

Étalonnage de caméra à partir d'images associées de sphères et de grilles

Camera Calibration from Spheres and Grid Images

par N. DAUCHER, M. DHOME, J.T. LAPRESTE

LASMEA
(URA CNRS n° 1793)
Université Blaise Pascal de Clermont-Ferrand
F-63177 Aubière Cedex

Résumé

A partir d'images de sphères, nous avons développé une nouvelle méthode d'étalonnage de caméra permettant de déterminer ses paramètres intrinsèques. Nous mettons en évidence une propriété géométrique intéressante concernant les primitives images utilisées. La prise en compte de la distorsion géométrique du système optique et une recherche au niveau de la détection des points de contour associés à une image de sphère nous ont permis d'arriver à de bons résultats.

Mots clés : Étalonnage, paramètres intrinsèques, sphère, ellipse, distorsion, détection de points de contour.

Abstract

From spheres images we have developed a new method for camera calibration in order to calculate its intrinsic parameters. We prove an interesting geometric propriety about ellipses extracted from images. Using the lens geometrical distortion introduced by the optical system and searching a precise points detection for spheres images, allow us to obtain satisfactory results.

Key words : Calibration, intrinsic parameters, sphere, ellipse, distortion, contour points extraction.

1. Introduction

Avant de débiter l'analyse d'images obtenues à l'aide d'une caméra liée à un système de digitalisation, il est nécessaire, si l'on envisage d'utiliser le système d'acquisition comme instrument de métrologie, d'en déterminer les caractéristiques.

L'étalonnage des systèmes d'acquisition d'images reste un problème crucial, sur lequel ont déjà porté divers travaux de recherche.

Traditionnellement, l'étalonnage distingue deux types de paramètres :

- intrinsèques : propres à la géométrie de la caméra (position de l'intersection de l'axe optique et du plan image, pas d'échantillonnage en x et y du système d'acquisition et distance focale).
- extrinsèques : quantifiant la position relative entre la "mire" d'étalonnage et la caméra.

Une première approche consiste à globaliser la détermination des paramètres extrinsèques et intrinsèques de la caméra, à

partir de la mise en correspondance d'un ensemble de points appartenant à une mire avec leur projection sur le plan image [14][6][9]. Une seconde consiste à découpler la recherche des paramètres extrinsèques et intrinsèques en exploitant des propriétés géométriques particulières, comme, par exemple, celle des points de fuite de faisceaux orthogonaux de droites parallèles de l'espace [15][9].

En général ces méthodes, contrairement à celles des photogrammètres, ne prennent pas en compte les phénomènes de distorsions optiques liés au capteur, et sont extrêmement sensibles à la précision de détection des primitives-images correspondant à la projection des éléments de la mire exploités par le processus d'étalonnage.

Pour notre part, nous avons conservé la deuxième approche en cherchant à minimiser ses travers. La mire d'étalonnage consiste en une simple sphère. La détermination des paramètres intrinsèques s'effectue à partir de plusieurs images de cette même sphère placée en diverses positions, les primitives choisies étant les ellipses provenant de la projetée des limbes de la sphère.

Notre méthode est basée sur l'exploitation d'une propriété géométrique liée à ce type de primitive (Cf. 2.2.1). Le choix de telles images est étayé par les considérations suivantes :

- le nombre important de points de contour appartenant à une primitive permet de déterminer l'équation de celle-ci avec une très bonne précision.
- les points des limbes de la sphère constituant le contour de l'image, ont l'avantage d'être situés à égale distance du centre optique et jouissent tous de la même netteté pour un réglage donné.
- enfin, ce type de primitive permet de s'affranchir en partie des problèmes liés à la distorsion, si celle-ci peut être considérée comme radiale à partir du point d'intersection de l'axe optique avec le plan image.

La connaissance du modèle de distorsion, utile pour le calcul de la distance focale utilise, quant à elle des images de grille.

L'article se subdivise en quatre parties. Dans le chapitre 2, nous décrivons les divers repères utilisés ainsi que les bases mathématiques de notre approche.

Le chapitre 3 décrit la mise en œuvre de la méthode et le dernier chapitre présente des résultats.

2. Bases mathématiques de la méthode.

Notre but est de déterminer les paramètres intrinsèques d'une caméra. Ces derniers, au nombre de quatre représentent :

- les coordonnées (u_0, v_0) de l'intersection de l'axe optique avec le plan image, que nous nommerons *point principal*.
- le rapport $k = dy/dx$ des pas d'échantillonnage en x et y du système d'acquisition.
- la distance focale f de la caméra.

2.1. LES DIFFÉRENTS REPÈRES UTILISÉS

Trois repères utiles à la modélisation des caméras doivent être définis avant d'aborder le problème de l'étalonnage.

Le repère *égocentrique* est lié à la mire, on le notera (R_e) .

Le repère *rétinocentrique* ou repère *caméra* est lié à la caméra et sera noté (R_c) . Son centre O_c est le centre optique de la caméra, les axes O_cX_c et O_cY_c sont parallèles respectivement aux lignes et aux colonnes de l'image, et l'axe O_cZ_c est perpendiculaire au plan image.

Le repère *naturel* est un repère bidimensionnel lié à l'image elle-même. Il sera noté (R_n) .

Un point P^i de l'espace peut être repéré soit dans le repère égocentrique, où il sera noté $P_e^i(X_e^i, Y_e^i, Z_e^i)$, soit dans le repère rétinocentrique, où il sera noté $P_c^i(X_c^i, Y_c^i, Z_c^i)$.

Sa projection sera le point $P_n^i(u_n^i, v_n^i)$ dans le repère naturel de l'image.

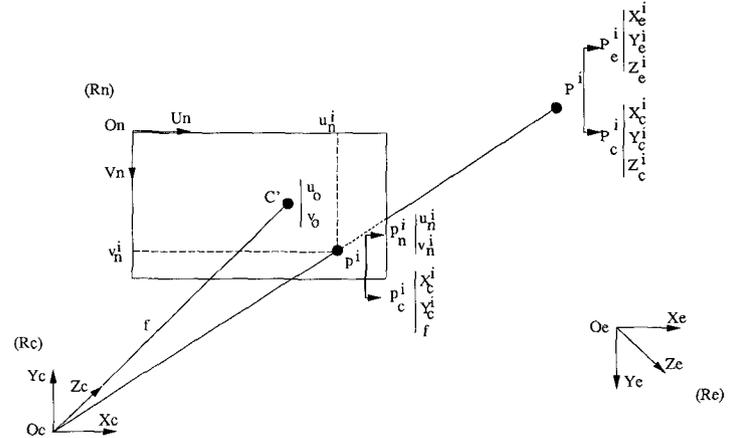


Figure 1.- Repères utiles.

2.2. MÉTHODE DE CALCUL

2.2.1. Détermination des paramètres (u_0, v_0, k)

Rappelons pour mémoire que les *limbes* d'une surface gauche sont les points de cette surface où le plan tangent passe par le centre optique de la caméra.

Ceci dit, le calcul des trois paramètres (u_0, v_0, k) est fondamentalement basé sur la propriété suivante des limbes d'une sphère.

Théorème :

La projetée des limbes d'une sphère dans le plan image est une ellipse dont le grand axe passe par le point (u_0, v_0) , intersection de l'axe optique avec le plan image.

Pour démontrer ce résultat, nous commencerons par exprimer l'équation de l'ellipse résultant de la projection des limbes d'une sphère, puis nous déterminerons l'équation de ses axes principaux, et nous en déduirons alors le résultat énoncé.

2.2.1.1. Vue perspective d'une sphère.

Reprenons le repère caméra défini en Figure 1, et soit S une sphère de rayon R centrée au point O de coordonnées (x_0, y_0, z_0) .

Considérons en outre un point P de coordonnées (X, Y, Z) situé sur les limbes de cette sphère. Ses coordonnées doivent vérifier les relations suivantes :

$$\begin{cases} \|\vec{CP}\|^2 + R^2 = \|\vec{OC}\|^2 \\ \vec{OP} \cdot \vec{CP} = 0 \end{cases}$$

Ces relations se traduisent par le système d'équations suivant :

$$\begin{cases} X^2 + Y^2 + Z^2 - x_0^2 - y_0^2 - z_0^2 + R^2 = 0 \\ (X - x_0)X + (Y - y_0)Y + (Z - z_0)Z = 0 \end{cases} \quad (1)$$

Recherchons l'équation du cône ayant pour sommet le centre optique C et s'appuyant sur les limbes de la sphère.

Cette équation est obtenue en considérant le lieu des points Q de coordonnées (x, y, z) tels que $\vec{CQ} = \lambda \vec{CP}$, c'est-à-dire tels que :

$$\begin{cases} x = \lambda X \\ y = \lambda Y \\ z = \lambda Z \end{cases} \quad (2)$$

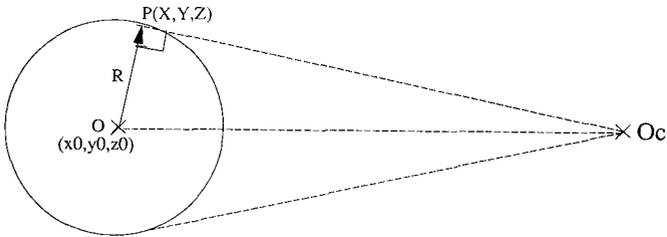


Figure 2.- Limbes d'une sphère.

L'équation du cône s'obtient alors en combinant (1) et (2) et s'écrit :

$$R_x x^2 + R_y y^2 + 2x_0 y_0 xy + 2x_0 z_0 xz + 2y_0 z_0 yz + R_z z^2 = 0$$

avec

$$\begin{cases} R_x = R^2 - y_0^2 - z_0^2 \\ R_y = R^2 - x_0^2 - z_0^2 \\ R_z = R^2 - x_0^2 - y_0^2 \end{cases} \quad (3)$$

Pour obtenir la projection de la sphère dans le plan image, il suffit de considérer l'intersection du cône et du plan situé à $z = f$, f étant la distance focale.

On aboutit à l'équation de l'ellipse suivante dans le repère caméra :

$$R_x x^2 + R_y y^2 + 2x_0 y_0 xy + 2x_0 z_0 xf + 2y_0 z_0 yf + R_z f^2 = 0 \quad (4)$$

qui s'écrit également :

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0 \quad (5)$$

avec

$$\begin{cases} A = R_x \\ B = x_0 y_0 \\ C = R_y \end{cases} \quad \begin{cases} D = x_0 z_0 f \\ E = y_0 z_0 f \\ F = R_z f^2 \end{cases} \quad (6)$$

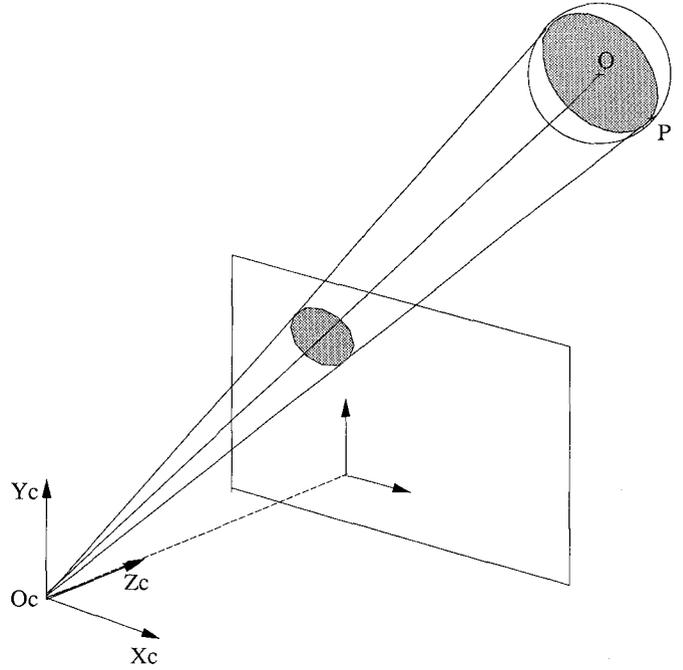


Figure 3.- Vue perspective d'une sphère.

2.2.1.2. Calcul de l'équation des axes principaux d'une ellipse

Considérons l'ellipse (E) centrée en (x_c, y_c) , d'équation (5) dans le repère caméra.

Pour déterminer l'équation résultante du produit des équations de ses deux axes principaux, nous recherchons les points de (E) minimisant ou maximisant la distance au centre de l'ellipse. Nous avons donc à résoudre un problème d'optimisation sous contrainte, que nous formulons à l'aide du Lagrangien :

$$L(x, y, \mu) = (x - x_c)^2 + (y - y_c)^2 + \mu(Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F)$$

Nous exprimons alors la condition du premier ordre en annulant le vecteur gradient de la fonction $L(x, y, \mu)$:

$$\frac{\delta L}{\delta x} = 0 \quad ; \quad \frac{\delta L}{\delta y} = 0 \quad ; \quad \frac{\delta L}{\delta \mu} = 0$$

En éliminant μ des deux premières équations, nous obtenons l'équation caractérisant les points dont la distance au centre est optimale. La troisième équation n'étant pas exploitée, ces points sont donc ceux appartenant aux axes principaux de l'ellipse. L'équation du produit des axes en fonction des coefficients de l'ellipse dans le repère caméra est donc :

$$Bx^2 + (C - A)xy - By^2 + (E + Ay_c - Bx_c)x + (By_c - Cx_c - D)y + Dy_c - Ex_c = 0 \quad (7)$$

De façon générale, cette équation s'écrit comme le produit des équations de deux droites orthogonales :

$$(ax + by + c_1)(-bx + ay + c_2)$$

Par identification avec l'équation (7), et en utilisant les relations (6), il est aisé de déterminer les valeurs de a et b et d'en déduire que les directions des axes principaux sont données par les vecteurs $(-y_0, x_0)$ et (x_0, y_0) .

Les équations des axes s'écrivent alors :

$$\begin{cases} -x_0x - y_0y + c_1 = 0 \\ -y_0x + x_0y + c_2 = 0 \end{cases}$$

La détermination des constantes c_1 et c_2 se fait en exprimant que les deux axes principaux passent par le centre (x_c, y_c) de l'ellipse. Or d'après [3], celles-ci vérifient les relations :

$$\begin{cases} Ax_c + By_c + D = 0 \\ Bx_c + Cy_c + E = 0 \end{cases}$$

On en déduit donc :

$$\begin{cases} x_c = \frac{BE - CD}{AC - B^2} \\ y_c = \frac{BD - AE}{AC - B^2} \end{cases}$$

En injectant ces valeurs dans les équations des deux axes principaux, on montre que la constante c_1 s'annule, ce qui exprime bien que l'un des axes principaux passe par le point $(0, 0)$ dans le repère caméra, qui n'est rien d'autre que le point (u_0, v_0) dans le repère image.

C'est sur cette propriété géométrique de la projection perspective des sphères qu'est basée notre méthode d'étalonnage.

2.2.1.3. Critère de minimisation

Considérons l'équation (7); si nous y injectons les coordonnées $(0, 0)$, il est clair, d'après la propriété que nous venons de démontrer, que l'expression obtenue doit s'annuler. Cette expression n'est rien d'autre que le produit des distances du point principal aux axes principaux de l'ellipse exprimée dans le repère caméra. C'est donc cette quantité, notée *prodist* que nous minimiserons.

$$prodist = DY_c - Ex_c = \frac{1}{AC - B^2} (D(BD - AE) - E(BE - CD)) \quad (8)$$

2.2.1.3.1. Calcul dans le repère image

Pour pouvoir utiliser le résultat, nous devons l'exprimer dans le repère naturel de l'image.

L'équation de l'ellipse (E') dans le repère image s'obtient à partir de l'équation (4) définie dans le repère caméra, en introduisant les formules de changement de repères :

$$\begin{cases} x = (u - u_0)dx \\ y = (v - v_0)dy \end{cases}$$

On aboutit à une équation de la forme :

$$A'x^2 + 2B'xy + C'y^2 + 2D'x + 2E'y + F' = 0$$

avec :

$$\begin{cases} A' = R_x dx^2 \\ B' = x_0 y_0 dx dy \\ C' = R_y dy^2 \\ D' = (x_0 z_0 f - u_0 R_x dx - x_0 y_0 v_0 dy) dx \\ E' = (y_0 z_0 f - v_0 R_y dy - x_0 y_0 u_0 dx) dy \\ F' = R_x dx^2 u_0^2 + 2x_0 y_0 u_0 v_0 dx dy + R_y dy^2 v_0^2 - 2x_0 z_0 u_0 f dx - 2y_0 z_0 v_0 f dy + R_z f^2 \end{cases}$$

On peut alors établir des relations entre les coefficients de l'ellipse (E) et ceux de l'ellipse (E') :

$$\begin{cases} A = A'/dx^2 \\ B = B'/dx dy \\ C = C'/dy^2 \\ D = (D' + B'v_0 + A'u_0)/dx \\ E = (E' + B'u_0 + C'v_0)/dy \\ F = F' + 2E'v_0 + 2D'u_0 + C'v_0^2 + 2B'u_0v_0 + A'u_0^2 \end{cases} \quad (9)$$

Ces relations nous permettent d'exprimer la valeur de *prodist* en fonction des paramètres intrinsèques de la caméra, et des coefficients de l'ellipse détectée dans l'image.

2.2.2. Calcul de la distance focale

Les développements mathématiques précédents font apparaître que le dernier paramètre intrinsèque de la caméra, qui est la distance focale, s'obtient analytiquement de façon très simple.

En effet, à partir de l'équation de l'ellipse exprimée dans le repère caméra, après normalisation des coefficients, sous la forme :

$$Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + 1 = 0$$

et des systèmes d'équations (3) et (6), nous avons à résoudre un système de 5 équations à 5 inconnues (x_0, y_0, z_0, R, f) . Après élimination des 4 autres inconnues, nous obtenons pour f la solution suivante :

$$f = \pm \sqrt{\frac{(B/DE) - 1}{(AB/DE) - (B/E)^2}}$$

Démonstration

Les coefficients A, B, C, D, E de l'ellipse dans le repère caméra, après normalisation, s'écrivent :

$$\begin{cases} A = R_x/R_z f^2 \\ B = x_0 y_0/R_z f^2 \\ C = R_y/R_z f^2 \\ D = x_0 z_0 f/R_z f^2 \\ E = y_0 z_0 f/R_z f^2 \end{cases} \quad \begin{cases} AR_z f^2 - R_x = 0 \\ BR_z f^2 - x_0 y_0 = 0 \\ CR_z f^2 - R_y = 0 \\ DR_z f^2 - x_0 z_0 f = 0 \\ ER_z f^2 - y_0 z_0 f = 0 \end{cases}$$

Le système d'équations (3) permet d'exprimer R_x et R_y en fonction de R_z , ce qui donne pour le système précédent :

$$\begin{cases} (Af^2 - 1)R_z + z_0^2 - x_0^2 = 0 \\ BR_z f^2 - x_0 y_0 = 0 \\ (Cf^2 - 1)R_z + z_0^2 - y_0^2 = 0 \\ DR_z f^2 - x_0 z_0 = 0 \\ ER_z f^2 - y_0 z_0 = 0 \end{cases} \quad (10)$$

Par ailleurs, x_0, y_0 et R_z s'expriment en fonction de z_0 par :

$$y_0 = \frac{B}{D} z_0 f \quad ; \quad x_0 = \frac{B}{E} z_0 f \quad ; \quad R_z = \frac{B}{DE} z_0^2$$

Par substitution dans la première équation du système (11), on obtient alors : $z_0^2 [(Af^2 - 1) \frac{B}{DE} + 1 - (\frac{B}{E})^2 f^2] = 0$

D'où :

$$f^2 = \frac{(B/DE) - 1}{(AB/DE) - (B/E)^2} \quad (11)$$

3. Résolution du problème d'étalonnage

3.1. MÉTHODE

L'estimation des paramètres intrinsèques de la caméra se fera en deux temps :

Première étape :

Nous déterminerons tout d'abord les paramètres $u_0, v_0, k = dy/dx$ par annulation du produit des distances du point (u_0, v_0) aux axes de l'ellipse détectée.

Ce produit s'annule si $N = D(BD - AE) - E(BE - CD) = 0$.

La méthode de résolution utilisée est celle de Newton-Raphson appliquée au développement limité au premier ordre de la quantité N au point (u_0^k, v_0^k, k^k) , vecteur solution à l'étape k du processus de minimisation :

$$N(u_0, v_0, k) = N(u_0^k, v_0^k, k^k) + \frac{\delta N}{\delta u_0} du_0 + \frac{\delta N}{\delta v_0} dv_0 + \frac{\delta N}{\delta k} dk$$

La valeur de N à la convergence étant nulle, l'équation se simplifie en :

$$N(u_0^k, v_0^k, k^k) = -\frac{\delta N}{\delta u_0} du_0 - \frac{\delta N}{\delta v_0} dv_0 - \frac{\delta N}{\delta k} dk \quad (12)$$

La détermination des 3 inconnues du_0, dv_0, dk nécessite donc d'avoir 3 équations au moins du type (12). Dans la pratique, nous utiliserons plus de trois images de sphères réparties autour de l'axe optique afin de moyenniser le "bruit" existant sur la valeur des coefficients des ellipses détectées.

Le système est résolu de façon itérative; à chaque itération, les corrections du_0, dv_0, dk à apporter sont alors utilisées pour estimer les paramètres A, B, C, D, E . La méthode de résolution utilisée est celle dite SVD [12] (Décomposition et Résolution par Valeurs Singulières).

Les dérivées partielles s'expriment simplement, en utilisant le système (9), en fonction des coefficients de l'ellipse (E) exprimée dans le repère caméra :

$$\begin{cases} \frac{\delta N}{\delta u_0} &= A(BD - AE) - B(BE - CD) - E(B^2 - AC) \\ \frac{\delta N}{\delta v_0} &= k(B(BD - AE) - C(BE - CD) + D(B^2 - AC)) \\ \frac{\delta N}{\delta k} &= \frac{1}{k}(2E(BE - CD) - N) \end{cases}$$

Deuxième étape :

Le calcul de la distance focale, s'il reste simple dans sa forme analytique, dépend cependant uniquement des paramètres des ellipses détectées. Sur des images de synthèse, une seule image suffit à déterminer exactement la focale; par contre, si nous bruitons ces images, les coefficients des ellipses sont modifiés et le calcul de la focale s'en trouve faussé. Sur des images réelles, nous déterminerons donc la focale en effectuant une moyenne des différents calculs associés à chaque ellipse.

Par ailleurs, comme dans tout système d'acquisition d'images, nous ne pouvons échapper au phénomène de distorsion inhérent au processus de formation de l'image.

Des discussions avec les calculateurs optiques de la société Angénieux nous ont amené à considérer que seule la distorsion radiale devait être prise en compte, les autres types de distorsion étant négligeables lorsque la fabrication des lentilles et leur ajustement sont effectués de façon rigoureuse.

Il faut noter que cette distorsion n'affecte pas la détermination des trois premiers paramètres; en effet, s'il est clair que la distorsion radiale nous amène à approximer des courbes qui ne sont plus réellement des ellipses, elle préserve cependant la symétrie de nos données par rapport à l'axe de l'ellipse passant par le point (u_0, v_0) (qui est le centre de distorsion). En conséquence, l'ellipse approximante que nous détectons aura le même axe que l'ellipse idéale (qui n'aurait pas subi la distorsion). Des expériences sur des images de synthèse confirment ce postulat (voir paragraphe 4.1.1).

Si donc cette distorsion n'affecte pas notre méthode en ce qui concerne la détermination de u_0, v_0, k , il n'en est pas de même pour la distance focale (puisque la distorsion affecte les dimensions de l'image). Il est donc indispensable d'effectuer une correction du positionnement des points de contour de l'image avant de lancer notre processus de détermination de ce paramètre.

Un deuxième facteur influant sur la détermination de la distance focale est la localisation de la projection des limbes.

Les deux paragraphes qui suivent exposent les principes que nous avons utilisés et les méthodes développées pour l'estimation de la distorsion d'une part, et la localisation de la projection des limbes d'une sphère d'autre part.

3.2. ESTIMATION DE LA DISTORSION RADIALE D'UN SYSTÈME OPTIQUE

La méthode utilisée ayant déjà fait l'objet d'un article [13], nous nous contenterons ici d'en rappeler l'essentiel.

Les images que nous nous proposons de traiter, pour estimer la distorsion radiale induite par l'objectif dont est muni notre système de prise de vue, sont des images de grilles planaires constituées de deux ensembles orthogonaux de lignes parallèles et équidistantes,

dont la projection dans le plan image est détectée précisément [11] et approximée par un ensemble de B-Splines [10] et [7].

L'algorithme développé est fondé sur l'utilisation du birapport de quatre points alignés, invariant projectif. La modélisation, proprement dite, de la fonction de distorsion est, quant à elle, réalisée également à l'aide d'une B-Spline.

3.2.1. Estimation de la distorsion radiale par utilisation de B-Splines

Par commodité, dans ce paragraphe les points de l'image seront exprimés en coordonnées polaires de pôle (u_0, v_0) .

3.2.1.1. Formulation du problème

Chaque point p de l'image a subi une distorsion radiale. Soit (r, θ) ses coordonnées polaires. Ces dernières sont liées aux coordonnées (r_c, θ_c) du point image idéal corrigé p_c par les relations :

$$r_c = r + \delta(r), \quad \theta_c = \theta.$$

Nous avons choisi d'utiliser une B-Spline de degré 3 pour modéliser la fonction δ .

Pour sa détermination, nous choisissons P abscisses R_i équiréparties sur l'intervalle des valeurs de distances représentatives dans l'image. On a alors :

$$\delta(r) = \sum_{i=1}^P \beta_i(s_r) \Delta_i,$$

s_r est la valeur du paramètre de parcours de la B-Spline, associée à p . Ici, s_r est proportionnelle à la distance r de p à (u_0, v_0) . Les Δ_i (ordonnées des points de contrôle d'abscisses R_i) sont fonction de la distorsion à estimer et représentent nos inconnues.

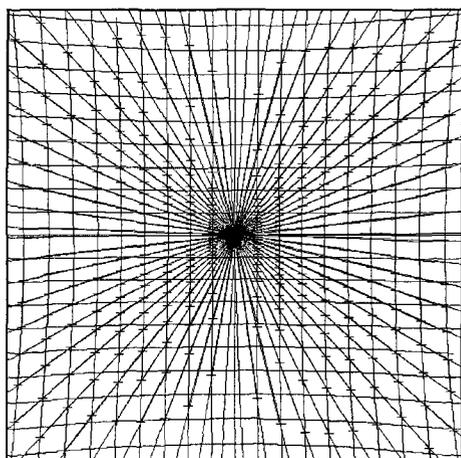


Figure 4.— Faisceau de 31 sécantes (sur les 80 utilisées) et points d'intersections participant aux quadruplets retenus sur les courbes "verticales".

3.2.1.2. Estimation à l'aide du birapport

Rappel : le birapport de 4 points alignés notés, $[A, B, C, D]$ est la quantité signée :

$$[A, B, C, D] = \frac{\overline{AC} \overline{DB}}{\overline{AB} \overline{DC}}$$

Le birapport est un invariant projectif. Si l'image résulte d'une projection centrale sans distorsion ce birapport est identique, qu'il soit calculé sur quatre points alignés d'un objet réel ou sur les quatre projections correspondantes dans l'image.

3.2.1.2.1. Choix des quadruplets

Nous considérons les birapports obtenus pour 4 points d'intersection q, m, p, v d'une droite D passant par (u_0, v_0) avec le faisceau des approximations dans l'image des lignes verticales (resp. horizontales) d'une grille placée devant la caméra.

Afin de supprimer une inconnue (la valeur du birapport) nous nous sommes restreints à choisir des quadruplets pour lesquels la valeur algébrique attendue du birapport K devait être 4.

$$K = \frac{\overline{qp} \overline{mv}}{\overline{qm} \overline{pv}} = 4$$

3.2.1.2.2. Construction du critère

Afin de suffisamment contraindre notre système, nous considérons dans notre image tous les sous-ensembles de quatre lignes, projections de quatre droites horizontales (resp. verticales), régulièrement espacées de notre grille. Tous les quadruplets de points à abscisses (resp. ordonnée) croissantes, résultants de l'intersection d'une droite passant par (u_0, v_0) avec les quadruplets de lignes, donneront le birapport 4 pour une distorsion corrigée.

Typiquement on retient les quadruplets d'intersection pour environ 80 droites passant par (u_0, v_0) . L'ensemble des quadruplets $Q = \{(q, m, p, v)\}$ d'intersection ainsi obtenus fournit un total de N birapports et nous pouvons écrire notre critère :

$$S_1 = \sum_{(q,m,p,v) \in Q} \left(4 - \frac{\overline{qp} \overline{mv}}{\overline{qm} \overline{pv}} \right)^2$$

ou encore :

$$S_1 = \sum_{(q,m,p,v) \in Q} F(q, m, p, v, \Delta_1, \dots, \Delta_p)^2$$

3.2.1.3. Contraintes supplémentaires

La minimisation du seul critère S_1 ne nous permettant pas d'obtenir des résultats satisfaisants, une contrainte supplémentaire est imposée au système portant sur l'alignement des points

appartenant à la projection d'une même droite de la grille. Nous avons donc procédé, pour chaque ligne, à la sélection aléatoire d'un ensemble de triplets de points. La condition d'alignement se traduit naturellement par la nullité du produit mixte de chaque triplet.

Si p_1, p_2 et p_3 sont trois points de l'image situés sur la projection d'une même droite de la grille, nous exprimons la nullité du déterminant suivant :

$$D(p_1, p_2, p_3, \Delta_1, \dots, \Delta_p) = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ t_1 & t_2 & t_3 \end{vmatrix}$$

où (x_i, y_i, t_i) est le système de coordonnées homogènes de p_i qui vérifie $x_i^2 + y_i^2 + t_i^2 = 1$.

Si $T = \{(p_1, p_2, p_3)\}$ désigne l'ensemble des M tels triplets, nous obtenons un nouveau critère à minimiser :

$$S_2 = \sum_{(p_1, p_2, p_3) \in T} D(p_1, p_2, p_3, \Delta_1, \dots, \Delta_p)^2$$

3.2.1.4. Critère final.

Pour tenir compte de nos deux conditions, nous nous ramenons finalement à un critère S combinaison de S_1 et S_2 selon A. Guézic [8].

La minimisation de S s'effectue par des techniques classiques, telles l'algorithme de Levenberg-Marquardt, ou même plus simplement celui de Newton-Raphson appliqué au critère obtenu en remplaçant F et D par leur développement limité au premier ordre.

3.3. LOCALISATION DES LIMBES D'UNE SPHÈRE.

La définition communément retenue pour la notion de contour dans une image de luminance peut être exprimée ainsi.

Définition :

Les contours dans une image de luminance se situent là où le module du gradient de l'intensité lumineuse est d'une part suffisamment grand et d'autre part maximal dans la direction de ce même gradient.

Divers algorithmes performants sont fondés sur cette notion, en particulier ceux de R. Deriche [5] et J.S. Chen [4].

Les figures ci-dessous représentent une vue en coupe de la ligne médiane d'une image de sphère blanche sur fond noir, ainsi qu'une image grise agrandie. La vue en coupe d'une ligne représente donc le niveau de gris des pixels de la ligne en fonction de leur position.

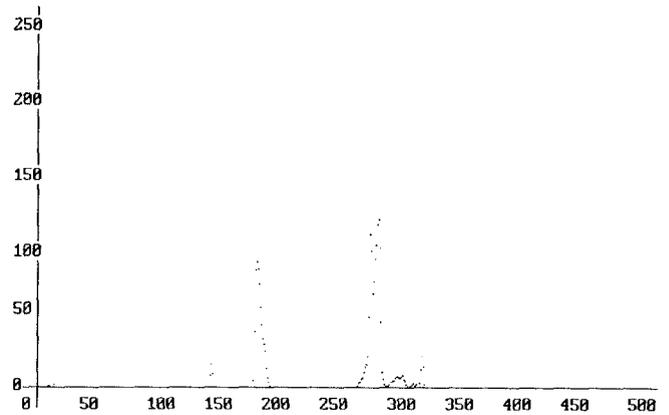


Figure 5.- Vue en coupe de la ligne médiane d'une image de sphère.

Les points des limbés d'une sphère étant les points où la normale à la surface est perpendiculaire à leur direction de projection, sont, par définition, des points non éclairés donc de niveau de gris nul. Nous avons donc cherché à détecter les points correspondant au début du front montant de la réponse de la caméra (d'abscisses approximatives 180 et 280); ces points sont en fait des points de courbure maximale.

Les algorithmes précédemment cités permettent de détecter les points situés au milieu du front montant de la réponse (là où le module du gradient est maximal).

L'expression de la courbure d'une courbe monodimensionnelle $y = f(x)$ est donnée par :

$$\rho = \frac{y''}{(1 + y'^2)^{3/2}}$$

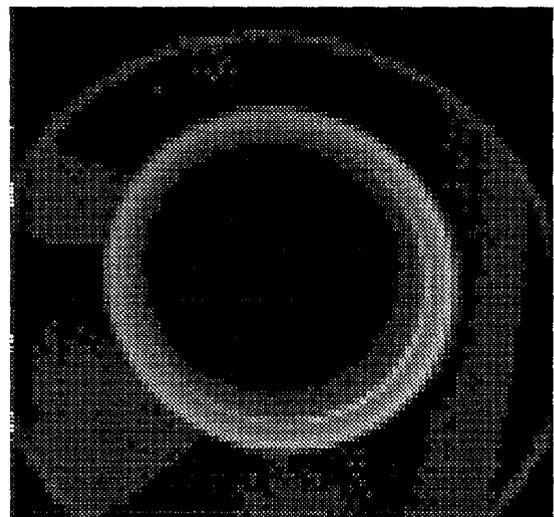


Figure 6.- Image grise agrandie d'une sphère.

Une approximation pour la recherche du maximum de courbure consiste à négliger le dénominateur et à rechercher le maximum de la dérivée seconde.

Nous avons étendu notre raisonnement à la courbe bidimensionnelle que constitue une image grise et avons appliqué l'algorithme de Deriche successivement à l'image grise et à l'image des modules du gradient obtenue à la suite du premier passage, avant de faire la recherche des maxima locaux.

A titre de comparaison, nous avons représenté sur la figure 7 le contour classiquement obtenu par l'application de l'algorithme de Deriche sur l'image grise (contour interne) et le contour obtenu par application de l'algorithme de Deriche sur l'image des modules du gradient (contour externe), superposés à une portion de l'image grise correspondante.

Nous observons bien un déplacement du contour vers le bord de l'image.

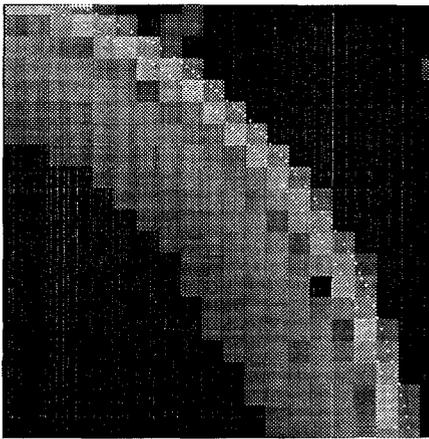


Figure 7.— Comparaison des contours détectés.

4. Expériences

Des expériences sur des images de synthèse ont été réalisées pour valider les deux approches théoriques développées précédemment. En ce qui concerne le processus de calibration par images de sphères, ces expériences permettent de mettre en évidence les avantages et également les limites de la méthode.

Quant au processus de l'estimation de la distorsion radiale, l'utilisation d'images de synthèse permet de générer de très fortes distorsions afin de tester la robustesse de la méthode proposée.

Les expériences sur des images réelles se décomposent de la façon suivante :

- validation de la méthode d'estimation de la distorsion radiale par comparaison de la modélisation par une B-Spline avec les courbes théoriques données par ANGENIEUX.
- acquisitions conjointes d'images de sphères et d'images de grilles permettant de mettre en œuvre le processus complet de calibration du système optique, c'est-à-dire : détermination du point principal et du rapport de digitalisation

en x et en y , estimation de la distorsion radiale, correction des images de sphères, détermination de la distance focale.

Le système d'acquisition d'images se compose d'une caméra CIDTEC CID 2250 équipée d'un zoom ANGENIEUX 14x9 BEAULIEU.

4.1. EXPERIENCE SUR LES IMAGES DE SPHERES.

4.1.1. Images de synthèse

Nous avons généré des images de synthèse 512x512 pixels correspondant à la projection des limbes d'une sphère de rayon 3 cm, située à une distance du centre optique de 0.8 m, 1 m et 1.2 m, pour des focales respectives de 1500 pixels, 2000 pixels et 2500 pixels, une position du point principal dans le repère image à $(u_0, v_0) = (255, 255)$ et un rapport k de digitalisation en x et y de 1. Le modèle de distorsion utilisé est inspiré de celui des photogrammètres [2] :

$$x_d = x(1 + k_1 r^2 + k_2 r^4)$$

$$y_d = y(1 + k_1 r^2 + k_2 r^4)$$

avec : (x, y) : coordonnées du point dans l'image de synthèse, (x_d, y_d) : coordonnées du point dans l'image distordue, r : distance du point (x, y) au centre (u_0, v_0) de l'image, k_1, k_2 : coefficients de distorsion.

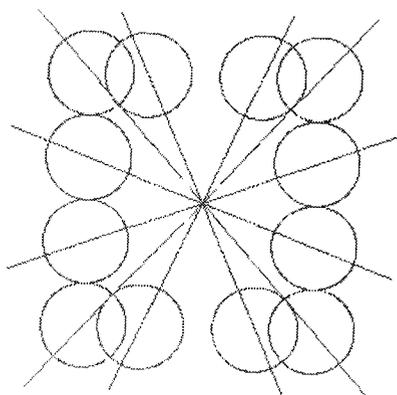
Cette fonction de distorsion présente l'avantage de donner directement les coordonnées (x_d, y_d) des points distordus en fonction de celles connues (x, y) des points synthétisés.

4.1.2. Première série d'expériences : mise en évidence de la non-influence de la distorsion radiale sur la détermination de (u_0, v_0, k) .

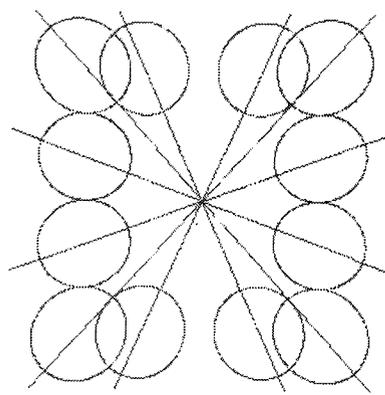
Les figures page suivante montrent un exemple de la convergence des grands axes pour des images non distordues et des images distordues avec des coefficients de distorsion $k_1 = 10^{-6}$, $k_2 = 10^{-12}$, non bruitées et bruitées à 1/10ème de pixel sur chacun des points.

Les tableaux suivants donnent des exemples de résultats obtenus pour différentes conditions d'expérimentations.

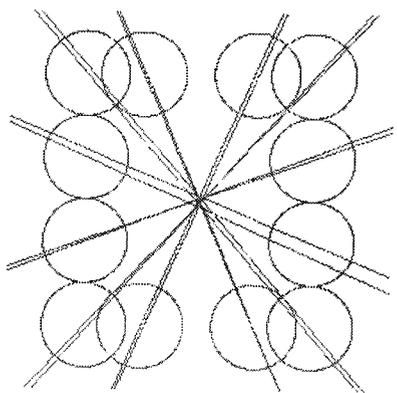
L'analyse de ces résultats montrent bien la non-influence de la distorsion radiale sur les images bruitées et non bruitées, quelle que soit la focale utilisée. Par contre, sur les images non distordues, l'influence du bruit est d'autant plus important que la focale est plus longue, ce qui s'explique par le fait que l'effet perspectif est alors moins fort que dans le cas de focales courtes; l'image résultante étant alors presque circulaire, il est plus délicat d'en déterminer les axes, et la précision sur les paramètres de calibration s'en trouve diminuée.



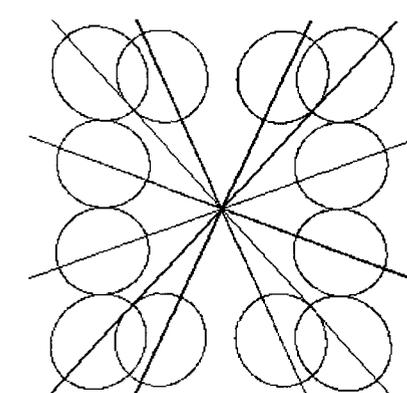
bruit nul, distorsion nulle.



bruit nul, $k_1 = 10^{-6}, k_2 = 10^{-12}$.



bruit = 1/10ème px, distorsion nulle.



bruit 1/10ème px, $k_1 = 10^{-6}, k_2 = 10^{-12}$.

	$f = 1500$	$f = 2000$	$f = 2500$
u_0	255.000	255.000	255.000
v_0	255.000	255.000	255.000
k	1.000	1.000	1.000

bruit nul, distorsion nulle.

	$f = 1500$	$f = 2000$	$f = 2500$
u_0	254.999	255.000	255.008
v_0	255.000	254.999	254986
k	1.000	1.000	1.000

bruit nul, $k_1 = 10^{-6}, k_2 = 10^{-12}$

	$f = 1500$	$f = 2000$	$f = 2500$
u_0	255.090	258.803	258.626
v_0	252.778	249.692	238.971
k	1.000	0.999	0.999

bruit = 1/10ème px, distorsion nulle.

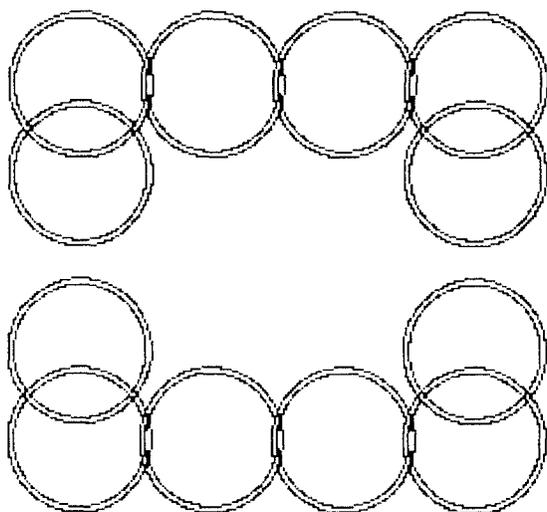
	$f = 1500$	$f = 2000$	$f = 2500$
u_0	254.867	254.772	255.239
v_0	255.899	255.869	254.459
k	1.001	1.001	0.999

bruit 1/10ème px, $k_1 = 10^{-6}, k_2 = 10^{-12}$

Sur des images réelles, nous aurons donc tout intérêt à lancer le processus de calibration des paramètres u_0, v_0, k sur des images brutes non corrigées. Par contre, nous serons limités quant au choix de la distance focale.

4.1.3. Deuxième série d'expériences : mise en évidence de l'influence d'une bonne détection de contours sur le calcul de la distance focale.

Nous avons simulé une erreur de détection de 3 pixels sur la distance au centre de l'ellipse de chacun des points de l'image. Nous obtenons donc une image concentrique à l'image initiale comme le montre la figure ci-dessous.



Le calcul de la distance focale se fait en moyennant toutes les valeurs calculées. Il apparaît ainsi une erreur de 2,5%, l'influence du bruit dans les images se faisant sentir essentiellement sur la valeur de l'écart-type associée à la valeur moyenne de la distance focale. Comme on pouvait l'imaginer, cet écart-type est d'autant plus faible que la distance focale est courte, pour les mêmes raisons que précédemment puisque seuls les coefficients de l'ellipse ramenés dans le repère de la caméra (donc dépendant de la bonne connaissance des paramètres u_0, v_0, k) interviennent dans le calcul de la focale.

4.1.4. Images réelles

La figure 8 montre le dispositif de prise de vue.

La sphère utilisée est une boule de billard placée au centre d'un cache plaqué contre un anneau lumineux (néon circulaire). Nous avons choisi ces conditions de prise de vue pour éclairer uniformément et avec un minimum de diffusion les points des limbes de la sphère. Pour un même réglage de la caméra (distance de mise au point et distance focale), nous faisons l'acquisition de plusieurs

images de cette sphère, réparties autour de la direction de l'axe optique de la caméra.

Les images page suivante donnent les résultats de l'étalonnage pour plusieurs séries de prises de vues.

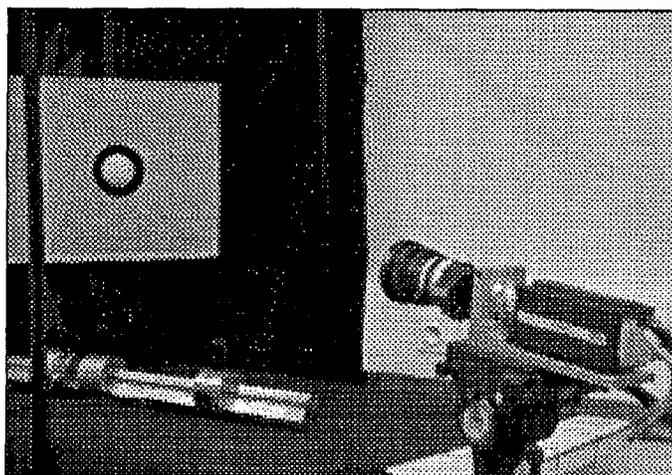


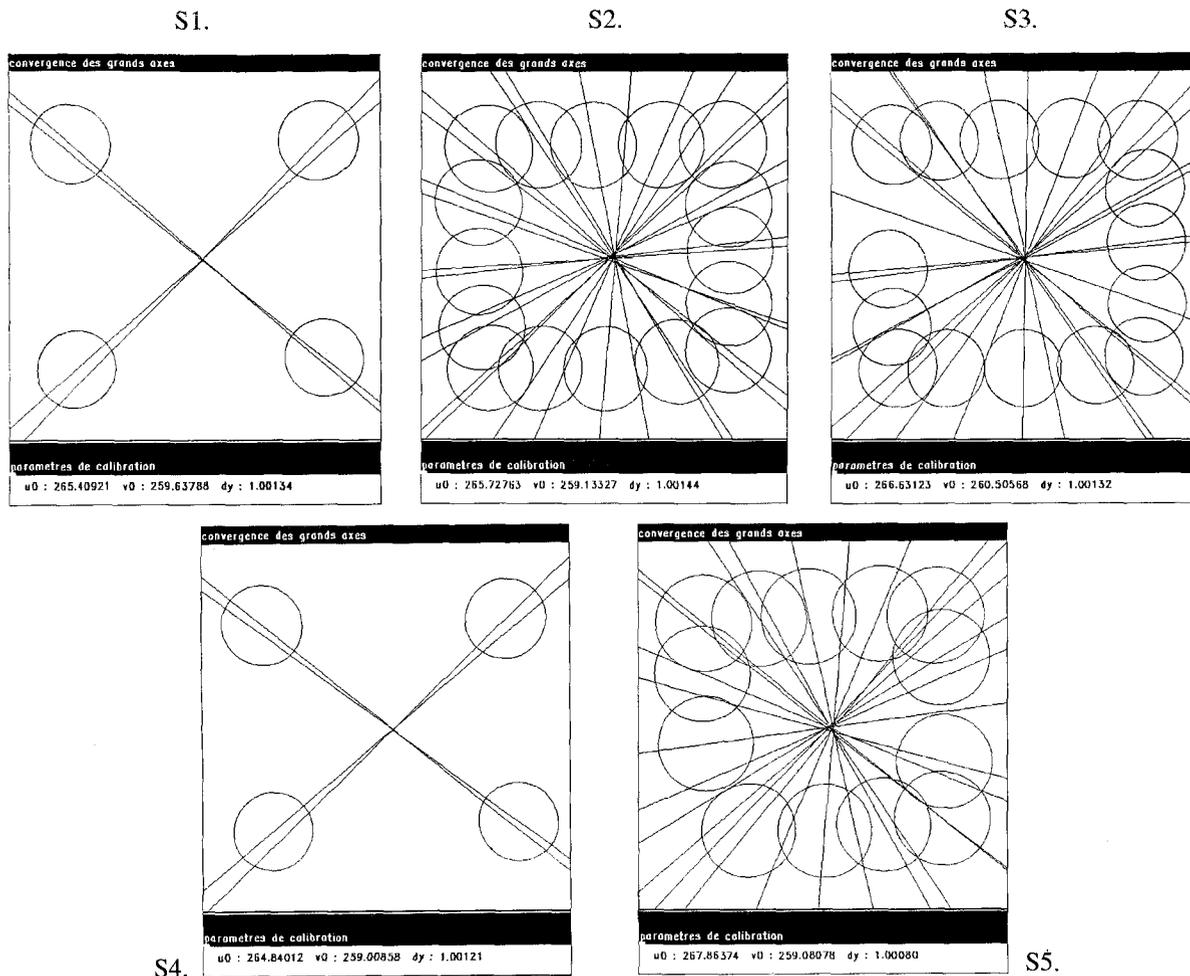
Figure 8.- Dispositif expérimental.

Le tableau suivant résume les résultats obtenus en fonction des valeurs de la distance de mise au point et de l'index focale, ces deux valeurs étant celles lues sur les graduations de l'objectif lors de la prise d'image :

Séries	u_0 (pixels)	v_0 (pixels)	$k = dy/dx$	Mise au point (m)	focale (mm)
S1	265,4	259,6	1,001	0,8	20
S2	265,7	259,1	1,001	0,8	20
S3	266,6	260,5	1,001	1	25
S4	267,9	259,1	1,001	1	30
S5	264,8	259,0	1,001	1	25

Les résultats S1 et S2 correspondent à la même série d'images de sphères, mais nous avons appliqué notre algorithme sur l'ensemble des images pour S2, et uniquement sur les quatre images de sphères les plus éloignées du centre de l'image pour S1. La série S5, quant à elle, n'est constituée que de 4 images de sphères.

La série S5 a été prise dans les mêmes conditions que la série S3, à quelques semaines d'écart, pour mettre en évidence la stationnarité des résultats.



4.2. EXPERIENCES SUR LES IMAGES DE GRILLE.

4.2.1. Images de synthèse.

Nous avons synthétisé l'image d'une grille parfaite déformée par un modèle de distorsion identique à celui utilisé pour la synthèse des images de sphères, et bruitée. Cette synthèse a été effectuée pour deux choix significatifs du couple de paramètres (k_1, k_2) (Cf. tableau 1), avec un bruit de $\pm 0,5$ pixels.

Tableau 1 : choix des valeurs de k_1 et k_2

	jeu n° 1	jeu n° 2
k_1	10^{-6}	-10^{-5}
k_2	10^{-12}	10^{-10}

Le jeu numéro 1 correspond à une faible distorsion et le numéro 2 à une distorsion forte. On peut voir sur les figures 10a et 10b les grilles déformées et bruitées et les courbes de correction ayant servi de modèle à leur construction figure 9.

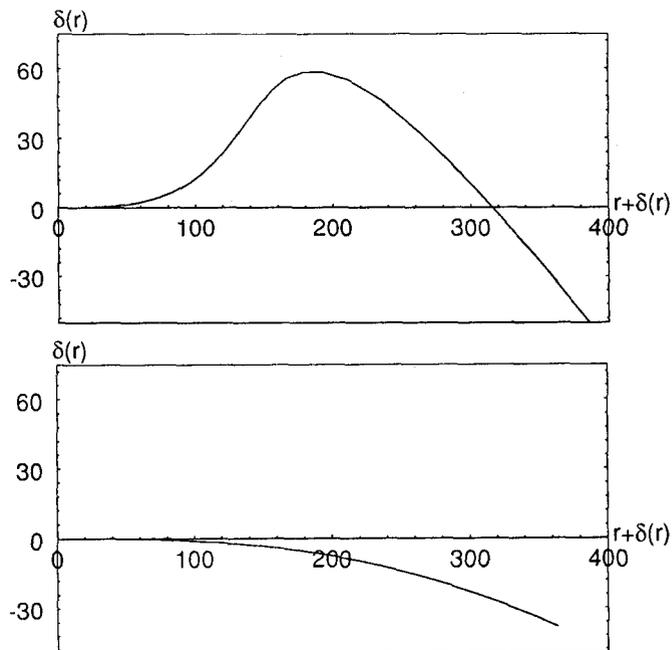


Figure 9.- Courbes $\delta(r) = k_1 r^3 + k_2 r^5$ pour les deux jeux de paramètres.

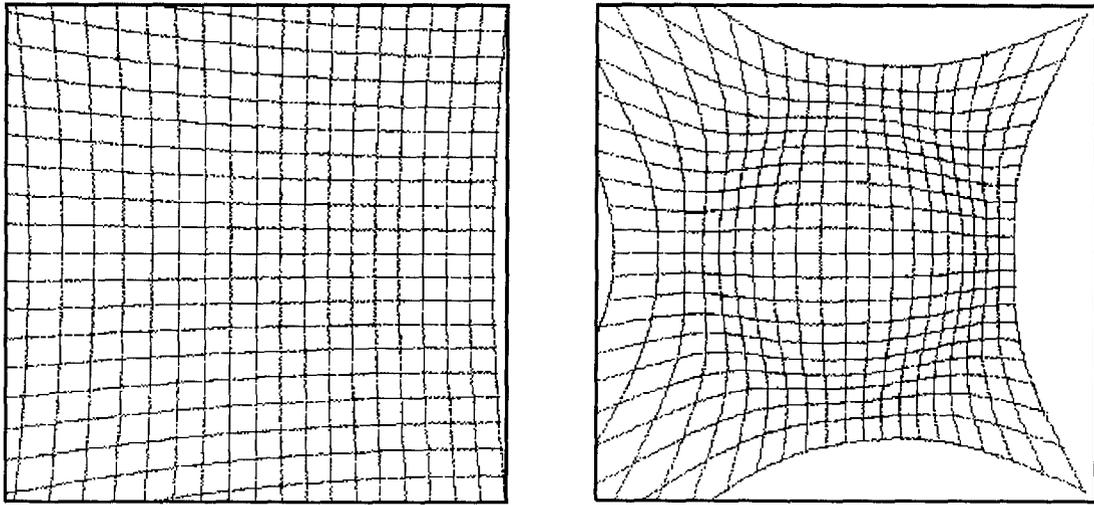


Figure 10.- Grilles de synthèse bruitées, (a) faiblement et (b) fortement distordues.

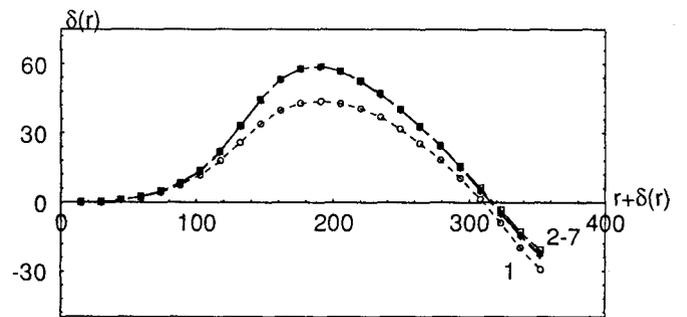
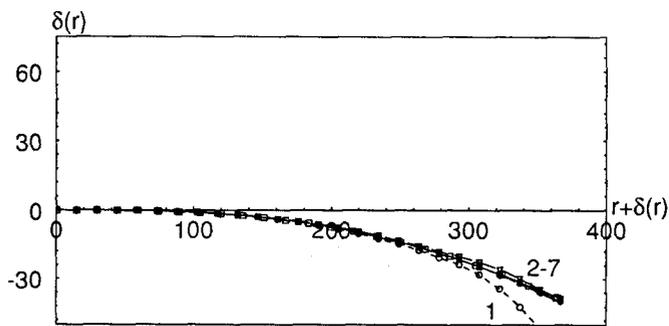


Figure 11.- Approximations successives dans le cas des grilles bruitées pour les deux jeux de paramètres.

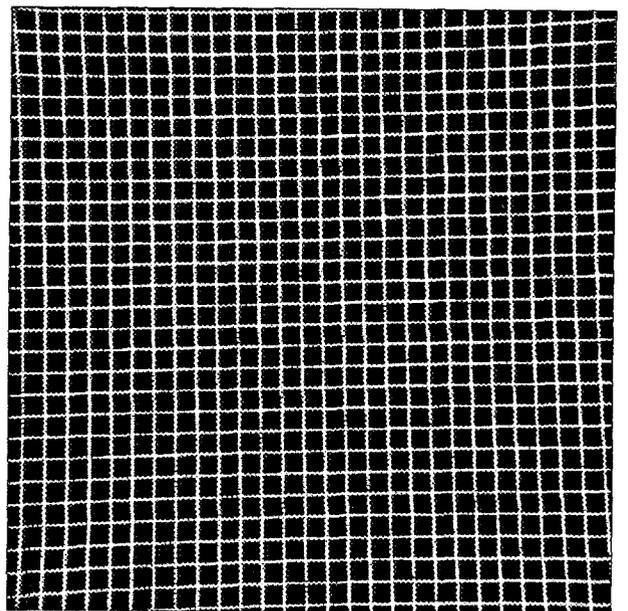
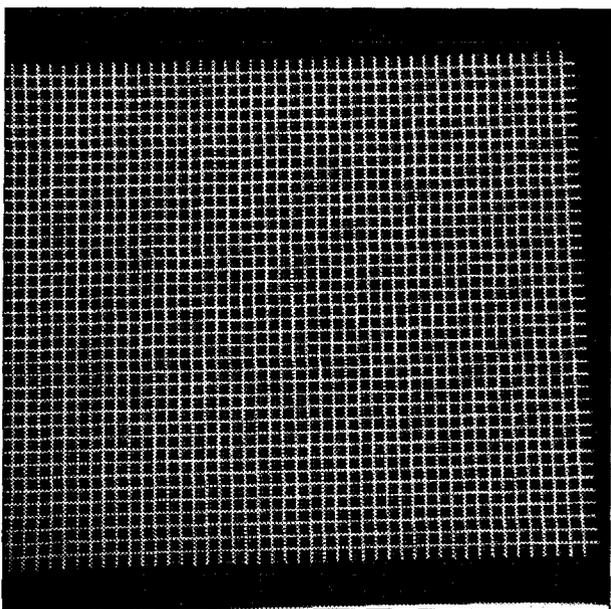


Figure 12.- Images réelles de grille pour une focale de (a) 9mm et (b) 15mm.

Notre processus de calcul de correction a été appliqué à ces grilles en prenant leurs intersections avec 80 droites passant par (u_0, v_0) . L'algorithme effectue moins de 10 itérations pour converger et l'estimation finale est pratiquement atteinte dès la deuxième itération.

La figure 11 présente la séquence des approximations successives de la spline de correction de distorsion.

Tableau 2 : moyenne et écart type des écarts entre les courbes de distorsion réelles et calculées.

	bruit ± 0.5
jeu n°1	$m = 0.04p, \sigma = 0.03p$
jeu n°2	$m = 0.2p, \sigma = 0.19p$

La qualité des résultats obtenus s'exprime dans le tableau précédent où pour chaque jeu de paramètres et chaque valeur du bruit, on peut lire la moyenne m et l'écart type σ des écarts absolus en pixels entre les courbes de distorsion réelles et calculées. Ces écarts sont calculés pour 25 points d'abscisses réparties entre 0 et 362, distance maximale d'un point de l'image à (u_0, v_0) .

4.2.2. Images réelles

Un processus analogue a été suivi sur deux images réelles de grille, prises à des focales de 9mm et 15mm (figure 12). Sur la figure 13, on trouvera la superposition des courbes de distorsion fournies par le constructeur du zoom et calculées par notre méthode.

De manière analogue nous avons calculé la moyenne et l'écart type des écarts absolus entre ces courbes de distorsion que l'on peut lire dans le tableau suivant.

focale 9mm	$m = 0.28, \sigma = 0.63$
focale 15mm	$m = 0.28, \sigma = 0.54$

Moyenne et écart type des écarts entre les courbes de distorsion réelles et théoriques.

4.3. Processus complet de calibration sur des images conjointes de sphères et de grilles.

Nous avons testé le processus de calibration complet sur 3 séries d'images de sphères et de grilles associées.

Nous avons donc pris simultanément des images de sphères et de grilles pour trois réglages différents de l'objectif.

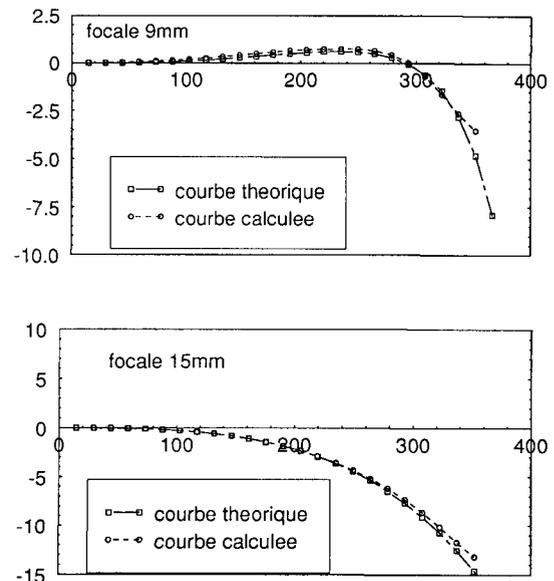
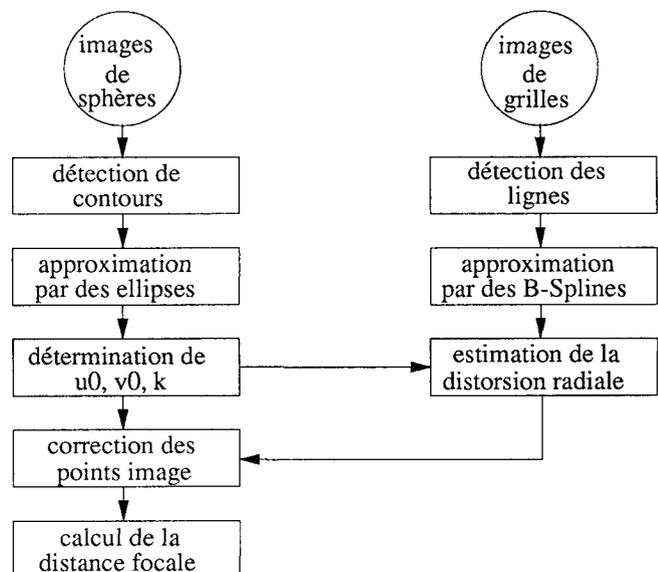


Figure 13.- Superposition des courbes de distorsion fournies par le constructeur du zoom et calculées par notre méthode pour une focale (a) 9mm et (b) 15mm.

Pour chacune des séries, la manipulation a été la suivante :



Le tableau ci-dessous rend compte des résultats obtenus :

Séries	u_0 (pixels)	v_0 (pixels)	k	focale calculée (mm)	focale constructeur (mm)
n°1	267,53	258,08	0,999	$23,97 \pm 3,12$	~ 22,54
n°2	265,16	258,27	0,999	$31,88 \pm 7,02$	~ 29,85
n°3	263,88	263,77	1,001	$38,09 \pm 14,78$	~ 40,47

– *focale calculée* : valeur calculée par notre méthode.

– *focale constructeur* : valeur théorique donnée par ANGENIEUX.

Cette valeur est obtenue à l'aide de jeux de tables fournies par le constructeur qui, à partir de la valeur de l'index de l'objectif (correspondant à la valeur de la focale pour une distance de mise au point infinie) et de la distance de mise au point utilisée fait correspondre la valeur réelle de la focale mise en jeu pour la série d'images considérée.

Si nous comparons les valeurs de la distance focale ainsi calculées avec les valeurs données par le constructeur, nous pouvons les considérer comme correctes. Cependant, les valeurs des écarts-types sont d'autant plus forts que la distance focale utilisée pour la prise d'images est longue, ce qui confirme les observations faites sur les images de synthèse.

Il est donc clair que, si nous avons confiance en la méthode de correction de la distorsion dans nos images, il reste encore sans doute des efforts à faire en ce qui concerne la détection des primitives elliptiques dans l'image pour apporter une meilleure stabilité des résultats lors de la détermination de la distance focale.

5. Conclusion

Cette nouvelle méthode d'étalonnage de système d'acquisition d'images que nous proposons, même si elle ne fait qu'augmenter le nombre de méthodes déjà existantes ou en cours d'étude [1], ne nous semble cependant pas dénuée d'intérêt.

En effet, nous avons choisi d'utiliser des primitives images jusqu'alors peu exploitées dans le domaine de l'étalonnage. Les points d'intérêt de l'image d'une sphère, qui sont les points de ses limbes, sont nombreux et jouissent tous de la même netteté. L'équation de l'ellipse à laquelle ils appartiennent peut donc être déterminée de façon précise.

Par ailleurs, la propriété géométrique mise en évidence présente l'avantage d'être insensible à la distorsion du système d'acquisition, dans la mesure où elle peut être considérée comme radiale autour de l'intersection de l'axe optique avec le plan image.

Notre méthode nous a alors permis de déterminer de façon simple et satisfaisante les coordonnées du point d'intersection de l'axe optique de la caméra avec le plan image (u_0, v_0) , ainsi que le rapport des pas d'échantillonnage en x et en y (dy/dx).

La détermination de la distance focale, quant à elle, est un peu plus délicate, même si la formulation mathématique est extrêmement simple.

En effet, interviennent alors deux facteurs modifiant la dimension et la forme des images acquises, la précision de la détection des points de contour et la distorsion géométrique du système d'acquisition. Si la correction de la distorsion ne pose pas de problème majeur, la précision de la détection des points des limbes d'une sphère est un problème délicat. Notre approche, bien qu'assez peu élaborée, nous amène à des résultats que l'on peut considérer comme corrects; les données théoriques en notre possession nous ont permis de les valider.

En conclusion, les résultats obtenus par cette nouvelle méthode de détermination des paramètres intrinsèques d'une caméra peuvent être considérés comme satisfaisants. Un des points forts de cette méthode est sans aucun doute la simplicité de mise en œuvre, tant du point de vue de la prise d'images que de la formulation mathématique.

BIBLIOGRAPHIE

- [1] P. BEARDSLEY, D. MURRAY, A. ZISSERMAN. *Camera Calibration Using Multiple Images*, In Proceedings of ECCV'92, Santa Margherita Ligure, Italie, Mai 92, pages 312-320.
- [2] H. BEYER. *Geometric and Radiometric Analysis of a CCD-Camera Based Photogrammetric Close-Range System*, Institut für Geodäsie und Photogrammetrie, Dissertation ETH n° 9701, Zürich, May 1992.
- [3] C. CAGNAC, E. RAMIS, J. COMMEAU, *Nouveau cours de Mathématiques Spéciales*, vol. n° 3 géométrie, éditions Masson & Cie, 1965.
- [4] J.S. CHEN and G. MEDIONI. *Detection, Localization and Estimation of Edges*. In *Proceedings of IEEE Workshop on Computer Vision*, Miami, November 1987, pages 215-217.
- [5] R. DERICHE, *Optimal Edge Detection Using Recursive Filtering*, In *Proceedings of First ICCV*, Londres, Juin 1987, pages 501-505.
- [6] O. FAUGERAS et G. TOSCANI, *Camera Calibration for 3D Computer Vision*, In *Proceedings of International Workshop on Machine Intelligence*, Tokyo, Février 1987.
- [7] Y. GARDAN, *Mathématique et CAO - Méthodes de base*, Hermès, Paris, 1985.
- [8] A. GUEZIEC, N. AYACHE, *Lissage et Reconnaissance de Courbes Gauches Bruitées*, *Revue Traitement du Signal*, Vol. 9, n° 6, pages 473-490, 1993.
- [9] P. LIMOZIN-LONG, *Présentation et Comparaison de trois méthodes de calibration*, *Convention de recherche ESA/INRIA*, Février 1988.
- [10] M.E. MORTENSON, *Geometric Modeling* John Wiley & Sons Editors, New-York, 1985.
- [11] B. PEUCHOT, *Camera Virtual Equivalent Model 0.012 Detectors*, *14th Int. conf. of IEEE Engineering in Medicine and Biology Society*, pages 41-45, Rennes, Novembre 1992.
- [12] H. PRESS, B. FLANNERY, S. TEUKOLSKY, W. VETTERLING, *Numerical Recipes, The Art of Scientific Computing*, pages 52-64.

- [13] S. RÉMY, M. DHOME, N. DAUCHER, J.T. LAPRESTÉ, Estimation de la distorsion radiale d'un système optique, *actes du 9ème Congrès RFIA de l'AFCEP*, Janvier 94, Paris.
- [14] R.Y. TSAI, An Efficient and Accurate Camera Calibration Techniques for 3D Machine Vision, In *Proceedings of Computer Vision and Pattern Recognition*, Miami, Juin 1986, pages 364-374.
- [15] G.Q. WEI, Z.Y. HE and S.D. MA, Camera Calibration for Stereo Vision by Vanishing Points, *Second Joint China-France Conference on Robotics*, Paris, 5-6-7 Décembre 1988.

Manuscrit reçu le 19 juillet 1992.

LES AUTEURS

Nadine DAUCHER Maître de Conférences à l'Université Blaise Pascal de Clermont-Ferrand, spécialité Automatique.

Recherche en vision monoculaire pour la robotique au sein du groupe GRAVIR.

Michel DHOME Chargé de recherches au CNRS; anime le groupe GRAVIR du LASMEA de l'Université Blaise Pascal de Clermont-Ferrand.

Jean-Thierry LAPRESTE Professeur à l'École Nationale Supérieure de Chimie de Clermont-Ferrand. Recherche effectuée au sein du groupe GRAVIR du LASMEA de l'Université Blaise Pascal à Clermont-Ferrand, dans le domaine de la vision monoculaire pour la robotique.