Calibration de Caméra à Base d'Ellipses

Ellipse-Based Camera Calibration

par J.-P. TAREL, A. GAGALOWICZ

INRIA
Domaine de Voluceau, Rocquencourt
B.P. 105, 78153 Le Chesnay Cedex, France.
Tél : (1) 39 63 54 79
E-mail: Jean-Philippe.Tarel@inria.fr, Andre.Gagalowicz@inria.fr

Résumé

En vision par ordinateur, il est souvent utile de connaître un modèle de la caméra pour pouvoir obtenir des informations 3D à partir des images. Nous présentons une méthode de calibration (ou calibrage) de caméra conçue pour fournir une estimation précise des paramètres du modèle de prise de vue. Pour avoir une bonne estimation, il est nécessaire avant tout d'avoir de bonnes mesures dans l'image. Nous avons choisi d'utiliser les moments géométriques d'une région comme les données 'image' à fournir à l'algorithme, car ils peuvent être calculés avec précision et robustesse en se fondant sur la photométrie des surfaces. Cela nous a conduits à développer une calibration à base d'ellipses. Pour cela, le lien exact entre la description d'une ellipse de l'espace 3D et les moments d'ordre inférieur ou égal à deux de sa projection a dû être explicité en fonction des paramètres de la calibration. Ces relations sont en partie étendues au cas où les distorsions sont modélisées.

Mots clés : Estimation sub-pixel, moments d'une forme, représentations de l'ellipse, distorsions, calibration de caméra, géométrie perspective.

Abstract

An important task in most 3D vision systems is camera modeling. We present a calibration method designed for a good accuracy in the parameters estimation. Accurate image measurments are required to perform the camera model estimation. We suggest a technique, based upon surface photometry analysis, to compute robustly and accurately the moments of a shape in an image, and use them as inputs for the estimation algorithm. As a consequences, an ellipse-based camera calibration is introduced, which uses the relationships between the specification of a 3D-ellipse, the moments of its image and the calibration parameters. These equations are partially extended to models including distortions.

Key words : *Sub-pixel estimation, shape moments, ellipse representations, distortions, camera calibration, perspective geometry.*

1. Introduction

Dans un processus de vision, il est utile de connaître la physique de l'instrument qui fournit les données de base : les images. Cette connaissance permet en particulier de calculer des informations tridimensionnelles de la scène à partir d'images, même si elle n'est pas forcément nécessaire [BMV93]. La calibration consiste à obtenir un modèle approché de la caméra. En toute généralité, les caractéristiques mécaniques, optiques, photométriques et électroniques doivent être modélisées. Pour le type d'applications généralement envisagées, on se restreint aux aspects géométriques de la calibration : digitalisation de l'image, distorsions géométriques dues à l'optique de l'objectif, localisation dans l'espace de la caméra. Les modèles considérés se fondent alors sur le modèle sténopé amélioré en tenant compte de différents types de distorsions. Le modèle permet d'expliciter la relation entre un objet de l'espace et son image. Comme les paramètres intrinsèques du modèle de caméra sont interdépendants (par exemple, le centre se déplace quand la focale change [LT88]), il est préférable d'estimer en même temps tous les paramètres du modèle sur une vue, comme dans la méthode proposée dans [FT86]. Une fois le modèle choisi, il reste alors à déterminer le type de données à utiliser pour l'estimation et le critère d'adéquation entre le modèle et ces données.

Celles-ci sont, le plus souvent, des points. Quelques méthodes utilisent d'autres types d'objets géométriques comme les lignes [CT90] [Vai90], ou comme les ellipses [HR92] [RZMF92] [Ma93]. Avec les points, selon la modélisation et le critère d'adéquation, le problème peut être formulé linéairement ou non, avec ou sans contraintes. Ceci conduit à des méthodes analytiques [Tsa86] [FT86] [PHYP93], ou à des algorithmes itératifs [CR89].

Pour avoir une estimation précise des paramètres, il faut contrôler un minimum l'environnement de la phase de calibration. Ainsi, une mire contrôlée géométriquement permet de disposer de données tridimensionnelles précises. Les conditions d'éclairage doivent être satisfaisantes pour que les données extraites de l'image soient suffisamment précises. Un grand nombre de données est généralement considéré pour améliorer l'estimation; néanmoins, la précision de celles-ci est directement liée à celle des mesures. Comme la précision de la méthode d'extraction des données qui sont fournies à l'algorithme de calibration est aussi importante que l'algorithme lui-même, une bonne approche du problème nécessite donc de concevoir la technique de mesure des données avec celle d'estimation du modèle.

Or l'utilisation des niveaux de gris d'une forme permet d'obtenir des informations précises sur celle-ci [Bey92][AATRM89]. Ainsi nous montrons que les moments géométriques d'une forme dans une image sont calculables avec précision, robustesse et simplicité. Après quelques rappels et notations (partie 2), la précision théorique de notre méthode de mesure des moments, fondée sur les niveaux de gris, est obtenue, de même que les conditions dans lesquelles ces mesures doivent être faites (partie 3). Nous avons pu relier la description d'un disque elliptique de l'espace aux moments géométriques de son image (partie 4). Ces équations permettent la conception d'une méthode de calibration basée non pas sur la correspondance entre points mais sur l'utilisation des moments d'une ellipse. D'une part, la méthode de résolution itérative employée est décrite dans la partie 5. D'autre part, nous présentons des résultats sur des images réelles et de synthèse (partie 6).

2. Modélisation de la caméra

Une caméra réalise une transformation des points de l'espace réel pour produire une image. On suppose généralement que la transformation réalisée est une simple projection centrale par rapport au centre de la caméra : c'est le modèle sténopé. Il peut être complété par la modélisation des diverses distorsions qui entachent généralement l'optique de la caméra. Voici un récapitulatif des modèles de caméra les plus utilisés.

2.1. Le modèle sténopé

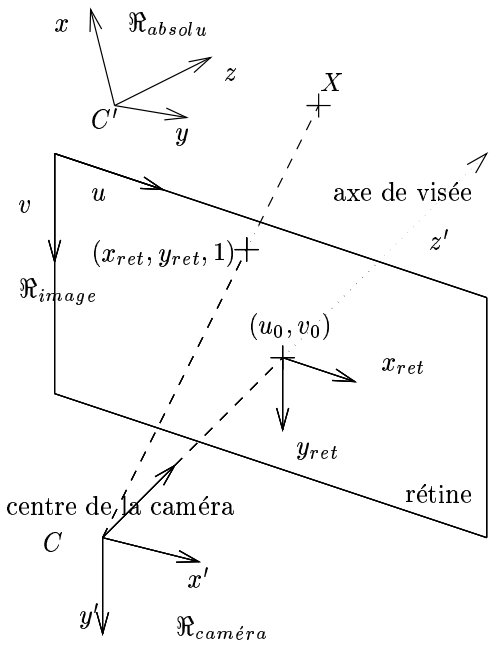


Fig. 1– *Le modèle sténopé de la caméra.*

Dans le modèle sténopé, l'optique est modélisée par une projection centrale sur le plan rétinien, de centre C (figure 2.1.). Un point P de l'espace est repéré par rapport au référentiel absolu R_{absolu}. Les paramètres extrinsèques de la caméra sont la rotation $R(p_1, p_2, p_3)$ et la translation $T(t_x, t_y, t_z)$ qui correspondent au déplacement entre les repères R_{absolu} et $R_{caméra}$.

La discrétisation est supposée s'effectuer sur une rétine rectangulaire avec un pas d'échantillonnage vertical $\frac{1}{a_v}$ et un pas horizontal $\frac{1}{a_u}$. Les paramètres intrinsèques sont les coordonnées (u_0, v_0) et l'inverse a_u et a_v des pas d'échantillonnage. (u_0, v_0) est la position dans l'image de l'intersection axe de visée-rétine. La relation entre les coordonnées (x^*, y^*) d'un point de la rétine et celles (u, v) dans R_{image} est donc :

$$\left.\begin{array}{l} u = u_0 + a_u x^* \\ v = v_0 + a_v y^* \end{array}\right\} \qquad (1)$$

En composant dans l'ordre - le déplacement rigide, la projection centrale et la numérisation (1) -, on obtient la relation entre un point de l'espace et son image. Cette relation s'exprime simplement en utilisant les coordonnées homogènes (en posant $X = (x\ y\ z\ 1)^t$ pour un point de l'espace) :

$$\left.\begin{array}{l} u = \frac{L_1 \cdot X}{L_3 \cdot X} \\ v = \frac{L_2 \cdot X}{L_3 \cdot X} \end{array}\right\} \text{ ou } U = \begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix} X \qquad (2)$$

La transformation complète peut donc être représentée par une matrice perspective 3x4, notée M.

$$M=\begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}=\begin{pmatrix} a_u & 0 & u_0 & 0 \\ 0 & a_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} R & T \\ 0 & 1 \end{pmatrix} \qquad (3)$$

On remarque que M doit vérifier deux contraintes pour être compatible avec son expression en fonction de R, T, u_0, v_0, a_u, a_v (3). Si on les oublie, la matrice est définie à un facteur d'échelle près. Dans ce cas, le modèle sténopé ne respecte plus l'orthogonalité des axes de la rétine.

2.2. Modèles avec distorsions

La modélisation précédente suppose que l'optique est parfaite. Les aberrations géométriques décrivent l'écart entre les positions observées et idéales de l'image d'un point de l'espace [Pér93]. Or la calibration peut estimer uniquement les distorsions puisque celles-ci provoquent des déformations indépendamment de la direction du rayon lumineux 'source'. Seuls les premiers termes du développement limité de cet écart sont retenus. Trois types de distorsions sont plus particulièrement considérés [Bey92] [WCH90] :

- La distorsion radiale, de paramètre α, est généralement la plus importante :

$$\left.\begin{aligned} x^d &= x^* + \alpha x^*(x^{*2} + y^{*2}) \\ y^d &= y^* + \alpha y^*(x^{*2} + y^{*2}) \end{aligned}\right\} \qquad (4)$$

- Le décentrage des lentilles par rapport à l'axe provoque une distorsion décrite par la relation suivante :

$$\left.\begin{aligned} x^d &= x^* + \beta(3x^{*2} + y^{*2}) + 2\gamma xy \\ y^d &= y^* + \gamma(x^{*2} + 3y^{*2}) + 2\beta xy \end{aligned}\right\} \qquad (5)$$

- Enfin, il résulte des petites erreurs de parallélisme entre lentilles une distorsion. La relation qui la décrit est :

$$\left.\begin{aligned} x^d &= x^* + \delta(x^{*2} + y^{*2}) \\ y^d &= y^* + \delta(x^{*2} + y^{*2}) \end{aligned}\right\} \qquad (6)$$

La prise en compte des distorsions conduit à une relation non linéaire entre un point de l'espace et son image. En effet, pour obtenir la relation qui lie un point à son image, il faut composer dans l'ordre le déplacement rigide, la projection centrale, les distorsions (4), (5) ou (6), et la numérisation (1).

3. Mesures dans l'image

Pour faire une bonne estimation des paramètres du modèle de caméra décrit précédemment, il est nécessaire d'avoir des données précises, notamment sur l'image, où la localisation doit être faite à au moins $1/10^{ème}$ de pixel près. Les caractéristiques géométriques d'une région de l'image comme l'aire, le barycentre et les moments d'inertie, sont faciles à mesurer avec une grande précision si on utilise les niveaux de gris. Sous certaines hypothèses, nous avons établi la précision théorique atteinte par notre méthode.

3.1. Principe

Un capteur CCD a une remarquable linéarité (les défauts de linéarité relatifs sont inférieurs à 10^{-3}, après une correction du gamma) sur une large plage de fonctionnement. La réponse d'une cellule s'avère donc proportionnelle à l'éclairement reçu par sa surface [Rag89]. Ainsi, si la caméra est pointée, avec un éclairage ambiant, vers un plan noir sur lequel une forme blanche est peinte, la valeur de chaque pixel de l'image obtenue est proportionnelle à l'aire de la surface blanche vue par le pixel. Or, connaître l'aire de l'intersection entre chaque pixel et l'image de la surface blanche est utile pour calculer une valeur approchée de ses caractéristiques géométriques. En particulier, l'utilisation des niveaux de gris permet d'obtenir le barycentre de la surface avec une précision supérieure au pixel.

Cette situation est idéale, mais il est possible de s'en approcher par un traitement approprié de l'image, si l'éclairage et la mire sont bien conditionnés. En effet, se placer dans des conditions telles que les gradients d'intensité soient linéaires n'est pas difficile. La surface de la mire doit être peu spéculaire pour être supposée à photométrie lambertienne, dans de bonnes conditions d'illumination. Ceci nécessite un traitement spécial pour les surfaces métalliques. On suppose donc que l'éclairage est ambiant ou qu'il provient de sources éloignées. Dans ces conditions, les pentes des gradients lumineux peuvent être corrigées sur l'image de la mire, pour se ramener à la situation idéale.

3.2. Correction des gradients d'intensité

Dans l'hypothèse lambertienne, une surface rayonne de façon diffuse une fraction de la lumière reçue. En chaque point du plan P (figure 3.2.), l'intensité émise dans une

direction quelconque est :

$$I = \frac{\rho E d}{p^3}$$

où ρ est le coefficient de réflection diffuse, et E l'émittance de la source.

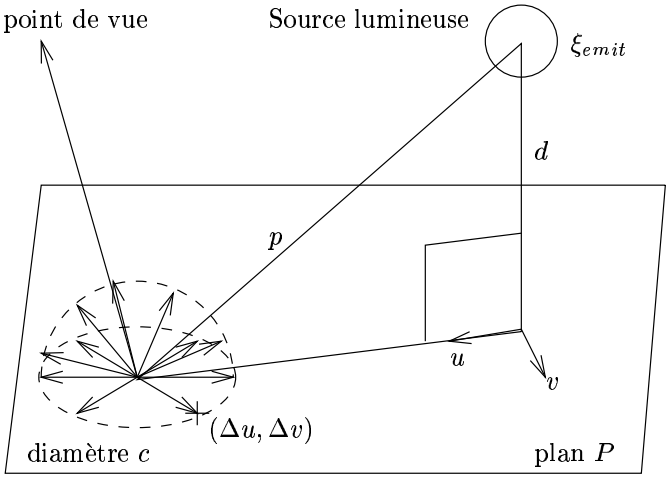


Fig. 2– *Réflexion diffuse d'une surface lambertienne.*

Si l'on approxime par un plan l'intensité réémise I, au voisinage d'un point (x_0, y_0), la fonction d'erreur relative au second ordre est :

$$\mathbf{e_{photo}} \simeq \frac{\Delta^2 I}{I} = \frac{\partial^2 I}{I \partial x^2}\Delta^2 x + \frac{\partial^2 I}{I \partial y^2}\Delta^2 y + \frac{\partial^2 I}{I \partial x \partial y}\Delta x \Delta y$$

Du fait de la symétrie circulaire, on peut se placer dans le repère de la facette qui annule y_0. Après développement :

$$\mathbf{e_{photo}} = \frac{3}{p^4}((4p^2 - 5d^2)\Delta^2 y - p^2 \Delta^2 x)$$

L'erreur commise sur un disque de diamètre c se majore, à partir de la formule précédente, par $3(\frac{c}{p})^2$, où p est l'éloignement du centre de la facette à la source lumineuse. Pour que l'erreur soit inférieure à un niveau de gris sur une image 8 bits, il faut que c soit 30 fois plus petit que p. L'approximation au premier ordre des gradients lumineux est donc justifiée dans nos expériences, car les facettes dessinées sur la mire ont un côté de 5 cm avec une source placée à 1.5 m.

En conséquence, l'intensité de la région 'image' d'une surface lambertienne varie linéairement en fonction des coordonnées des pixels. Le plan qui approxime le mieux la distribution d'intensité de ces régions est calculé par une méthode aux moindres carrés. Dans le cas où une surface claire est observée sur un fond sombre, le calcul est réalisé sur le fond $I_{fond}(u, v)$ et sur la région claire $I_{forme}(u, v)$. Ces deux équations suffisent pour normaliser l'image entre 0 et 1 pour chaque pixel (u, v) :

$$\frac{I(u,v) - I_{fond}(u,v)}{I_{forme}(u,v) - I_{fond}(u,v)}$$

La valeur 0 indique que le pixel de l'image normalisée ne voit pas la surface claire, alors que la valeur 1 indique au contraire qu'il s'intersecte en totalité avec l'image de la surface claire. On obtient une image dont la valeur d'un pixel donne le pourcentage de son aire couverte par la projection de la surface claire. L'image normalisée est donc celle qui serait observée si l'éclairage était à l'infini.

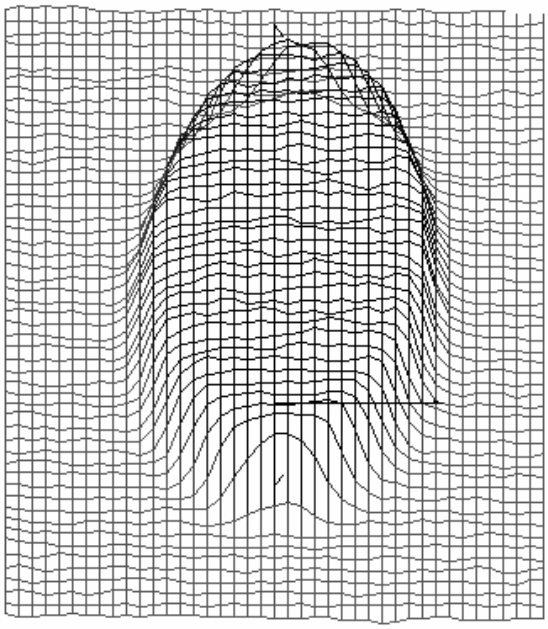


Fig. 3– *Distribution réelle de l'intensité d'un disque blanc sur fond noir.*

Sur la figure 3.2., la rupture d'intensité au niveau du contour n'est pas brutale, mais au contraire assez douce. Ce lissage est provoqué par différents phénomènes physiques : la réponse électronique du capteur, l'effet de profondeur, les aberrations et la diffusion de l'optique, les interactions électronique entre pixels. En particulier, avec notre caméra, un étalement de trois pixels de chaque côté du contour se produit.

3.3. Mesures des moments géométriques

La réponse d'une cellule CCD est proportionnelle à l'intensité reçue sur sa surface. En conséquence, pour une distribution d'intensité constante λ, la contribution au moment d'ordre (p, q) du pixel (u, v), où u et v sont les coordonnées

de son coin supérieur gauche, est:

$$M_{p,q}^{pixel} = \lambda \int_{aire\ pixel} x^p y^q \, dx \, dy = \lambda(\frac{(u+1)^{p+1} - u^{p+1}}{p+1})(\frac{(v+1)^{q+1} - v^{q+1}}{q+1})$$

En pondérant le moment de chaque pixel (u, v) par la valeur $\lambda_{u,v}$ de l'image normalisée, les caractéristiques géométriques de la forme se calculent avec précision. Le moment d'ordre (p, q) est alors donné par l'équation suivante :

$$M_{p,q} = \sum_{u,v} \lambda_{u,v} (\frac{(u+1)^{p+1} - u^{p+1}}{p+1})(\frac{(v+1)^{q+1} - v^{q+1}}{q+1}) \tag{7}$$

L'aire A, le barycentre G et la matrice inertie surfacique I de la région 'image' sont alors respectivement donnés par :

$$A = \sum_{u,v} \lambda_{u,v}$$

$$G = \frac{1}{A} \sum_{u,v} \lambda_{u,v} \begin{pmatrix} u + \frac{1}{2} \\ v + \frac{1}{2} \end{pmatrix}$$

$$I = \frac{1}{A} \sum_{u,v} \lambda_{u,v} \begin{pmatrix} u^2 + u + \frac{1}{3} & (u + \frac{1}{2})(v + \frac{1}{2}) \\ (u + \frac{1}{2})(v + \frac{1}{2}) & v^2 + v + \frac{1}{3} \end{pmatrix}$$

Pour estimer les caractéristiques géométriques, il suffit de calculer une somme pondérée avec les niveaux de gris normalisés. L'obtention de ceux-ci ne présente pas non plus de difficulté sur une image de mire. En effet, le calcul et la correction des gradients d'intensité sont simples quand on a pu délimiter avec certitude, grâce à un détecteur de contours [Der87], une zone 'intérieure' à chaque région et une 'extérieure' voisine. Avec notre caméra, une bande de 3 pixels de chaque côté du contour est exclue.

3.4. Précision du calcul des moments

L'analyse des erreurs sur les moments doit tenir compte de l'erreur $\mathbf{e_{photo}}$, conséquence de l'approximation photométrique. Mais il y a d'autres causes d'erreurs : le bruit de l'image ($\mathbf{e_{bruit}}$), la quantification de l'intensité des pixels ($\mathbf{e_{digit}}$) et l'indétermination de localisation provoquée par la discrétisation spatiale ($\mathbf{e_{local}}$). Ces sources d'erreurs sont supposées indépendantes.

$\mathbf{e_{bruit}}$: Le bruit en intensité présent sur chaque pixel est supposé gaussien centré de variance σ^2. Alors, l'écart type de l'erreur produite par ce bruit pour l'aire A est $\sqrt{s}\sigma$, où s est le nombre de pixels de la région. Pour les composantes du barycentre G, l'erreur est approximativement $\sigma\sqrt{\frac{M_{2,0}}{A^2}}$ et $\sigma\sqrt{\frac{M_{0,2}}{A^2}}$. Celle-ci, sur les éléments $I_{u,u}$, $I_{u,v}$ et $I_{v,v}$ de la matrice d'inertie, a une forme analogue en remplaçant les moments par $M_{4,0}$, $M_{2,2}$ et $M_{0,4}$ respectivement. Une forme aplatie provoque donc des erreurs non équivalentes sur chaque composante. Avec un disque de rayon R, l'erreur sur le barycentre est inférieure à σ, alors que, pour l'inertie, elle est majorée par σR. En conséquence, un bruit dont l'écart type est de quelques niveaux de gris produit une erreur inférieure à 0.01 pixel sur le barycentre, pour des images 8 bits très contrastées.

$\mathbf{e_{photo}}$: Dans nos expériences, comme cela est décrit précédemment, l'erreur photométrique est inférieure à un niveau de gris sur des images contrastées. Elle est beaucoup plus petite que celle provoquée par le bruit.

$\mathbf{e_{digit}}$: L'erreur de quantification de l'intensité peut être décrite comme une variable aléatoire uniforme centrée de variance $\sigma^2 = \frac{1}{12g^2}$ sur l'intensité de chaque pixel, où g est la hauteur du saut d'intensité entre la région claire et la région sombre [Der87]. Pour des images 8 bits, avec un contraste maximum ($g = 256$), l'écart type est de l'ordre de 10^{-3}. Comme pour $\mathbf{e_{photo}}$, $\mathbf{e_{digit}}$ est petite.

$\mathbf{e_{local}}$: L'erreur provoquée par la discrétisation spatiale est nulle sur A, puisque l'aire n'est pas localisée. Dans le cas d'une dimension, nous montrons que [Tar94], indépendamment de la réponse de la caméra, l'erreur de résolution est de $\frac{1}{8A}$ sur la position du barycentre, et de $\frac{1}{4}$ sur I. À partir de ces résultats, un passage à la dimension deux peut s'effectuer. Soit n_c le nombre de pixels du contour de la région considérée, la moitié des pixels du contour peuvent biaiser la mesure dans le même sens, donc l'erreur $\mathbf{e_{local}}$ est majorée par $\frac{n_c}{16A}$. La forme qui minimise le rapport de la longueur du contour avec l'aire est un disque. Soit R son rayon, l'erreur $\mathbf{e_{local}}$ est alors majorée par $\frac{1}{8R}$. Il suffit donc d'un cercle de rayon 13 pixels pour avoir une précision de l'ordre du centième de pixel sur la position du barycentre.

L'erreur totale théorique obtenue est donc, pour un cercle, inférieure à $5/100^{ème}$ de pixel sur le barycentre, si le bruit est inférieur à 10 niveaux de gris. Une aussi grande précision ne peut pas être obtenue uniquement par la moyenne du contour de la région. En pratique, il est donc possible d'atteindre une erreur inférieure à $1/10^{ème}$ de pixel, sur des images réelles de formes arrondies et compactes, si l'environnement est bien conditionné.

3.5. Conclusion

Il est possible de calculer les caractéristiques géométriques d'une forme plane dans une image avec une bonne précision, si elle est la projection d'une facette lambertienne de taille petite par rapport à la distance de la source lumineuse, et si sa taille est supérieure à une vingtaine de pixels. La forme qui optimise l'erreur de localisation sur le barycentre est une forme arrondie. La méthode proposée a deux propriétés remarquables : l'estimation est assez robuste au bruit et au lissage par un filtre symétrique quelconque de l'image (en particulier pour les coordonnées du barycentre). Enfin, la méthode est relativement simple à mettre en place.

4. Transformation perspective d'une ellipse

Les caractéristiques géométriques d'une région de l'image peuvent être mesurées avec précision. Mais pour pouvoir utiliser ces données dans la calibration, il faut être capable de les relier à la description 3D de la surface observée. L'ellipse présente un grand intérêt de part ses nombreuses propriétés vis-à-vis de la transformation perspective (ou transformation conique) [Her92]. Cependant, comment une ellipse se transforme-t-elle par le modèle sténopé ? La caméra doit voir l'ellipse entièrement et son image rester elliptique. De fait, les conditions de la calibration permettent ces restrictions, mais la généralisation de ce qui suit aux coniques est aisée. À partir des différentes représentations d'une ellipse dans le plan, le lien entre l'équation d'une ellipse et ses moments géométriques d'ordre inférieur ou égal à deux sera établi. Ces équations permettent de faire la calibration, à partir du modèle de la mire et des données 2D extraites des images.

4.1. Caractérisations d'une ellipse

Une ellipse peut être caractérisée dans le plan par son centre C et les vecteurs demi-grand axe A_1 et demi-petit axe A_2 donnés dans le repère orthonormal R (figure 4.1.). Les points de l'ellipse sont alors décrits par :

$$X = C + \cos(\theta)A_1 + \sin(\theta)A_2 \qquad \theta \in [0, 2\pi]$$

Cette équation se traduit en coordonnées homogènes par la relation suivante :

$$X = X_g + \cos(\theta)R_1 + \sin(\theta)R_2 \qquad \theta \in [0, 2\pi]$$

où $X_g = \begin{pmatrix} C \\ 1 \end{pmatrix}$, $R_1 = \begin{pmatrix} A_1 \\ 0 \end{pmatrix}$ et $R_2 = \begin{pmatrix} A_2 \\ 0 \end{pmatrix}$. Soit B la matrice 3×3 constituée, dans l'ordre, des vecteurs R_1, R_2 et X_g. Cette matrice est une première caractérisation de l'ellipse dans le repère R.

L'équation dans un repère donné représente autrement l'ellipse dans le plan. L'équation implicite de celle-ci s'écrit en coordonnées homogènes sous la forme suivante :

$$X^t E X = 0$$

E est une matrice symétrique 3×3 de signature $(\epsilon, \epsilon, -\epsilon)$ avec $\epsilon = \pm 1$. La matrice E dans le repère R est une autre caractérisation de l'ellipse dans le plan. Cette matrice est définie à un facteur près.

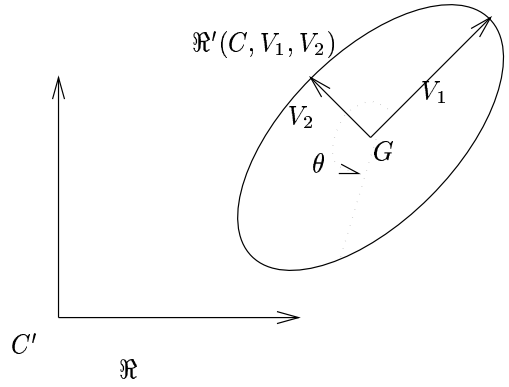


Fig. 4– *Repère absolu $R(O, i, j)$ et canonique $R'(C, A_1, A_2)$ d'une ellipse.*

Il existe un lien entre la représentation par B et celle par E. Pour l'expliciter, on se place dans le repère orthogonal R' d'origine C et de vecteurs A_1 et A_2 (figure 4.1.). Dans ce repère, l'équation implicite est réduite à son expression canonique :

$$X'^t E' X' = X'^t \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} X' = 0 \qquad (8)$$

La matrice E' associée est alors particulièrement simple.

La transformation entre les deux repères R et R' s'explicite en fonction de la matrice B. Ainsi, un point de coordonnées X' dans le repère R' a pour nouvelles coordonnées $X = BX'$ dans le repère R. En réalisant le changement de repère dans l'équation implicite réduite (8), la relation entre la matrice E et B se déduit, par identification :

$$E = B^{t^{-1}} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} B^{-1}$$

Ce qui peut s'écrire, après inversion, puisque E est de rang 3 :

$$E^{-1} = B \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} B^t \qquad (9)$$

Alors, l'inverse de la matrice de la forme quadratique se décompose sous la forme suivante :

$$E^{-1} = \begin{pmatrix} CC^t - A_1A_1^t - A_2A_2^t & C \\ C^t & 1 \end{pmatrix} \qquad (10)$$

On remarque que la dernière colonne de E^{-1} (ou la dernière ligne), quelque soit le repère choisi, est égale, à un facteur près, au vecteur homogène X_g qui donne la position du centre de l'ellipse dans ce repère.

4.2. Matrice d'inertie homogène

La matrice d'inertie homogène du disque elliptique est définie par :

$$H = \int_{disque\ elliptique} XX^t dS$$

Pour un disque de rayon unité dans le repère barycentrique, le calcul donne :

$$H' = \begin{pmatrix} \frac{\pi}{4} & 0 & 0 \\ 0 & \frac{\pi}{4} & 0 \\ 0 & 0 & \pi \end{pmatrix}$$

Comme précédemment, en passant de R' au repère R, on en déduit que la matrice d'inertie homogène de l'ellipse, dans le cas général, s'écrit :

$$H = det(B)BH'B^t$$

Cette relation est vraie uniquement si la transformation B est affine. Après un développement semblable à celui de (10), on obtient l'expression suivante :

$$H = \pi det(A_1\ A_2) \begin{pmatrix} \frac{1}{4}(A_1A_1^t + A_2A_2^t) + CC^t & C \\ C^t & 1 \end{pmatrix}$$

L'aire de l'ellipse A s'avère bien être $\pi det(A_1\ A_2)$. La relation entre la matrice d'inertie homogène H et E se déduit, par comparaison avec (10) :

$$\frac{H}{A} = \frac{1}{4}(5X_gX_g^t - E^{-1}) \qquad (11)$$

Pour cette relation, E est supposée normalisée, c'est-à-dire que le dernier élément de la matrice est 1.

En conséquence, il suffit d'inverser la matrice de l'équation implicite homogène pour connaître la matrice d'inertie homogène de cette ellipse. Enfin, la seule connaissance des caractéristiques géométriques G et I permet de déterminer complètement l'ellipse. Une ellipse est donc caractérisée par ses seuls moments d'ordre inférieur ou égal à deux.

4.3. projection d'une ellipse

Les caractéristiques d'une ellipse E_{image}, qui est la projection perspective d'une ellipse E_{espace}, peuvent être obtenues à partir des formules précédentes. Soient R_1, R_2 et X_g les trois vecteurs (demi-grand axe, demi-petit axe et centre) qui caractérisent l'ellipse E_{espace} dans le repère absolu de l'espace. Soit $P_{ellipse}$ le plan qui contient l'ellipse, et P un point de $P_{ellipse}$ de coordonnées $(x, y, 1)$ dans le repère orthogonal où l'équation de l'ellipse E_{espace} est canonique (8). Les coordonnées de P dans le repère de référence de l'espace sont :

$$(R_1 \quad R_2 \quad X_g) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

L'image de ce point est donc, d'après (2) :

$$U = M(R_1 \quad R_2 \quad X_g) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = MB \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

La relation entre le plan de l'ellipse $P_{ellipse}$ et le plan 'image' est donc une transformation homogène linéaire. En conséquence, puisque (9) reste vraie pour une transformation homogène quelconque, la matrice E de la forme quadratique de l'ellipse E_{image} a pour inverse :

$$E^{-1} = MB \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} (MB)^t \qquad (12)$$

À partir de (12), grâce à la remarque faite sur (10), nous pouvons déduire le centre de l'ellipse E_{image} en fonction des éléments de la matrice MB, c'est-à-dire selon les caractéristiques (R_1, R_2, X_g) de l'ellipse E_{espace} et selon les paramètres (L_i) de la transformation perspective M. Les coordonnées homogènes $U_g = (s_gu_g\ s_gv_g\ s_g)^t$ du barycentre sont donc données, après calcul, par :

$$\left.\begin{array}{rcl} s_gu_g & = & R_1 \cdot L_1R_1 \cdot L_3 + R_2 \cdot L_1R_2 \cdot L_3 - X_g \cdot L_1X_g \cdot L_3 \\ s_gv_g & = & R_1 \cdot L_2R_1 \cdot L_3 + R_2 \cdot L_2R_2 \cdot L_3 - X_g \cdot L_2X_g \cdot L_3 \\ s_g & = & (R_1 \cdot L_3)^2 + (R_2 \cdot L_3)^2 - (X_g \cdot L_3)^2 \end{array}\right\} \qquad (13)$$

Une fois que le centre $C = (u_g\ v_g\ 1)^t$ de l'ellipse 'image' est connu, la matrice d'inertie I s'obtient en fonction de M et B grâce à (11) :

$$I = \frac{1}{4}(5CC^t - \frac{1}{s_g} \begin{pmatrix} e_{uu} & e_{uv} \\ e_{uv} & e_{vv} \end{pmatrix}) \qquad (14)$$

avec

$$\begin{array}{rcl} e_{uu} & = & (R_1 \cdot L_1)^2 + (R_2 \cdot L_1)^2 - (X_g \cdot L_1)^2 \\ e_{vu} & = & R_1 \cdot L_1R_1 \cdot L_2 + R_2 \cdot L_1R_2 \cdot L_2 - X_g \cdot L_1X_g \cdot L_2 \\ e_{vv} & = & (R_1 \cdot L_2)^2 + (R_2 \cdot L_2)^2 - (X_g \cdot L_2)^2 \end{array}$$

4.4. Conclusion

Une ellipse est caractérisable par ses moments géométriques d'ordre inférieur ou égal à deux, et le lien entre sa description dans le plan et ses moments est relativement simple. Par l'utilisation des coordonnées homogènes, cette relation a été généralisée à la transformation perspective d'une ellipse de l'espace. Alors, le barycentre et les termes d'inertie s'expriment comme le rapport de polynômes de degré deux en fonction des caractéristiques de l'ellipse 3D et de la transformation perspective.

5. Résolution

Une méthode précise de calcul des moments d'une forme dans une image a été élaborée. Quand la forme est un disque elliptique, les relations entre ces moments et l'ellipse 3D ont été explicitées. Ces relations qui lient les données mesurées sur l'image et les données tridimensionnelles de la mire permettent l'estimation des paramètres de la calibration, si la mire est supposée de géométrie parfaitement connue.

5.1. Critère d'adéquation

Le critère d'adéquation du modèle avec les données est choisi du type moindres carrés. Si les données sont des points U_i^r dans l'image, le critère est basé sur la distance euclidienne. Ainsi, si l'on désire minimiser la distance en pixel, le critère à choisir est :

$$\min_{paramètres} \sum_i (u_i^r - u_i)^2 + (v_i^r - v_i)^2$$

Par contre, pour l'erreur réelle sur la rétine, par exemple en millimètres, le critère à minimiser est :

$$\min_{paramètres} \sum_i \left(\frac{u_i^r - u_i}{a_u}\right)^2 + \left(\frac{v_i^r - v_i}{a_v}\right)^2$$

Ce critère a donné de meilleurs résultats dans nos expérimentations. U_i s'exprime en fonction des paramètres de la calibration et des caractéristiques de la mire. Dans le cas de la correspondance point à point, U_i est donné par l'équation (2). Si les données mesurées sont les centres des ellipses images, l'utilisation des équations (13) s'avère alors nécessaire.

Il est possible de généraliser ce qui précède à un modèle de caméra autre que sténopé. En effet, du fait de la faiblesse des distorsions géométriques, l'image de l'ellipse peut être considérée comme uniformément distordue. Il suffit alors, pour avoir le lien entre une ellipse 3D et le centre de son image, de composer les formules de distorsions (4) (5) ou (6) avec (13).

Pour les termes de la matrice d'inertie, le critère d'adéquation est la distance des trois termes indépendants de la matrice. Avec le modèle sténopé les termes d'inertie fournissent trois contraintes supplémentaires par ellipse (14). Mais il semble difficile de les généraliser au cas avec des distorsions.

5.2. Méthode itérative

Le problème de la calibration d'une caméra peut s'exprimer, sous forme minimale, comme un problème non linéaire de moindres carrés sans contrainte. Pour cela, la rotation doit être représentée sous une forme minimale. Les paramètres (p_1, p_2, p_3) choisis sont ceux d'Euler-Rodrigues, car ils permettent d'écrire les coordonnées de la projection d'un point comme le rapport de deux polynômes de degré trois.

L'algorithme utilisé pour minimiser le critère est dit de Gauss-Newton [DS83]. Il nécessite le calcul du vecteur des dérivées partielles du critère par rapport aux paramètres de la calibration, ce qui peut être facilement fait avec un calculateur formel. Les valeurs des paramètres initiaux sont obtenues par la méthode de Faugeras et Toscani [FT86]. Ceci assure une convergence rapide de l'algorithme.

6. Résultats

La vérification expérimentale d'une méthode de calibration est difficile, car elle demande de mesurer des distances et des orientations avec une précision importante. Une première alternative à cette difficulté est de créer des images de synthèse d'une scène de calibration artificielle où la géométrie et la photométrie sont reproduites. Une autre possibilité est de vérifier la stabilité de l'estimation des paramètres intrinsèques.

6.1. Images de synthèses

6.1..1 Précision sur les centres des ellipses

Avec une scène de synthèse, la maîtrise du processus permet de connaître les erreurs faites, notamment sur la mesures des données dans l'image. Pour cela, les mesures sont

comparées avec les données obtenues par calcul. Dans le tableau suivant, les erreurs de mesures, sur les barycentres de 280 ellipses de diamètre 40 pixels environ, sont calculées :

erreur	contours	intensité
maximum	1.2 pixel	0.07 pixel
moyenne	0.3 pixel	0.03 pixel
écart type	0.3 pixel	0.02 pixel

Ce tableau montre la supériorité de la méthode fondée sur l'intensité des pixels, sur la technique qui utilise les contours (rapport dix en précision). L'ordre de grandeur des erreurs ($5/100^{ème}$) est respecté par rapport au calcul théorique d'erreur (partie 3).

6.1..2 robustesse au bruit du centre

La méthode de calcul du barycentre fait une moyenne sur un grand nombre de pixels. Elle doit donc être assez résistante au bruit de l'image. La robustesse a été testée et vérifiée sur des images de disques de différents diamètres, bruitées avec un bruit gaussien centré de variance σ^2. Le tableau suivant donne l'écart type de l'erreur provoquée uniquement par le bruit, pour deux diamètres sur une centaine de disques.

diamètre	$\sigma = 2$	$\sigma = 4$	$\sigma = 8$	$\sigma = 16$
10 pixels	0.004	0.01	0.02	0.03
20 pixels	0.0006	0.001	0.003	0.005

L'erreur sur le barycentre, due au bruit, diminue en fonction du diamètre du disque. Elle est très inférieure à la majoration prévue (partie 3), dès que le rayon est supérieur à une dizaine de pixels.

6.1..3 Corrélation des estimations

Dans la matrice de régression, un coefficient proche de un indique qu'il existe un lien entre deux des paramètres de la calibration. Or, il existe un lien entre les paramètres t_z, a_u et a_v, de même que pour les groupes de paramètres t_x, u_0, p_2 et t_y, v_0, a_v, p_1, p_3. Ces liaisons s'expliquent quand on examine les modifications produites sur l'image de la mire par de petites variations des paramètres de la calibration. Certaines combinaisons de ces variations produisent de très petites modifications de l'image. Par exemple, une faible variation de profondeur peut être approximativement compensée par une dilatation de la rétine, d'où la liaison entre t_z, a_u et a_v. Les deux autres groupes de liens s'expliquent par le fait qu'une petite translation de la mire ou une petite rotation (proche d'une translation) peut être compensée par un déplacement du centre de l'image.

Un développement limité, supposant une erreur uniquement sur u_0 et t_x, montre que les erreurs se compensent à peu près quand l'épaisseur de la mire est petite relativement à sa profondeur z_m, si elles vérifient la relation suivante:

$$\delta_{t_x} = -\frac{z_m \delta_{u_0}}{a_u} \tag{15}$$

Ceci se traduit aussi géométriquement par l'existence de directions privilégiées où le critère varie peu, au voisinage de son minimum. En conséquence, la qualité de l'estimation est directement fonction du rapport de la profondeur de la mire par sa distance à la caméra. Ce rapport doit être le plus grand possible.

6.1..4 Précision de la calibration

Quand la caméra de synthèse est un sténopé parfait, les paramètres sont tous retrouvés avec une erreur relative de moins de 0.07% (voir 6.1..5). Pour initialiser l'algorithme d'estimation, la méthode de Faugeras-Toscani est utilisée. Elle est basée sur la correspondance de points. Avec le type de mire utilisé (figure 6.2.), les points mis en correspondance sont le centre des ellipses 3D et 'image'. Malgré cette approximation, la méthode de Faugeras-Toscani donne une estimation avec une erreur relative maximum de 0.7%. En conséquence, si le maximum de précision n'est pas nécessaire, la correspondance des barycentres peut être supposée. Ces résultats ont été obtenus, en calculant l'écart type sur 6 positions assez différentes de la caméra (figure 6.2.).

Pour se rapprocher de la réalité, une importante distorsion radiale a été ajoutée à la caméra synthétique ($\alpha = 1.0$). Si la calibration est réalisée avec l'hypothèse du modèle sténopé, les résultats sont fortement perturbés. Et le biais augmente avec l'importance des distorsions. Sur 6 positions de la caméra, le centre de l'image (u_0, v_0) est biaisé en moyenne d'une vingtaine de pixels, comme le montre le tableau suivant :

	sténopé	décentré+radial	radial
u_0	24 pixel	-0.3 pixel	0.1 pixel
v_0	27 pixel	0.2 pixel	0.01 pixel
a_u	390 p/mm	1 p/mm	0.46 p/mm
a_v	380 p/mm	0.7 p/mm	0.28 p/mm
t_x	-23 mm	0.4 mm	-0.1 mm
t_y	-30 mm	-0.2 mm	-0.02 mm
t_z	491 mm	1.9 mm	1.3 mm
θ	1.6^o	0.01^o	0.008^o
α		-0.002	-0.0008

Ce tableau donne les biais observés, sur les différents paramètres estimés, pour trois hypothèses de modèles : sténopé, avec décentrage et distorsions radiales (4&5), avec distorsions radiales (5). Le modèle pertinent est le dernier. On note que l'introduction de paramètres non pertinents (paramètres β et γ du décentrage) induit une moins bonne estimation des autres paramètres. De plus une dispersion 10 fois plus grande des résultats est observée.

De même, la méthode point à point avec distorsions donne des résultats biaisés. Le biais est particulièrement sensible sur t_z (voir aussi le paragraphe 6.2..1).

Même pour un objet pris dans de bonnes conditions à une distance de 1 mètre, une erreur d'estimation de 25 pixels du centre de la caméra, comme celle observée en faisant l'hypothèse du modèle sténopé, produit d'après la formule (15) un biais de 2.5 cm dans la connaissance de la position absolue de l'objet. Placé à une distance de 10 mètres le biais produit atteint 25 cm. De plus, le tableau précédent confirme expérimentalement l'équation (15), puisque les biais sur u_0 et t_x (ou v_0 et t_y) sont proportionnels.

6.1..5 Avec des moments d'inertie

Dans le cas où le modèle sténopé est valide, trois contraintes supplémentaires sont données par les trois moments d'inertie de l'image (14). L'apport éventuel de l'utilisation de ces contraintes a été testé sur 6 images de synthèse de la mire.

	sans inertie	avec inertie
u_0	0.04%	0.04%
v_0	0.04%	0.01%
a_u	0.03%	0.02%
a_v	0.02%	0.02%
t_x	0.07%	0.08%
t_y	0.07%	0.03%
t_z	0.03%	0.02%
θ	0.03%	0.03%

Le tableau précédent montre l'erreur relative faite dans l'estimation des paramètres, avec ou sans l'utilisation des termes d'inertie. Sur certains paramètres, l'estimation est améliorée d'un facteur 2 à 3 avec les moments d'inertie. L'apport n'est pas très important, mais il faut remarquer que les moments d'ordre 2 sont plus sensibles aux perturbations que ceux d'ordre inférieur.

6.2. Images réelles

La mire utilisée est constituée de deux plans métalliques où sont fixés des disques de couleur unie (figure 6.2.). La surface est traitée pour être faiblement spéculaire.

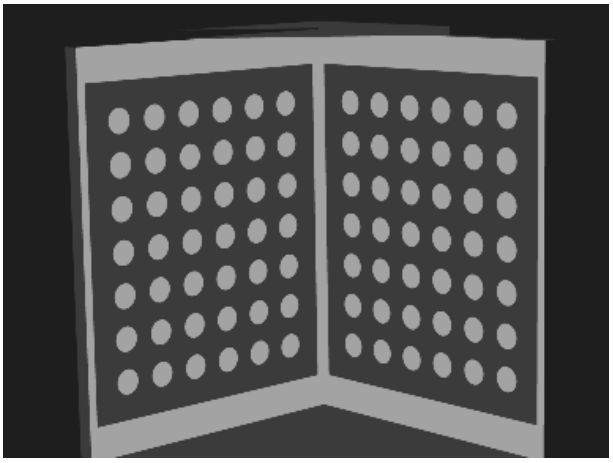


Fig. 5– Image synthétique de la mire.

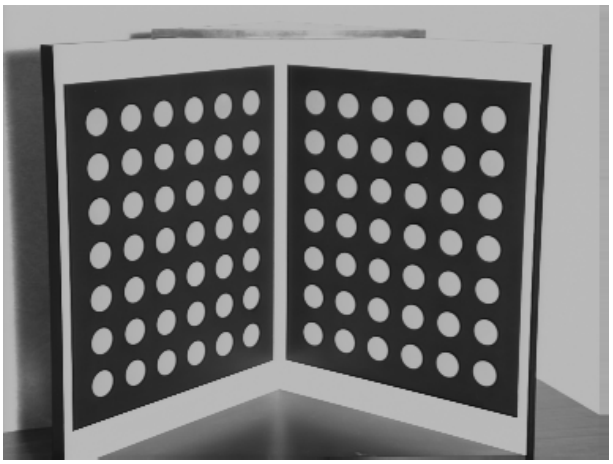


Fig. 6– Image de la mire métallique.

6.2..1 Comparaison des méthodes

Les spécifications d'une caméra réelle ne permettent généralement qu'un calcul approximatif des paramètres exacts. Sans référence, il est difficile de tirer beaucoup d'informations de la comparaison des résultats des différentes méthodes de calibration. Dans le tableau suivant, se trouvent les résultats d'une calibration par trois méthodes différentes: la méthode point à point sans distorsion [FT86], la méthode itérative point à point avec distorsions radiales [LT88], et notre méthode avec des ellipses et distorsions radiales.

	Toscani	points	ellipses
u_0	376.1 pixel	380.2 pixel	379.6 pixel
v_0	261.3 pixel	288.6 pixel	288.0 pixel
a_u	1698.5 p/mm	1810.1 p/mm	1804.1 p/mm
a_v	1648.9 p/mm	1755.4 p/mm	1750.4 p/mm
t_x	-216.5 mm	-219.9 mm	-219.5 mm
t_y	-98.2 mm	-121.9 mm	-121.4 mm
t_z	1428.6 mm	1535.2 mm	1530.6 mm
θ	51.00^o	51.13^o	51.11^o
α		-0.351	-0.348
résidu	0.23 pixel	0.058 pixel	0.056 pixel

Le résidu moyen, entre les mesures 'images' et leur estimées après minimisation, est aussi indiqué. Ce résidu est pour les images réelles d'environ $1/20^{ème}$ de pixel avec les points comme avec les ellipses. L'ordre de grandeur du résidu est cohérent avec les estimations d'erreurs précédentes (voir 6.1..1). En effet, si le modèle de caméra est adéquat, ce résidu doit être inférieur ou égal à l'erreur faite sur les barycentres, suivant l'importance du biais présent dans les mesures.

Dans cet exemple, l'erreur produite en supposant la mise en correspondance des barycentres d'une ellipse et de son image est du même ordre de grandeur que l'erreur de mesure. Pourtant, les méthodes avec ou sans l'utilisation de cette approximation donnent des résultats proches mais distinguables (différence de 5 mm en t_z par exemple). En effet, si la plus grande précision est désirée, on peut faire l'hypothèse de la mise en correspondance des barycentres seulement quand elle produit une erreur négligeable par rapport à l'erreur de mesure.

Sur les images de synthèse, le résidu moyen observé est encore plus petit. Ceci indique que sur les images de synthèse, la méthode de mesure des barycentres donne des résultats avec une précision inférieure à $1/10^{ème}$ de pixel, et explique donc l'obtention de meilleurs calibrations que dans le cas réel.

6.2..2 Stabilité des paramètres intrinsèques

Il est possible de pallier partiellement au manque de référence sur les paramètres réels. En effet, si l'estimation des paramètres intrinsèques n'est pas stable, le modèle de caméra considéré n'est pas pertinent. En effet, quand la mire est déplacée dans le champ de vision de la caméra, sans toucher aux réglages de cette dernière, seuls les paramètres extrinsèques doivent changer d'une estimation à l'autre. Plusieurs modèles ont été testés, suivant le choix et la combinaison des divers types de distorsions (partie 2). Les estimations, avec la modélisation des distorsions radiales, se sont avérées les plus stables. Par exemple, sur une série de 12 images de la mire dans différentes positions, les résultats suivants ont été obtenus :

sténopé	u_0	v_0	a_u	a_v	
moyenne	248.8	250.8	1189	1690	
écart type	7.6	9.8	5.4	7.5	
écart en %	3.1%	3.9%	0.46%	0.45%	
radial	u_0	v_0	a_u	a_v	α
moyenne	252.4	253.5	1276	1812	-0.38
écart type	0.25	0.45	1.9	2.5	0.008
écart en %	0.1%	0.18%	0.15%	0.14%	2.1%

Le choix du modèle est très important. Car, s'il ne possède pas suffisamment de degrés de liberté (sur-contraint), l'estimation risque d'être biaisée. À l'opposé, si le modèle est trop général (sous-contraint), l'estimation des paramètres, pertinents ou non, peut être instable. Ainsi, quand les distorsions radiales et de décentrage sont modélisées, l'estimation du coefficient de décentrage β varie de 35% sur la série des 12 images. Avec notre caméra CCD, une Sony couleur XC-007 et un objectif Canon J15x9.5B, le modèle le plus adéquat parmi la dizaine de modèles testés est celui qui a des distorsions radiales du premier ordre.

7. Conclusion

La méthode de calibration présentée utilise non pas la correspondance entre points mais entre des surfaces caractérisées par leurs moments. Une technique de calcul des moments, fondée sur la photométrie des surfaces, a été développée. Elle est robuste au bruit de l'image et permet de fournir à l'algorithme de calibration des données précises (moins $1/10^{ème}$ de pixel sur le barycentre).

Nous avons explicité les équations qui lient les moments d'ordre inférieur ou égal à deux de l'image d'une ellipse à sa description 3D. Dans nos résultats, quelque soit le modèle, le barycentre des ellipses 'images' peut être utilisé comme donnée pour l'estimation. Les termes d'inertie fournissent trois contraintes supplémentaires dans le cas du modèle sténopé. Il semble donc intéressant d'étudier la généralisation, à un modèle quelconque, de la relation avec les moments d'ordre supérieur à deux. Des types de formes différentes du disque elliptique peuvent aussi être utilisés.

Si le modèle est bien choisi, la méthode de calibration proposée permet une bonne estimation des paramètres de la caméra (la précision est supérieure à 0.1% sur les images de synthèse). Suite à des tests avec notre caméra, le modèle avec distorsions radiales est le plus adéquat. Il permet d'estimer les paramètres de la caméra avec une stabilité d'environ 0.1%.

7. Références

AATRM89 A. Abi-Ayad, B. Thiesse, C. Ragi, and Ph. Marthon. Les multiples facettes du problème de calibrage de caméras. In *7ème congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle*, 1989.

Bey92 H.A. Beyer. Accurate calibration of ccd-cameras. In *IEEE International Conference on Computer Vision*, pages 96–101, Champaign, June 1992.

BMV93 B. Boufama, R. Mohr, and F. Veillon. Euclidean constraints for uncalibrated reconstruction. In *IEEE International Conference on Computer Vision*, pages 466–470, Berlin, March 1993.

CR89 F. Chaumette and P. Rives. Modélisation et calibration d'une caméra. In *7ème congrès AFCET, Reconnaissance des Formes et Intelligence Artificielle*, 1989.

CT90 B. Caprile and V. Torre. Using vanishing points for camera calibration. *International Journal of Computer Vision*, 4:127–140, 1990.

Der87 R. Deriche. Using canny's criteria to derive an optimal edge detector recursively implemented. *International Journal of Computer Vision*, 1(2), 1987.

DS83 J.E. Denis and R.B. Schnabel. *Numerical methods for unconstrained optimization and non linear equations.* Prentice-Hall, Englewood Cliffs, NJ, 1983.

FT86 O.D. Faugeras and G. Toscani. The calibration problem for stereo. In *Proceedings, CVPR '86 (IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami Beach, FL, June 22–26, 1986)*, IEEE Publ.86CH2290-5, pages 15–20. IEEE, 1986.

Her92 I. Herman. *The Use of Projective Geometry in Computer Graphics.* Springer-Verlag, 1992.

HR92 M. Han and S. Rhee. Camera calibration for three-dimensional measurement. *Pattern Recognition*, 25(2):155–164, 1992.

LT88 R.K. Lenz and R.Y. Tsai. Techniques for calibration of the scale factor and image center for high accuracy 3-d machine vision metrology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5):713–720, 1988.

Ma93 S. De Ma. Conics-based stereo, motion estimation and pose determination. *International Journal of Computer Vision*, 10(1), 1993.

Pér93 J.Ph. Pérez. *Optique géométrique matricielle et ondulatoire.* Masson, 1993.

PHYP93 T.Q. Phong, R. Horaud, A. Yassine, and D.T. Pham. Optimal estimation of object pose from a single perspective view. In *IEEE International Conference on Computer Vision*, pages 439–446, Berlin, March 1993.

Rag89 C. Ragi. *N.O.I.S.I.V un système pour l'expérimentation en vision passive tridimensionnelle par ordinateur.* PhD thesis, Institut national polytechnique de Toulouse, 1989.

RZMF92 C.A. Rothwell, A. Zisserman, C.I. Marinos, and D.A. Forsyth. Relative motion and pose from arbitrary plane curves. *Image and Vision Computing*, 10(4):250–262, 1992.

Tar94 J.P. Tarel. Calibration de caméra fondée sur les ellipses. Rapport de recherche 2200, INRIA, 1994.

Tsa86 R.Y. Tsai. An efficient and accurate camera calibration technique for 3d machine vision. In *Conference on Computer Vision and Pattern Recognition*, pages 364–374, Miami Beach, USA, June 1986.

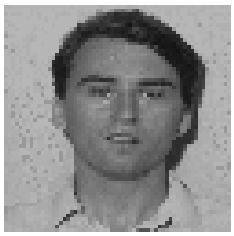
Vai90 R. Vaillant. *Géométrie différentielle et vision par ordinateur: détection et reconstruction des contours d'occultation de la surface d'un objet non-polyédrique.* PhD thesis, Université Paris-Sud, 1990.

WCH90 J. Weng, P. Cohen, and M. Herniou. Calibration of stereo cameras using a non linear distortion model. In *Proceedings, Tenth International Conference on Pattern Recognition*, pages 246–253, Atlantic City, USA, June 1990.

Manuscrit reçu le 6 septembre 1994.

LES AUTEURS

J.-P. TAREL



Jean-Philippe Tarel est un ancien élève de l'École Nationale des Ponts et Chaussées (promotion 91). A l'heure actuelle, il prépare sa thèse au sein du Projet synthèse et analyse d'images de l'Institut National de Recherche en Informatique et Automatique (INRIA). Il s'intéresse particulièrement à la calibration de caméra, à la reconstruction et au recalage 3D sur des paires stéréoscopiques.

A. GAGALOWICZ

André Gagalowicz est directeur de recherches à l'INRIA-Rocquencourt où il dirige le projet Syntim. Ingénieur de l'École Supérieure d'Électricité et docteur d'état, il est actuellement intéressé par les problèmes de vision par ordinateur, d'infographie et la mise en œuvre de schémas coopératifs entre ces deux domaines. Il enseigne ces deux disciplines à l'université d'Orsay, à l'École Nationale Supérieure des Techniques Avancées et à l'Université de Paris VI.