméra étalonnée :

- longueur de focale horizontale f_x en pixel ;
- longueur de focale verticale f_y en pixel ;
- la position (c_x, c_y) en pixels du centre optique dans l'image.

La matrice de projection complète P décrivant une caméra est donc composée d'une matrice de paramètres extrinsèques T et d'une matrice de paramètres intrinsèques K , de sorte que :

$$s.X_s = [K][T][X_w]^T = [P][X_w].$$

La matrice de projection P peut ainsi être exprimée comme une matrice 3×4 contenant 12

Étalonnage d'une caméra

termes p_{ij} paramètres indépendants extrinsèques et intrinsèques présentés précédemment. Sous la forme , P s'écrit alors :

$$[P] = \begin{bmatrix} r_{11} \cdot f_x + r_{31} \cdot c_x & r_{12} \cdot f_x + r_{32} \cdot c_x & r_{13} \cdot f_x + r_{33} \cdot c_x & t_x \cdot f_x + t_z \cdot c_x \\ r_{21} \cdot f_y + r_{31} \cdot c_y & r_{22} \cdot f_y + r_{32} \cdot c_y & r_{23} \cdot f_y + r_{33} \cdot c_y & t_y \cdot f_y + t_z \cdot c_y \\ & r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \quad (0.7)$$

Vision omnidirectionnelle

0.1 Introduction

Cette partie est consacré à introduire, dans un premier lieu, la vision omnidirectionnelle et particulièrement de type catadioptrique où nous présentons brièvement les différents systèmes omnidirectionnels utilisés dans la littérature. Ensuite, nous nous penchons sur la vision catadioptrique en détaillant la théorie du Point de Vue Unique visant à définir les systèmes permettant une projection centrale.

0.2 Capteurs omnidirectionnels

Un système de vision omnidirectionnelle fournit une perception de l'environnement avec un champ de vue très large qui peut atteindre 360° autour de son axe. Plusieurs systèmes permettent l'obtention d'une image omnidirectionnelle, cette dernière pourrait être réduite à une bande donnant ce qu'on appelle une image panoramique.

Parmi les premiers systèmes omnidirectionnels, nous trouvons la caméra rotative. Il s'agit d'une caméra conventionnelle qui effectue une rotation autour d'un axe vertical à vitesse angulaire constante. L'image panoramique est construite après l'acquisition de plusieurs images successives couvrant l'environnement observé. Aussi système polydioptrique permet également une vision omnidirectionnelle, c'est un regroupement de caméras pointant vers différentes directions. Un autre système, largement utilisé ces dernières années, s'appuie sur l'utilisation des lentilles grand-angle comme la caméra fisheye et système catadioptrique c'est la combinaison d'une caméra perspective dirigée vers un miroir de révolution réfléchissant.

0.2.1 Camera polydioptrique

On peut noter les principaux inconvénients de la caméra rotative le temps d'acquisition assez long et l'inadaptabilité aux scènes dynamiques. Pour cette raison, le système de caméras multiples a été conçu. Il repose sur l'assemblage des images acquises avec plusieurs caméras, chacune orientée vers une direction et recouvrant un champ de vue restreint. Pour ce type de systèmes, nous donnons l'exemple de deux caméras récentes :



Figure 1 – Caméra Sphericam et Caméra Gear 360 Samsung

0.2.2 Camera fisheye

Afin d'obtenir un champ de vue plus large qu'une caméra perspective, un objectif appelé fisheye (Œil de poisson) a été proposé au début du vingtième siècle. Sa particularité d'avoir une distance focale très courte induit un angle de champ très important de 180° et même 190° . La figure suivante présente un exemple de caméra fisheye récente ainsi que l'image résultante.



Figure 2 – Caméra fisheye et exemple d'image acquise

Un objectif fisheye présente l'avantage de sa simplicité d'utilisation, puisqu'il s'adapte comme toute autre optique sur les caméras. Cependant, il introduit de fortes distorsions radiales. En plus, ce type de capteur ne vérifie pas la contrainte du PVU (Point de Vue Unique), ce qui rend difficile la modélisation de sa projection.

0.2.3 Capteurs catadioptriques

L'association d'une caméra conventionnelle et d'un miroir de révolution permet la construction d'un système de vision omnidirectionnelle : les rayons lumineux provenant de l'environnement 3D sont réfléchis par la surface du miroir et dirigés vers la lentille de la caméra. Ces dispositifs sont référencés dans la littérature comme des systèmes catadioptriques :

- Le mot "cata" renvoie à la catoptrique, c'est la partie de l'optique géométrique qui s'intéresse aux problèmes de réflexion sur des miroirs.
- Le mot "dioptrique" renvoie à la réfraction des lentilles.

Il existe deux types des capteurs catadioptriques, centraux et non-centraux, les caméras catadioptriques centrales sont celles qui possèdent un centre de projection unique. Cette théorie du point de vue unique, publiée par Nayar et Baker signifie que chaque rayon lumineux de la scène passe dans une seule direction à travers ce point de vue ou de façon équivalente, le système associe à chaque point de l'espace 3D un unique point projeté sur l'image.

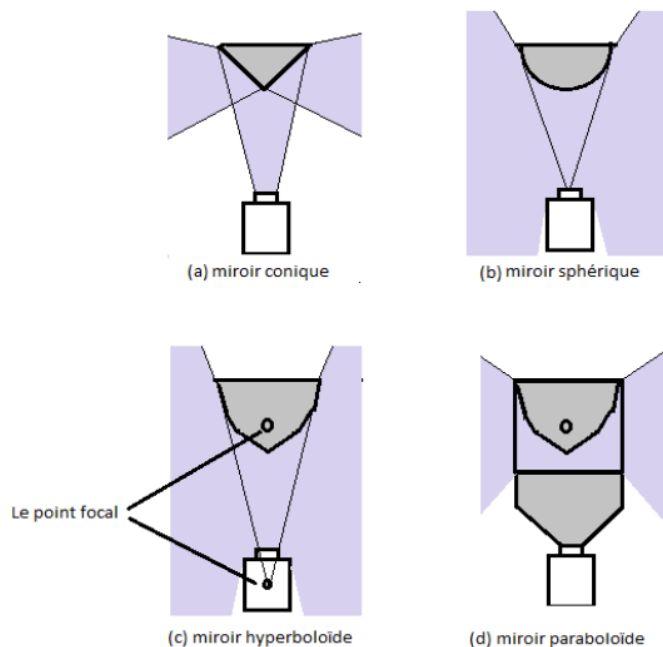


Figure 3 – Miroirs catadioptriques

Capteurs catadioptriques centraux

Les capteurs catadioptriques centraux, combinant des miroirs de révolution avec des caméras conventionnelles, constituent la majorité des systèmes utilisés en vision omnidirectionnelle. La contrainte du point central implique que les droites liant un point de l'espace et sa projection sur le miroir passent par un point unique. Sous cette contrainte, chaque pixel sur le plan image mesure la luminance du rayon réfléchi par le miroir passant par le point central dans une direction particulière et connue.

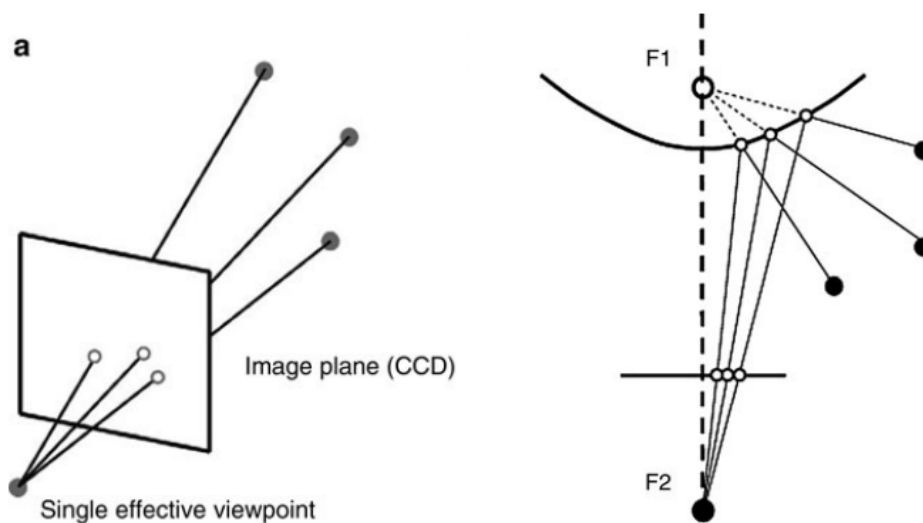


Figure 4 – Exemple de caméras centrales : Projection perspective et projection catadioptrique à travers un miroir hyperbolique

Capteurs catadioptriques non centraux

Les capteurs catadioptriques non centraux ont, par définition, plus d'un centre de projection (plusieurs points de vue). L'ensemble de ces points de vue définit un lieu particulier de l'espace, délimité par une surface 3D appelée surface caustique. D'un point de vue géométrique, cette surface est formée par l'ensemble des points d'intersection des rayons réfléchis par le miroir.

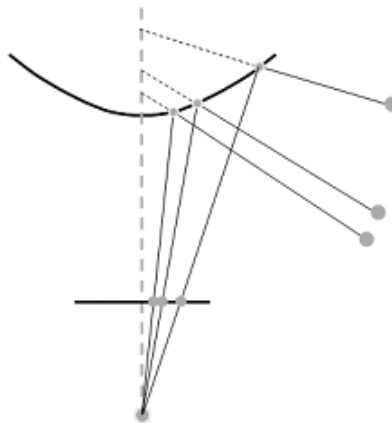


Figure 5 – Exemple de caméras non centrales

0.2.4 Contrainte du PVU

La contrainte du point de vue unique (PVU) implique la convergence des rayons vers un point unique, appelé le centre optique du capteur. La caméra perspective est un exemple de système central à PVU. Cette propriété est intéressante et souhaitable pour les systèmes catadioptriques parce que chaque pixel dans l'image mesure l'intensité de lumière (l'irradiance) provenant d'une unique direction. Par conséquent, la modélisation du processus de projection devient plus simple. Ainsi, on a la possibilité d'une construction, géométriquement correcte, d'une image panoramique ou planaire perspective à partir d'une image catadioptrique.

Dans [Baker 1999], Baker et Nayar ont conclu les formes des miroirs à combiner avec la caméra conventionnelle permettant d'assurer cette contrainte, qui sont : le miroir plan, le miroir elliptique, le miroir parabolique et le miroir hyperbolique. Ils ont mis au point les équations définissant la classe géométrique complète ces miroirs.

$$(z - \frac{c}{2})^2 - r^2(\frac{k}{2} - 1) = \frac{c^2}{4}(\frac{k-2}{k}) \quad (k \geq 2) \quad (0.1)$$

$$(z - \frac{c}{2})^2 + r^2(1 + \frac{c^2}{2k}) = (\frac{2k + c^2}{4}) \quad (k > 0) \quad (0.2)$$

Où : $r = \sqrt{x^2 + y^2}$, k une constante strictement positive et c est la distance entre le point de vue et le centre optique de la caméra Fig.6.

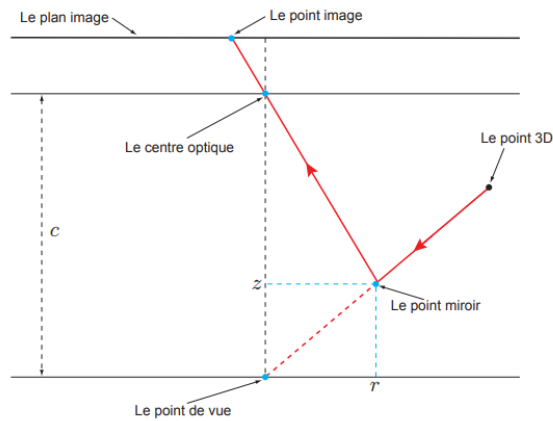


Figure 6 – La géométrie utilisée pour dériver la classe des miroirs satisfaisant la contrainte du point de vue unique (adapté de [Baker 1999]).

0.2.5 Modèle d'équivalence sphérique

Le modèle de projection sphérique [Geyer and Daniilidis, 2000] unifie tout type de caméras conventionnelles ou systèmes catadioptriques centraux respectant la contrainte du PVU [Baker and Nayar, 1999]. Ce modèle est illustré sur la Fig. 7 et peut être décrit en deux étapes :

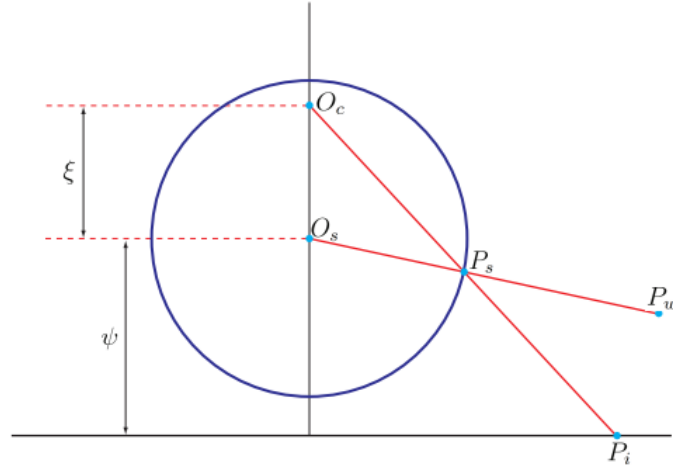


Figure 7 – Modèle sphérique unifié

- 1- Le point 3D $P_w(X_w, Y_w, Z_w)$ est projeté sur la sphère unitaire en un point $P_s(X_s, Y_s, Z_s)$ obtenu de l'intersection de la sphère et la droite joignant son centre O_w et le point P_w .
- 2- Le point P_s est ensuite projeté sur le plan image, en utilisant une projection perspective de centre O_c , donnant ainsi l'image $P_i(x, y)$ de point 3D P_w sur l'image omnidirectionnelle.

Ce modèle est applicable à tous les systèmes catadioptriques centraux qui sont définis par les paramètres ζ et ψ qui sont respectivement la distance entre le centre de projection et le centre de la sphère, et la distance entre le plan image et le centre de la sphère. Le tableau ci-dessous montre les paramètres ζ et ψ correspondant à chaque type de capteur catadioptrique central en fonction des paramètres de son miroir. $4p$ et c sont respectivement le latus rectum et la distance entre les deux foyers du miroir.

| Captuer | Parabolique | Hyperbolique | Planaire |
|---------|-------------|------------------------------------|----------|
| ζ | 1 | $\frac{c}{\sqrt{c^2 + 4p^2}}$ | 0 |
| ψ | $1 + 2p$ | $\frac{c + 2p}{\sqrt{c^2 + 4p^2}}$ | 1 |

Scaramuzza et al. ont aussi proposé un modèle unifié pour les systèmes catadioptriques centraux et les caméras fisheye. La fonction de projection est approximée par un développement

en série de Taylor, et les coefficients sont estimés par un ajustement linéaire par la méthode des moindres carrés, ce qui facilite d'estimer les paramètres intrinsèques et extrinsèques (calibrage). Dans le chapitre suivante on va détailler ce modèle proposé Scaramuzza et al.

0.3 Le modèle de caméra omnidirectionnelle

Étalonnage d'une caméra omnidirectionnelle implique de trouver la relation entre un point de pixel 2D donné q' et le vecteur 3D X émanant du point de vue effectif du miroir :

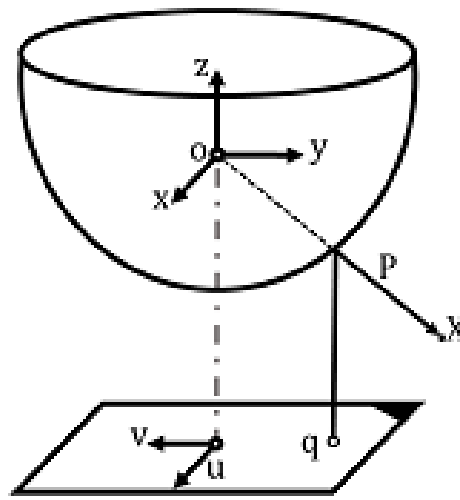


Figure 8 – Système de coordonnées dans le cas catadioptrique

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} u \\ v \\ f(u, v) \end{bmatrix}$$

À ce stade, nous pouvons introduire la fonction d'imagerie g , qui capture la relation entre un point q , dans le plan du capteur, et le vecteur P émanant du point de vue O vers un point de la scène X .

Ce modèle de caméra omnidirectionnelle traite le système d'imagerie comme un système compact unique ; c'est-à-dire qu'il n'a pas d'importance si vous utilisez un miroir ou une lentille fisheye en combinaison avec votre caméra.

Ce modèle repose sur les 4 hypothèses suivantes :

- 1- Le système de caméra miroir est un système central, donc, il existe un point dans le miroir où chaque rayon réfléchi se croise. Ce point est considéré comme l'origine des axes du système de coordonnées XYZ de la caméra.
- 2- Les axes de la caméra et du miroir sont bien alignés, c'est-à-dire que seules de légères déviations de la rotation sont prises en compte dans ce modèle.
- 3- Le miroir est symétrique de rotation par rapport à son axe.
- 4- La distorsion de l'objectif de la caméra n'est pas prise en compte dans ce modèle. La distorsion de l'objectif de la caméra n'a pas été incluse car les caméras omnidirectionnelles utilisant des miroirs ont généralement besoin d'une grande distance focale pour focaliser l'image sur le miroir. Ainsi, la distorsion de l'objectif peut être négligée. Si vous utilisez des lentilles fisheye, la distorsion de l'objectif de la caméra est déjà intégrée dans la fonction de projection f .

0.3.1 Explication partielle du modèle de caméra omnidirectionnelle

Pour l'instant, supposons que l'hypothèse 2 soit parfaitement vérifiée, c'est-à-dire que les axes de la caméra et du miroir sont parfaitement alignés. Plus tard, nous verrons comment surmonter également cette contrainte.

Soit q' un point de pixel de l'image et $[u', v']^T$ ses coordonnées de pixel par rapport au centre de l'image omnidirectionnelle :

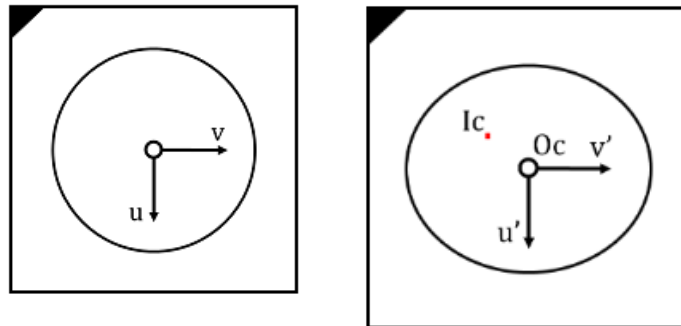


Figure 9 – (b) Plan du capteur, en coordonnées métriques (c) Plan de l'image de la caméra, exprimé en coordonnées de pixels.

Soit X (un point de la scène) son vecteur 3D correspondant émanant d'un unique point de vue effectif, et (x, y, z) ses coordonnées par rapport à l'origine des axes. On suppose que $q = [u, v]^T$ est la projection de X sur le plan du capteur. les deux systèmes (b) et (c) sont liés par une transformation affine qui intègre le processus de numérisation et de petits désalignements d'axes, ainsi $q = Aq' + t$, où $A \in \mathbb{R}^{2 \times 2}$ et $t \in \mathbb{R}^{2 \times 1}$.

Ainsi, la relation entre un point de pixel q' et un point de scène X est :

$$\lambda.p = \lambda.g(q) = \lambda.g(Aq' + t) = P.X \quad \lambda > 0 \quad (1)$$

où $X \in \mathbb{R}^4$ est exprimé en coordonnées homogènes et $P \in \mathbb{R}^{3 \times 4}$ est la matrice de projection perspective $P=K[R \ T]$. Par étalonnage de la caméra omnidirectionnelle, on entend l'estimation des matrices A et t et de la fonction non linéaire g , de sorte que tous les vecteurs $g(Aq'+t)$ satisfont l'équation de projection (1). Nous supposons pour g l'expression suivante :

$$g(u, v) = \begin{bmatrix} u \\ v \\ f(u, v) \end{bmatrix} \quad (2)$$

De plus, nous supposons que la fonction f dépend de u et v uniquement à travers de $\rho = \sqrt{v^2 + u^2}$. Cette hypothèse correspond à supposer que la fonction g est symétrique par rotation par rapport à l'axe du capteur.

La fonction f peut avoir différentes formes en fonction de la construction du miroir ou de la lentille. Contrairement à l'utilisation d'un modèle spécifique pour le capteur en cours d'utilisation, nous choisissons d'appliquer un modèle paramétrique généralisé de f , qui convient à différents types de capteurs. La raison en est que nous voulons que ce modèle compense tout désalignement entre le point focal du miroir (ou de la lentille fisheye) et le centre optique de la caméra.

De plus, nous souhaitons que notre fonction généralisée tienne approximativement avec ces capteurs pour lesquels la propriété de point de vue unique n'est pas exactement vérifiée (par exemple, les caméras fisheye génériques).

Nous proposons la forme polynomiale suivante pour f :

$$f(\rho) = a_0 + a_1\rho + a_2\rho^2 + a_3\rho^3 + \dots + a_n\rho^n \quad (3)$$

En fait, plus on augmente l'ordre du polynôme, plus la précision de la calibration augmente. Ce n'est pas vrai pour les polynômes d'ordre élevé. Les polynômes de quatrième ordre donnent les meilleurs résultats de calibration.

Les coefficients a_i , $i = 0, 1, 2, \dots, n$ et le degré polynomial n sont les paramètres de calibration que nous voulons déterminer.

Cette description polynomiale de f peut être simplifiée en considérant que toutes les définitions précédentes de f satisfont toujours ce qui suit :

$$\frac{\partial f}{\partial \rho}_{\rho=0} = 0$$

Cette propriété est vraie pour les miroirs hyperboliques et paraboliques ou les caméras fisheye. Cette simplification nous permet de supposer que $a_1 = 0$, et donc (3) peut être réécrit comme suit :

$$f(u, v) = f(\rho) = a_0 + a_2\rho^2 + a_3\rho^3 + \dots + a_n\rho^n \quad (5)$$

En conséquence, nous avons réduit le nombre de paramètres à estimer. En résumé, l'équation (1) peut être réécrite comme suit :

$$\lambda \cdot \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \lambda \cdot g(A \cdot q' + t) = \lambda \cdot \begin{bmatrix} A \cdot q' + t \\ f(u, v) \end{bmatrix} = P \cdot X \quad \lambda > 0 \quad (6)$$

0.4 Étalonnage de caméra

0.4.1 Résolution des paramètres intrinsèques et extrinsèques

Estimation des paramètres intrinsèques : Les paramètres intrinsèques de la caméra comprennent la distance focale, le point principal et les coefficients de distorsion de l'objectif. Pour estimer ces paramètres, nous pouvons utiliser un algorithme de calibration de caméra comme la méthode de Zhang ou la méthode de Bouguet. Ces algorithmes utilisent les points caractéristiques pour estimer les paramètres intrinsèques de la caméra.

Estimation des paramètres extrinsèques : Les paramètres extrinsèques de la caméra comprennent sa position et son orientation par rapport au motif de calibration. Pour estimer ces paramètres, nous pouvons utiliser un algorithme d'estimation de pose comme PnP (Perspective-n-Point) ou EPnP (Efficient Perspective-n-Point). Ces algorithmes utilisent les paramètres intrinsèques, et les points caractéristiques pour estimer la position et l'orientation de la caméra.

Selon ce que nous avons dit jusqu'à présent, pour calibrer une caméra omnidirectionnelle, nous devons estimer les paramètres A, t, a_0, a_2, \dots et a_n .

Dans notre approche, nous avons décidé de séparer l'estimation de ces paramètres en deux étapes. Dans le premier, nous estimons les paramètres affines A et t .

Dans l'autre, nous estimons les coefficients a_0, a_2, \dots et a_n .

Les paramètres A et t décrivent la transformation affine qui relie le plan du capteur au plan de la caméra (figures 9.b et 9.c). A est la matrice d'étirement et t est le vecteur de translation (figure 9.c).

Pour estimer A et t , nous introduisons une méthode qui, contrairement aux travaux précédents, ne nécessite pas la visibilité de la limite externe circulaire du miroir (esquissée par l'ellipse dans la figure 9.c).

Cette méthode repose sur une procédure itérative qui commence en fixant A à la matrice identité Eye et t à 0. Cette hypothèse signifie que le plan de la caméra et le plan du capteur coïncident initialement. Les éléments corrects de A seront ensuite estimés par affinement non linéaire, tandis que t sera estimé par un algorithme de recherche itérative. Cette approche sera détaillée dans la section 0.4.3. À partir de maintenant, nous supposons que $A = Eye$ et $t = 0$, ce qui signifie que $u'' = u'$. Ainsi, en substituant cette relation dans (6) et en utilisant (5), nous

obtenons l'équation de projection suivante :

$$\lambda \cdot \begin{bmatrix} u'' \\ v'' \\ w'' \end{bmatrix} = \lambda \cdot g(u') = \lambda \cdot \begin{bmatrix} u' \\ v' \\ a_0 + a_2 \rho'^2 + \dots + a_n \rho'^n \end{bmatrix} = P \cdot X \quad \lambda > 0 \quad (7)$$

Où u' et v' sont les coordonnées en pixels d'un point d'image par rapport au centre de l'image, et ρ' est la distance euclidienne. De plus, il faut maintenant estimer seulement n paramètres (a_0, a_2, \dots, a_n). Dès maintenant, nous ferons référence à ces paramètres comme étant les paramètres intrinsèques.

Pendant la procédure de calibration, un motif plan de géométrie connue est montré à différentes positions inconnues, qui sont liées au système de coordonnées du capteur par une matrice de rotation $\mathbf{R} \in \mathbf{R}^{3 \times 3}$ et une translation $\mathbf{T} \in \mathbf{R}^{3 \times 1}$. R et T seront appelés paramètres extrinsèques.

Soit I_i une image observée du motif de calibration, $M_{ij} = \begin{bmatrix} X_{ij} \\ Y_{ij} \\ Z_{ij} \end{bmatrix}$ les coordonnées 3D de ses points dans le système de coordonnées du motif, et $m_{ij} = \begin{bmatrix} u_{ij} \\ v_{ij} \end{bmatrix}$ les coordonnées de pixels correspondantes dans le plan de l'image. Comme nous avons supposé que le motif soit plan, sans perte de généralité, nous avons $Z_{ij} = 0$. Ensuite, l'équation (7) devient :

$$\lambda_{ij} \cdot p_{ij} = \lambda_{ij} \cdot \begin{bmatrix} u_{ij} \\ v_{ij} \\ a_0 + a_2 \rho'^2 + \dots + a_n \rho'^n \end{bmatrix} = P^i \cdot X = [r_1^i \cdot r_2^i \cdot r_3^i T^i] \cdot \begin{bmatrix} X_{ij} \\ Y_{ij} \\ 0 \\ 1 \end{bmatrix} = [r_1^i \cdot r_2^i \cdot T^i] \cdot \begin{bmatrix} X_{ij} \\ Y_{ij} \\ 1 \end{bmatrix} \quad (8)$$

Où r_1, r_2 et r_3 sont les vecteurs colonnes de \mathbf{R} . Par conséquent, afin de résoudre la calibration de la caméra, les paramètres extrinsèques doivent également être déterminés pour chaque pose du motif de calibration. En observant l'équation (8), nous pouvons éliminer la dépendance de l'échelle de profondeur λ_{ij} en multipliant les deux côtés de l'équation par le vecteur p_{ij} .

Cela implique que chaque point p_j contribue à trois équations non linéaires homogènes

$$\begin{cases} v_j.(r_{31}Xj + r_{32}Yj + t_3) - f(\rho_j)(r_{21}Xi + r_{22}Yj + t_2) = 0 & (9.1) \\ f(\rho_j).(r_{11}Xi + r_{12}Yj + t_1) - u_j.(r_{31}Xj + r_{32}Yj + t_3) = 0 & (9.2) \\ u_j.(r_{21}Xj + r_{22}Yj + t_2) - v_j.(r_{11}Xj + r_{12}Yj + t_1) = 0 & (9.3) \end{cases}$$

Où le sous-index i a été supprimé pour alléger la notation, et t_1, t_2 et t_3 sont les éléments de T . Remarquez que dans (9), X_j, Y_j , et Z_j sont connus, ainsi que u_j, v_j . De plus, observez que seulement (9.3) est linéaire dans les inconnues $r_{11}, r_{12}, r_{21}, r_{22}, t_1, t_2$.

Dès maintenant, les détails pour la résolution de l'équation (9) peuvent être trouvés dans (Scaramuzza et al., 2006). Le principe de la technique consiste d'abord à résoudre les paramètres $r_{11}, r_{12}, r_{21}, r_{22}, t_1$ en résolvant linéairement l'équation (9.3).

Ensuite, nous utilisons la solution de (9.3) comme entrée pour (9.1) et (9.2), et résolvons pour les paramètres restants a_0, a_2, \dots, a_n et t_3 . Dans les deux étapes, la solution est obtenue en utilisant la minimisation linéaire des moindres carrés.

Jusqu'à présent, nous n'avons pas spécifié le degré polynomial n à utiliser. Pour calculer le meilleur n , nous commençons en fait avec $n = 2$. Ensuite, nous augmentons n par des étapes unitaires et nous calculons la valeur moyenne de l'erreur de reprojection de tous les points de calibration. La procédure s'arrête lorsqu'une erreur minimale est trouvée. Les valeurs empiriques typiques pour n sont généralement $n = 3$ ou $n = 4$.

0.4.2 Détection du centre de l'image

Comme indiqué dans les sous sections précédant, une particularité de notre boîte à outils de calibration est qu'elle nécessite un minimum d'interaction utilisateur. L'un des outils qui accomplit cette tâche est sa capacité à identifier le centre de l'image omnidirectionnelle O_o (figure 9.b), même lorsque la limite externe du capteur n'est pas visible dans l'image.

Au début, nous avons fait les hypothèses suivantes pour A et t , à savoir $A = Eye$ et $t = 0$. Ensuite, nous avons dérivé les équations pour résoudre les paramètres intrinsèques et extrinsèques qui ne sont valables que sous ces hypothèses.

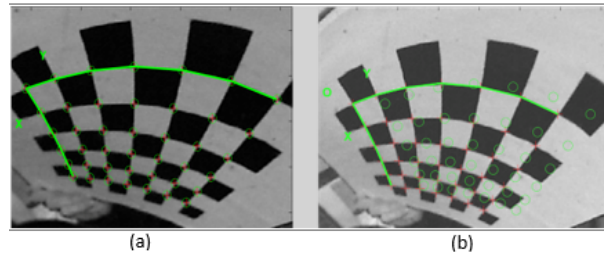


Figure 10 – Lorsque la position du centre est correcte, les points 3D du damier se projettent correctement (ronds verts) sur les points de calibration (croix rouges) (a). Au contraire, lorsque la position du centre est incorrecte, les points ne se projettent pas sur les points de calibration réels (b).

Dans la figure 10.a, le lecteur peut voir ce qui se passe lorsque la position du centre est correcte. Les croix rouges sont les points de calibration d'entrée sélectionnés par l'utilisateur.

Les ronds verts sont les points 3D reprojectés sur les images selon les paramètres intrinsèques et extrinsèques estimés par la calibration. Comme le lecteur peut le constater, les points 3D recouvrent parfaitement les points d'entrée, ce qui signifie que la calibration a fonctionné correctement.

La figure 10.b montre le résultat lorsque la position d'entrée du centre est incorrecte, c'est-à-dire que l'erreur de reprojction est grande.

Motivés par cette observation, nous avons effectué de nombreux essais de notre procédure de calibration pour différentes positions de centre, et, pour chaque essai, nous avons calculé la somme des erreurs de reprojction au carré (SSRE).

En conséquence, nous avons vérifié que le SSRE a toujours un minimum global à la position correcte du centre. Ce résultat nous conduit à une recherche exhaustive du centre O_c , qui s'arrête lorsque la différence entre deux emplacements potentiels de centre est inférieure à un certain (nous avons utilisé $= 0,5$ pixels). L'algorithme est le suivant :

- 1- À chaque étape de cette recherche itérative, un nombre fixe d'emplacements de centre candidats est sélectionné de manière uniforme à partir d'une région d'image donnée (voir la figure 3).
- 2- Pour chacun de ces points, la calibration est effectuée en utilisant ce point comme emplacement de centre potentiel et le SSRE est calculé.
- 3- Le point fournissant le SSRE minimum est pris comme centre potentiel.

- 4- La recherche se poursuit en sélectionnant d'autres emplacements candidats dans la région autour de ce point, et les étapes 1, 2 et 3 sont répétées jusqu'à ce que la condition d'arrêt soit satisfaite.

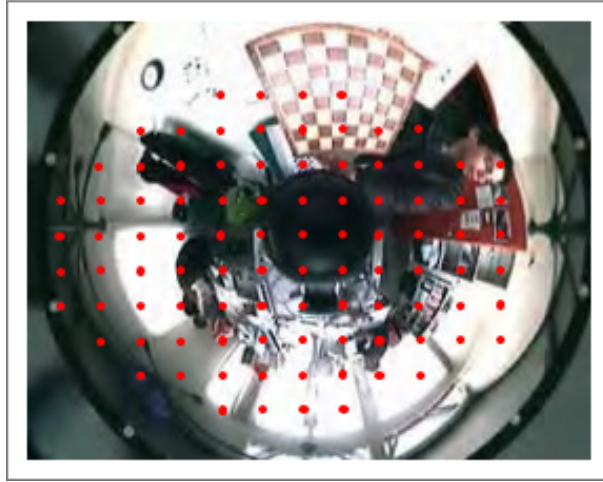


Figure 11 – Une image omnidirectionnelle utilisée pour l'étalonnage avec un échiquier utilisé comme motif de calibration. Les points rouges identifient les emplacements de centre candidats pris lors de la première étape de l'algorithme. À chaque étape, les points candidats occupent une région de plus en plus petite autour du point de convergence final.

Remarquez que le coût de calcul de cette recherche itérative est si faible qu'elle prend moins de 3 secondes pour s'arrêter. À ce stade, le lecteur pourrait se demander comment nous estimons les éléments de la matrice A . En fait, au début, nous avons supposé que $A = Eye$. L'algorithme itératif mentionné ci-dessus recherche de manière exhaustive l'emplacement du centre (à savoir O_c) en laissant A inchangé.

La raison en est que l'excentricité de la frontière externe d'une image omnidirectionnelle est généralement proche de zéro, ce qui signifie que $A \approx Eye$.

Par conséquent, nous avons choisi d'estimer A dans une deuxième étape en utilisant une méthode de minimisation non linéaire, qui est décrite dans suivante.

0.4.3 Affinement non-linéaire

La solution linéaire présentée dans la section (Résolution des paramètres intrinsèques et extrinsèques) est obtenue en minimisant une distance algébrique qui n'a pas de signification physique. Pour remédier à cela, nous avons choisi d'affiner les paramètres de calibration par inférence de maximum de vraisemblance.

Supposons que nous disposons de K images d'un plan modèle, chacune contenant L points de coin. Ensuite, supposons que les points d'image sont corrompus par un bruit indépendant et identiquement distribué.

L'estimation de maximum de vraisemblance peut alors être obtenue en minimisant la fonctionnelle suivante :

$$E = \sum_{i=1}^K \sum_{j=1}^L \|m_{ij} - m(R_i, T_i, A, O_c, a_0, a_2, \dots, a_n, M_j)\|^2 \quad (10)$$

Où $m(R_i, T_i, A, O_c, a_0, a_2, \dots, a_n, M_j)$ est la reprojection du point M_j du plan i selon l'équation (1). R_i et T_i sont les matrices de rotation et de translation de chaque pose de plan. R_i est paramétré par un vecteur de 3 paramètres liés à R_i par la formule de Rodrigues. Remarquez que nous incorporons maintenant dans la fonctionnelle à la fois la matrice d'étirement A et le centre de l'image omnidirectionnelle O_c .

En minimisant la fonctionnelle définie en (10), nous trouvons en fait les paramètres de calibration qui minimisent l'erreur de reprojection. Afin d'accélérer la convergence, nous avons décidé de diviser la minimisation non linéaire en deux étapes.

La première affine les paramètres extrinsèques, en ignorant les intrinsèques. Ensuite, la deuxième étape utilise les paramètres extrinsèques récemment estimés et affine les intrinsèques. En effectuant de nombreuses simulations, nous avons constaté que cette division n'affecte pas le résultat final par rapport à une minimisation globale.

Pour minimiser (10), nous avons utilisé l'algorithme de Levenberg-Marquadt (Levenberg, 1944 ; Marquardt, 1963), tel qu'implémenté dans la fonction Matlab **lsqnonlin**.

L'algorithme nécessite une estimation initiale des paramètres. Ces paramètres initiaux sont ceux obtenus en utilisant la technique linéaire décrite dans la sous-section (Résolution des paramètres intrinsèques et extrinsèques). Comme première estimation pour A , nous avons utilisé la matrice identité, tandis que pour O_c , nous avons utilisé la position estimée par la procédure itérative expliquée dans la sous-section (Détection du centre de l'image).

Conclusion

En conclusion, la calibration de la caméra est une étape importante pour permettre aux systèmes de vision par ordinateur et de robotique de comprendre et d'analyser les images. Une calibration précise peut améliorer considérablement la précision et la fiabilité des systèmes de perception. Il existe plusieurs méthodes pour effectuer la calibration de la caméra, notamment la calibration à partir de motifs, la calibration à partir de points de référence et la calibration autonome. La méthode appropriée dépend des exigences spécifiques de l'application et des ressources disponibles.