

Utilisation d'un objectif à focale variable en vision monoculaire en vue de la reconstruction 3D

Using a Zoom Lens in Monocular Vision For 3D Reconstruction



Jean-Marc LAVEST

Équipe Vision Artificielle
Laboratoire d'Électronique
Université Blaise Pascal
63177 Aubière Cedex
email lavest@le-eva.univ-bpclermont.fr

Jean-Marc Lavest est né à Aurillac, le 12 août 1964. Il obtient le diplôme d'Ingénieur en Électronique de l'Institut des Sciences pour l'Ingénieur CUST de Clermont-Ferrand, en 1988, et le Diplôme de Docteur de l'Université Blaise Pascal de Clermont-Ferrand en 1992. Actuellement boursier post-doctoral dans l'équipe de recherche en Vision Artificielle du Laboratoire d'Électronique de l'Université Blaise Pascal de Clermont-Ferrand, ses principaux sujets de recherche regroupent la reconnaissance et la modélisation d'objets tridimensionnels.



Gérard RIVES

Équipe Vision Artificielle
Laboratoire d'Électronique
Université Blaise Pascal
63177 Aubière Cedex
email rives@le-eva.univ-bpclermont.fr

Gérard Rives est né à Clermont-Ferrand, le 4 janvier 1946. Docteur de Troisième Cycle d'Électronique de l'Université de Clermont-Ferrand en 1972, il est actuellement Maître de Conférence en Physique à l'Institut Universitaire de Technologie de l'Université d'Auvergne. En 1978, il rejoint l'équipe de recherche de Vision Artificielle du Laboratoire d'Électronique de l'Université Blaise Pascal. Ses principaux domaines de recherche sont la reconnaissance de formes et le traitement d'images.



Michel DHOME

Équipe Vision Artificielle
Laboratoire d'Électronique
Université Blaise Pascal
63177 Aubière Cedex
email dhome@le-eva.univ-bpclermont.fr

Michel Dhome est né à Vichy, le 2 février 1957. Il reçoit le diplôme d'Ingénieur de l'École Nationale des Mines et Techniques Industrielles d'Alès, en 1981 et le diplôme de Docteur-Ingénieur de l'Université Blaise Pascal de Clermont-Ferrand en 1984. De 1984 à 1985, il effectue une année en tant que boursier post-doctoral du gouvernement Canadien au Conseil National de la Recherche, à Ottawa. Il est actuellement Chargé de Recherche au CNRS, responsable du groupe de recherche Vision Artificielle du Laboratoire d'Électronique de l'Université Blaise Pascal. Habilité à diriger les recherches depuis 1992, ses principaux travaux regroupent la vision artificielle, la localisation tridimensionnelle d'objets et la reconstruction.

RÉSUMÉ

L'utilisation d'un zoom en vision monoculaire, en vue de la reconstruction tridimensionnelle d'un univers observé, nous paraît une idée intéressante. En effet les propriétés optiques d'un objectif à focale variable permettent de déterminer, facilement et avec précision, les coordonnées de l'intersec-

tion de l'axe optique et du plan image, mais également de concevoir une reconstruction par triangulation. Dans cet article, nous montrons que l'approximation de cet objectif par le modèle sténopé nécessite de ne plus considérer la distance entre l'objet et le plan image comme un paramètre

fixe. Afin d'appréhender plus globalement les phénomènes optiques mis en jeu par la variation de distance focale, il est alors nécessaire de recourir au modèle optique épais.

Des résultats expérimentaux de reconstruction d'objets polyédriques, provenant d'images réelles, montrent qu'il est possible d'inférer des informations tridimensionnelles relatives à une scène statique, sans déplacement de la caméra dès lors que l'on utilise un objectif à focale variable.

MOTS CLÉS

Vision Monoculaire, Objectif à Focale Variable, Zoom, Points de Fuite, Modèle Sténopé, Modèle Épais, Calibrage d'un Zoom, Triangulation.

ABSTRACT

Using a zoom lens in monocular vision seems to be an interesting idea. The principal aim of this article is to prove that it is possible to infer 3D information from a set of images taken with a zoom lens. In the first part of this article, a precise study of the optical properties of such a lens gives the two following major results :

- *The intersection between the optical axis and the image plane can be independently and accurately determined,*
- *the pin-hole model can be used for the zoom lens only if the object-to-image distance is not considered as fixed.*

Then, to explain the optical phenomena occurring during a focal length change, we show that a thick optical model must be considered.

Finally, experimental results of reconstruction, from a set of real images, are given.

KEY WORDS

Monocular Vision, Focal Change Lens, Zoom, Vanishing Point, Pin-Hole Model, Thick Lens Optical Model, Zoom Calibration, Triangulation.

Remerciements

Les auteurs remercient tous les re-lecteurs qui ont permis, par l'ensemble de leurs critiques et remarques, d'améliorer la rédaction de cet article.

Introduction

Un des problèmes majeurs, en vision artificielle, est d'inférer des informations tridimensionnelles à partir de l'analyse d'images de luminance. Dans le domaine de la vision pour la robotique, la technique couramment utilisée est basée sur la mise en correspondance des projetées d'une même entité spatiale (point, droite, courbe gauche, ...) extraites dans différentes images et la reconstruction de cette primitive 3D par triangulation. Ces images sont acquises soit par des systèmes binoculaires, voire trinoculaires, soit par déplacement d'un système de vision monoculaire.

Le but de cet article est de montrer que l'utilisation d'un objectif à focale variable (zoom), monté sur une caméra fixe, permet, également par triangulation, d'obtenir des données tridimensionnelles relatives à la scène observée supposée fixe. De plus la modélisation de cet objectif complexe est abordée.

Le zoomage ou zooming peut être assimilé à un déplacement virtuel de la caméra selon une trajectoire particulière. En effet le changement de focale se traduit, si on considère le modèle sténopé, par une translation δf du centre optique le long de l'axe optique et donc par une modification de la distance entre l'objet et ce centre. D'après la figure 1 [9], on peut noter qu'à une augmentation de focale doit correspondre un rapprochement de l'objet.

Pour utiliser un objectif de façon précise il faut, avant toute chose, réaliser une étape de calibrage afin de déterminer

avec exactitude les paramètres intrinsèques du système de prise de vue, dont le modèle couramment utilisé sera rappelé dans la première partie. Dans le cas d'un zoom, cette étape doit être réalisée pour les différents réglages de focale et de mise au point.

Nous démontrons dans la deuxième partie qu'il est possible, à partir de l'analyse des propriétés des points de fuite en projection perspective, de déterminer de façon simple et robuste l'intersection de l'axe optique, supposé fixe, avec le plan image. Ce point particulier (u_0, v_0) fait partie des paramètres intrinsèques de la caméra.

Une méthode de calibrage originale sera développée dans la troisième partie. Sa mise en œuvre, sur des images réelles, montrera que la modélisation du changement de distance focale par le modèle sténopé, telle qu'elle est présentée sur la figure 1, ne permet pas d'expliquer les résultats obtenus.

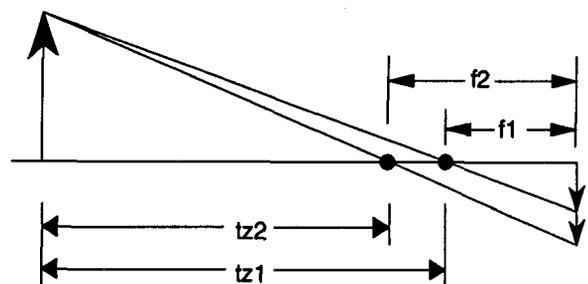


Figure 1. — Modèle sténopé pour deux focales différentes.

Pour appréhender l'ensemble des phénomènes optiques, il sera nécessaire de modéliser le zoom par un système optique épais, présenté dans la quatrième partie, utilisant la notion des plans dits principaux ou unitaires.

Un processus de triangulation, prenant en compte les déplacements virtuels induits par l'utilisation du zoom,

sera développé en dernière partie. Sa mise en œuvre montrera que la reconstruction d'un univers, ici polyédrique, à partir d'un zoom est alors envisageable.

1. Modélisation d'un système de prise de vue

Trouver le modèle d'un système de prise de vue consiste à évaluer la transformation permettant de passer des coordonnées 3D de la scène observée définies dans un référentiel quelconque $R_m : (O_m, X_m, Y_m, Z_m)$, aux coordonnées 2D liées au référentiel de l'image $R_n : (O_n, u_n, v_n)$.

Ce calcul se décompose en deux phases :

— la première revient à estimer le changement de repère permettant d'exprimer les données, formulées dans le repère R_m , dans un repère lié à la caméra $R_c : (O_c, X_c, Y_c, Z_c)$. L'origine de ce dernier est le centre optique de la caméra O_c , son axe z est perpendiculaire au plan image et confondu avec l'axe optique, ses axes x et y sont respectivement parallèles aux lignes et aux colonnes de l'image. Le passage du repère R_m au repère R_c peut être décomposé en une rotation $[R_{\alpha\beta\gamma}]$ suivie d'une translation $[T_{lx, ly, lz}]$ respectivement autour et le long des axes x, y, z .

$$\begin{bmatrix} X_c^i \\ Y_c^i \\ Z_c^i \\ 1 \end{bmatrix} = [R_{\alpha, \beta, \gamma}] \cdot [T_{lx, ly, lz}] \cdot \begin{bmatrix} X_m^i \\ Y_m^i \\ Z_m^i \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} r_{11} & r_{12} & r_{13} & 0 \\ r_{21} & r_{22} & r_{23} & 0 \\ r_{31} & r_{32} & r_{33} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & lx \\ 0 & 1 & 0 & ly \\ 0 & 0 & 1 & lz \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_m^i \\ Y_m^i \\ Z_m^i \\ 1 \end{bmatrix}$$

— la deuxième phase consiste à modéliser la projection perspective $[F]$ et les caractéristiques de l'échantillonnage de la caméra $[J]$ pour estimer la transformation permettant de passer du repère R_c au repère R_n lié au plan image.

$$\begin{bmatrix} su_n^i \\ sv_n^i \\ s \end{bmatrix} = [J] \cdot [F] \cdot \begin{bmatrix} X_c^i \\ Y_c^i \\ Z_c^i \\ 1 \end{bmatrix} =$$

$$= \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c^i \\ Y_c^i \\ Z_c^i \\ 1 \end{bmatrix}$$

Les paramètres dx, dy prennent en compte la forme de la cellule élémentaire de la matrice CCD et la fréquence d'échantillonnage du signal vidéo. On posera $f_x = f/dx$ et $f_y = f/dy$.

La figure 1.1 montre que, contrairement à la représentation du modèle sténopé de la figure 1, on place arbitrairement le

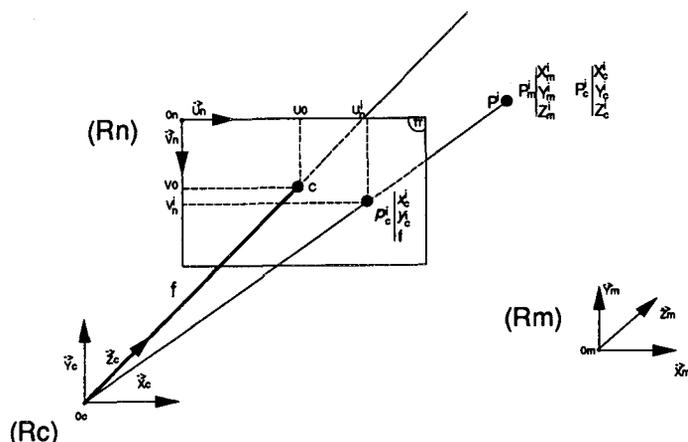


Figure 1.1. — Les différents repères utilisés.

plan image devant le centre optique de manière à obtenir une image non inversée de la scène. Cette convention et le fait que l'origine du repère image se situe en haut et à gauche conduisent à des valeurs positives pour f_x et négatives pour f_y .

Les paramètres u_0 et v_0 sont les coordonnées du centre de l'image C , intersection de l'axe optique et du plan image. Au lieu de calculer de façon globale les paramètres intrinsèques de la caméra, nous allons montrer que l'utilisation du zoom permet de déterminer indépendamment et facilement u_0 et v_0 .

2. Détermination, à l'aide d'un zoom, des coordonnées du centre de l'image

De nombreux auteurs ont étudié les propriétés des points de fuite en projection perspective [13]. Leurs propriétés permettent la détermination des paramètres intrinsèques et extrinsèques d'une caméra. Nous allons démontrer, qu'en utilisant les propriétés des points de fuite et une séquence d'images prises à des distances focales différentes, il est possible de calculer de façon simple et robuste l'intersection de l'axe optique et du plan image.

La détermination des coordonnées du centre de l'image (u_0, v_0) , par une variation de la distance focale a été proposée par Lenz et Tsai [4]. Comme les auteurs ne possédaient pas de zoom adaptable sur leur caméra lors de leur expérimentation, la modification de focale a été obtenue soit par la variation de la mise au point soit par un changement d'objectifs. Dans le premier cas, les valeurs de variation de focale sont très faibles, dans le second, le changement d'objectif en cours de manipulation suppose que la caméra soit solidement fixée et que les axes optiques des deux systèmes soient confondus.

Dans leur conclusion les auteurs soulignent le fait que le centre des lentilles ne se déplace pas précisément le long de l'axe optique quand la distance focale est modifiée. La nature du déplacement observé s'apparente à celui d'une spirale.

L'utilisation d'un objectif à focale variable apporte une nette amélioration et nous verrons au cours de l'interprétation de nos résultats expérimentaux comment prendre en compte les remarques formulées par Lenz et Tsai.

2.1. POINTS DE FUITE

Soit un ensemble S de droites L parallèles dans l'espace (fig. 2.1), définies dans le repère caméra et de vecteur directeur $\vec{V} = (A, B, C)$.

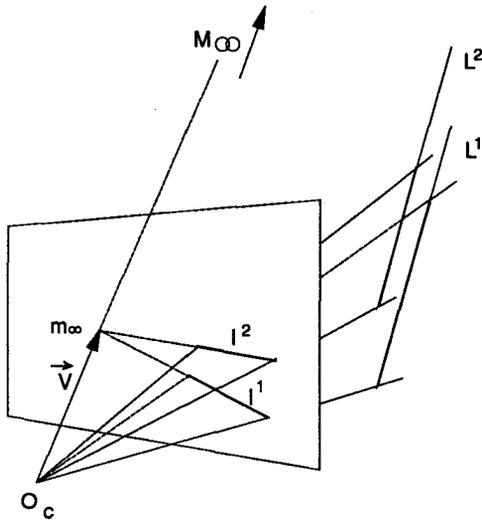


Figure 2.1. — Point de fuite.

Pour chaque droite L de S on peut écrire l'équation paramétrique suivante :

$$\begin{cases} X_c(t) = X_c(0) + A * t \\ Y_c(t) = Y_c(0) + B * t \\ Z_c(t) = Z_c(0) + C * t \end{cases}; \quad t \in [-\infty, +\infty].$$

En projection perspective, le projeté m_∞ du point de fuite M_∞ de L sur le plan image, est :

$$\begin{cases} x_c = \lim_{t \rightarrow \infty} f \frac{X_c(t)}{Z_c(t)} = f \frac{A}{C} \\ y_c = \lim_{t \rightarrow \infty} f \frac{Y_c(t)}{Z_c(t)} = f \frac{B}{C} \\ z_c = f. \end{cases}$$

Il résulte que le vecteur passant par le centre optique et le projeté m_∞ du point de fuite s'écrit :

$$\vec{O_c m_\infty} = \begin{pmatrix} x_c \\ y_c \\ f \end{pmatrix} = \frac{f}{C} \begin{pmatrix} A \\ B \\ C \end{pmatrix}.$$

Ce vecteur est donc parallèle à l'ensemble des droites (L). Ce résultat est exprimé sous la forme de la propriété 1.

Propriété 1 : Le vecteur passant par le centre optique et le projeté du point de fuite d'un ensemble de droites parallèles, est colinéaire au vecteur directeur de ces droites.

Une démonstration analogue peut être faite en utilisant les coordonnées homogènes en projection perspective. En effet, m_∞ représente la projection de M_∞ situé à l'infini dans la direction (A, B, C) , les coordonnées de M_∞ sont donc $(A, B, C, 0)$. En supposant que la coordonnée en Z de m_∞ soit égale à f , le vecteur $\vec{O_c m_\infty}$ est équivalent à :

$$\vec{O_c m_\infty} = \frac{f}{C} \begin{pmatrix} A \\ B \\ C \end{pmatrix}.$$

2.2. DÉTERMINATION DES COORDONNÉES DU CENTRE OPTIQUE

Par convention, l'axe optique coupe le plan image en son centre, de coordonnées $(0, 0, f)$ dans le repère caméra et (u_0, v_0) dans le repère image. Ce centre n'est autre, d'après la propriété 1, que le projeté du point de fuite d'un faisceau de droites parallèles à l'axe optique.

Propriété 2 : Le point de coordonnées (u_0, v_0) dans le repère image, correspond au projeté du point de fuite d'un faisceau de droites parallèles à l'axe optique.

Supposons qu'un objet subisse une translation selon $O_c Z_c$ (fig. 2.2) ; tout point de cet objet se déplace sur une droite parallèle à l'axe optique. La projetée de cette droite passe par le centre de l'image. La connaissance d'au moins deux projetées permet alors de déterminer, par le calcul de leur intersection, les coordonnées (u_0, v_0) du centre de l'image.

Pratiquement, ces projetées sont calculées en suivant des points caractéristiques d'un objet à travers une séquence d'images prises à différentes focales.

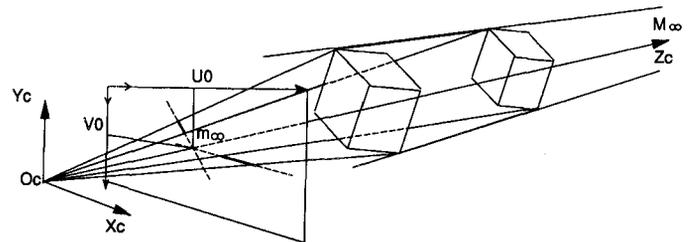


Figure 2.2. — Détermination de (u_0, v_0) .

2.3. RÉSULTATS EXPÉRIMENTAUX

Nous avons utilisé un zoom T 14 x 9 B1ESM d'ANGENIEUX monté sur une caméra CIDTEC CID2250. Les spécificités de la variation de distance focale et de l'ouverture sont respectivement de 9 à 140 mm et $f/1,6$. La résolution de l'ensemble d'acquisition est de 512×512 et la taille du pixel de $15 \mu\text{m} \times 15 \mu\text{m}$. Chaque pixel dans la technologie CID (Charge Injection Device) est adressé et digitalisé de façon individuelle. Contrairement à la caméra CCD les données ne sont pas échantillonnées à partir du

signal vidéo composite mais représentent l'énergie des photons collectée sur chaque site élémentaire.

La détermination de (u_0, v_0) a été effectuée pour une séquence d'images, figure 2.3, d'une grille carrée de dimension 100 mm. Nous avons utilisé comme points caractéristiques les projetés de toutes les intersections des lignes et des colonnes de la grille.

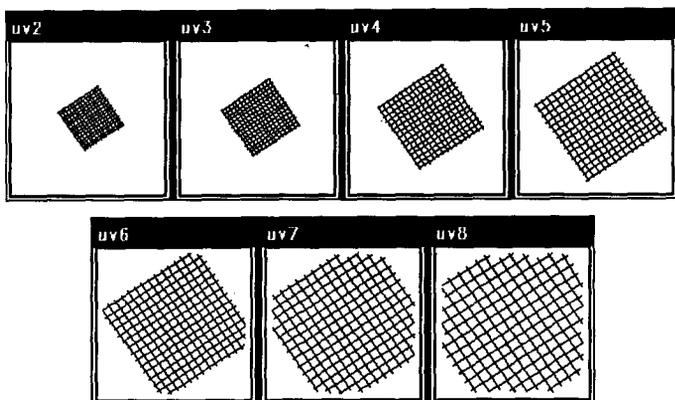


Figure 2.3. — Séquence d'images de grille.

Pour un point caractéristique donné, ses différentes projections sont repérées sur l'ensemble des images. Un processus de moindres carrés nous permet alors d'estimer la droite passant par ces points, correspondant à la projetée de la trajectoire virtuelle et parallèle à l'axe optique, du point pendant la séquence de zoom.

Ce processus est réitéré pour toutes les intersections de la grille. Le point le plus proche, au sens des moindres carrés, des différentes droites ainsi obtenues, nous donne les coordonnées (u_0, v_0) du centre de l'image.

La vue 2.4 montre l'intersection de la projection des droites correspondant aux translations de chacune des intersections de la grille dans l'espace (image de 512×512 pixels). Cette séquence a été réalisée à partir d'une variation de distance focale de 25 à 80 mm.

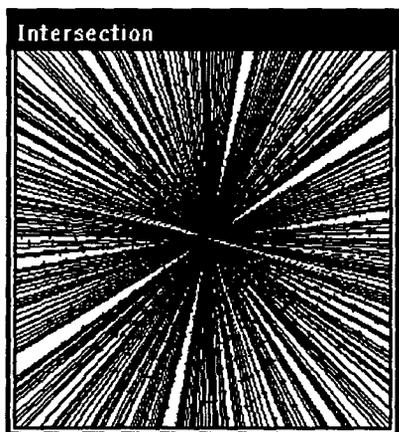


Figure 2.4. — Résultat du calcul de (u_0, v_0) .

La distance moyenne entre la position du centre de l'image et les différentes droites est de 0,007 pixel pour un écart-type de 0,090 pixel.

2.4. STABILITÉ DE (u_0, v_0)

Afin de tenir compte des remarques formulées par Lenz et Tsai sur la stabilité du centre de l'image, nous avons réalisé deux expérimentations pour voir si le centre des lentilles se déplaçait réellement le long de l'axe optique.

L'étude de la stabilité des coordonnées (u_0, v_0) sera réalisée pour une variation de la distance focale, à mise au point constante. En effet pour un zoom parfaitement réglé, la défocalisation n'entraîne aucune modification de la netteté de l'image.

La première expérience consiste à évaluer les coordonnées du centre de l'image pour une faible variation de distance focale à partir de deux images. Si on se réfère aux considérations mathématiques faites précédemment, on peut noter que le faisceau de droites convergentes ne sera pas estimé, dans ce cas-là, avec une grande précision. En effet, pour deux images données, seulement deux points p_1 et p_2 seront utilisés pour estimer la projection de la translation virtuelle d'un point particulier. De plus si la variation de focale est petite, la distance entre les deux points p_1 et p_2 sera faible.

A partir de la même séquence d'images de grille, nous avons combiné les images deux à deux pour des distances focales croissantes (voir tableau 2.2).

Tableau 2.1. — Coordonnées de (u_0, v_0)

combinaison d'images	\bar{u}_0	\bar{v}_0	distance moyenne	écart type
uv2-uv3	263.877	253.619	0.0016	0.0248
uv3-uv4	263.401	253.717	0.0008	0.0289
uv4-uv5	264.149	253.477	-0.0005	0.0530
uv5-uv6	263.801	253.701	0.0026	0.0975
uv6-uv7	263.999	254.096	-0.0103	0.1399
uv7-uv8	263.627	253.791	-0.0162	0.2269

Tableau 2.2. — Correspondance des différentes distances focale indiquées sur le zoom

Image	Focale zoom(mm)	Image	Focale zoom(mm)
uv2	20	uv6	60
uv3	30	uv7	70
uv4	40	uv8	80
uv5	50	-	-

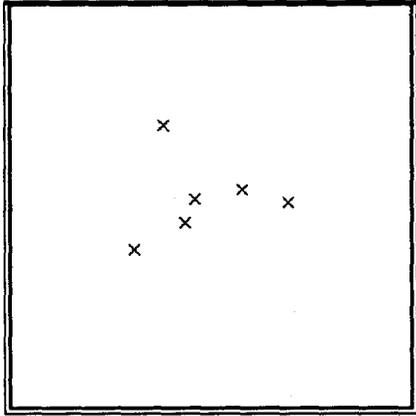


Figure 2.5. — Position des différentes coordonnées (u_0, v_0) dans l'image (taille de la fenêtre : 2×2 pixels).

Comme nous pouvons le noter, sur la figure 2.5 et le tableau 2.1, aucun modèle de déplacement particulier semble se dégager. Les erreurs rencontrées sur les coordonnées de (u_0, v_0) sont dues au bruit sur la localisation des points nécessaires au calcul.

Dans une seconde expérience, tableau 2.3, toutes les combinaisons de deux, trois, ... n images parmi les sept images, sont testées afin d'estimer une valeur moyenne de (u_0, v_0) .

Nous pouvons noter que les écarts-types σ_{u_0} , σ_{v_0} et le paramètre « dist » qui représente la distance moyenne du point estimé (u_0, v_0) au faisceau de droites convergentes, décroissent en fonction du nombre d'images et que l'ensemble des résultats donne une stabilité satisfaisante. L'hypothèse que nous avons émise sur la stabilité du centre de l'image semble vérifiée.

Tableau 2.3. — Coordonnées de (u_0, v_0) pour plusieurs combinaisons d'images.

nb d'images	nb combinaisons	\bar{u}_0 pixel	\bar{v}_0 pixel	σ_{u_0} pixel	σ_{v_0} pixel	dist pixel
2	21	263.795	253.712	0.161	0.129	-0.0015
3	35	263.793	253.706	0.090	0.075	0.0098
4	35	263.792	253.704	0.059	0.045	0.0098
5	21	263.787	253.703	0.038	0.027	0.0091
6	7	263.783	253.700	0.020	0.012	0.0072
7	1	263.782	253.700			0.0074
	TOUTES	263.791	253.706	0.090	0.073	0.0075

2.5. Conclusion

La stabilité des résultats que nous avons obtenue nous semble principalement due à la qualité du matériel utilisé. En effet la précision mécanique du déplacement des blocs mobiles de lentilles, à l'intérieur d'un zoom, est primordial. Le changement de focale pour l'objectif ANGENIEUX que nous avons utilisé, est réalisé par une translation pure

de deux blocs mobiles de lentilles. Cette translation n'est couplée à aucune rotation ce qui n'est généralement pas le cas pour des zoom de qualité standard. Nous pensons que c'est la raison pour laquelle Lenz et Tsai ont observé un déplacement du centre de l'image en forme de spirale pendant leurs expérimentations.

Nous supposons donc au cours des développements à venir que le changement de distance focale est équivalent à une translation du centre optique le long de l'axe optique.

3. Calibrage

La plupart des méthodes de calibrage utilisent la mise en correspondance entre des points tridimensionnels d'un modèle connu et des points bidimensionnels détectés dans une image. Peu de méthodes utilisent comme primitive des droites. La méthode que nous présentons a été développée dans un premier temps pour la localisation de modèle connu, en vision monoculaire, puis généralisée à la calibration de caméra.

3.1. MÉTHODE DE CALIBRAGE

Cette méthode permet de déterminer à la fois les paramètres intrinsèques et extrinsèques de la caméra. Nous supposons que parmi les dix paramètres $\alpha, \beta, \gamma, l_x, l_y, l_z, f_x, f_y, u_0, v_0$, les deux derniers ont été précédemment déterminés. Cette méthode nécessite la connaissance d'un modèle polyédrique de l'objet observé et l'appariement des arêtes de ce dernier aux segments détectés dans l'image. Elle reprend le principe de la procédure de localisation spatiale présentée par Lowe [8], mais en minimisant un critère différent (calculs plus simples) et en estimant conjointement certains paramètres intrinsèques de la caméra. En résumé, si on appelle plan d'interprétation le plan passant par le centre optique et s'appuyant sur un segment de l'image, cette méthode calcule itérativement f_x, f_y et l'attitude spatiale du modèle qui minimisent la somme des distances des arêtes de ce dernier au plan d'interprétation des segments associés. L'algorithme de minimisation adopté est celui de Newton-Raphson.

3.2. Formulation générale du problème

Soient n segments l^i de l'image appariés respectivement avec n arêtes L^i du modèle.

Supposons que chaque segment l^i soit caractérisé, dans le repère caméra R_c , par son vecteur unitaire \vec{v}_c^i et un point p_c^i , nous pouvons facilement calculer la normale unitaire \vec{N}_c^i au plan d'interprétation I^i de l^i . En utilisant les notations suivantes :

$$\vec{v}_c^i = \begin{pmatrix} a_i \\ b_i \\ 0 \end{pmatrix} \quad \vec{O_c p_c^i} = \begin{pmatrix} x_c^i \\ y_c^i \\ f \end{pmatrix} \quad \vec{N}_c^i = \begin{pmatrix} A_i \\ B_i \\ C_i \end{pmatrix}$$

nous obtenons :

$$\vec{N}_c^i = \frac{\vec{v}_c^i \wedge \vec{O}_c \vec{P}_c^i}{\|\vec{v}_c^i \wedge \vec{O}_c \vec{P}_c^i\|} \Leftrightarrow \begin{cases} A_i = \frac{b_i f}{\sqrt{f^2 + d_i^2}} \\ B_i = \frac{-a_i f}{\sqrt{f^2 + d_i^2}} \\ C_i = \frac{d_i}{\sqrt{f^2 + d_i^2}} \end{cases}$$

où $d_i = a_i y_c^i - b_i x_c^i$ est la distance de la droite support de l^i au centre de l'image (u_0, v_0) .

Le plan d'interprétation I^i a pour équation :

$$A_i x_c + B_i y_c + C_i z_c = 0.$$

Considérons les points P_m^{1i} et P_m^{2i} comme étant les extrémités de l'arête L^i . Le but, comme il est montré figure 3.1, est de trouver la rotation $[R_{\alpha, \beta, \gamma}]$, la translation $[T_{lx, ly, lz}]$ et les focales f_x, f_y qui minimisent la somme suivante :

$$S = \sum_{i=1}^n \sum_{j=1}^2 (\text{distance de } P_c^{ji} \text{ à } I^i).$$

P_c^{ji} étant le transformé de P_m^{ji} par $[R_{\alpha, \beta, \gamma}]$ et $[T_{lx, ly, lz}]$, il vient :

$$\begin{aligned} P_c^{ji} &= [R_{\alpha, \beta, \gamma}] P_m^{ji} + [T_{lx, ly, lz}] = \\ &= \begin{pmatrix} X_c^{ji} \\ Y_c^{ji} \\ Z_c^{ji} \end{pmatrix} = [R_{\alpha, \beta, \gamma}] \begin{pmatrix} X_m^{ji} \\ Y_m^{ji} \\ Z_m^{ji} \end{pmatrix} + [T_{lx, ly, lz}]. \end{aligned}$$

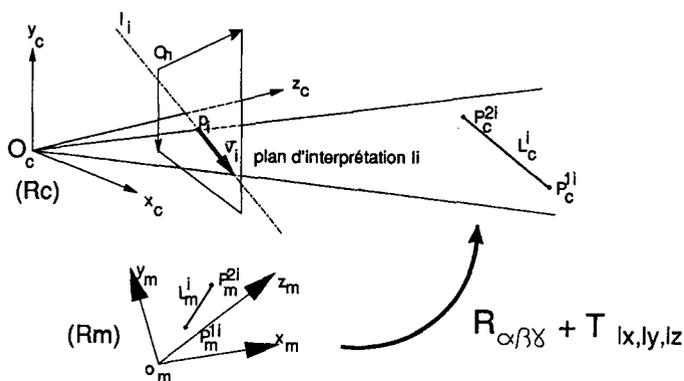


Figure 3.1. — Plan d'interprétation et les différents repères.

Considérons la fonction $F(\alpha, \beta, \gamma, lx, ly, lz, f_x, f_y, P_m^i)$ qui associe à tout point P_m^i , appartenant à une arête L_m^i , la distance de P_c^i au plan d'interprétation I^i . Comme \vec{N}_c^i est un vecteur unitaire, nous avons :

$$F(\alpha, \beta, \gamma, lx, ly, lz, f_x, f_y, P_m^i) = A_i X_c^i + B_i Y_c^i + C_i Z_c^i.$$

Nous recherchons donc le vecteur $(\alpha, \beta, \gamma, lx, ly, lz, f_x, f_y)$ qui minimise :

$$\begin{aligned} S &= \sum_{i=1}^n \sum_{j=1}^2 F(\alpha, \beta, \gamma, lx, ly, lz, f_x, f_y, P_m^{ji})^2 \\ &= \sum_{i=1}^n \sum_{j=1}^2 (A_i X_c^{ji} + B_i Y_c^{ji} + C_i Z_c^{ji})^2. \end{aligned}$$

Il est clair, à travers cette écriture, que les paramètres intrinsèques de la caméra n'interviennent que dans le calcul du vecteur \vec{N}_c^i . De même pour les paramètres extrinsèques qui se retrouvent dans le calcul de P_c^i .

Notons \vec{A} , le vecteur d'état solution composé des trois paramètres de rotation α, β, γ , des trois paramètres de translation lx, ly, lz , et des focales f_x, f_y recherchés et \vec{A}_k le vecteur d'état obtenu à l'étape k du processus itératif

$$\vec{A} = \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ lx \\ ly \\ lz \\ f_x \\ f_y \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_8 \end{pmatrix} \quad \text{et} \quad \vec{A}_k = \begin{pmatrix} \alpha_k \\ \beta_k \\ \gamma_k \\ lx_k \\ ly_k \\ lz_k \\ f_{xk} \\ f_{yk} \end{pmatrix} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{8k} \end{pmatrix}.$$

Ecrivons le développement limité à l'ordre 1 de la fonction $F(\alpha, \beta, \gamma, lx, ly, lz, f_x, f_y, P_m^i)$ au voisinage de \vec{A}_k

$$\begin{aligned} F(\vec{A}, P_m^i) &= F(\vec{A}_k, P_m^i) + \frac{\partial F(\vec{A}_k, P_m^i)}{\partial \vec{A}} (\vec{A} - \vec{A}_k) + \epsilon \\ &= F(\vec{A}_k, P_m^i) + \sum_{l=1}^8 \frac{\partial F(\vec{A}_k, P_m^i)}{\partial a_l} (a_l - a_{lk}) + \epsilon. \end{aligned}$$

En supposant que \vec{A} soit la solution recherchée ($F(\vec{A}, P_m^i) \approx 0$ pour $i \in [1, n]$), nous pouvons faire l'approximation suivante :

$$F(\vec{A}_k, P_m^i) = \sum_{l=1}^8 \frac{\partial F(\vec{A}_k, P_m^i)}{\partial a_l} (a_{lk} - a_l).$$

En posant $\Delta a_l = a_{lk} - a_l$, nous obtenons :

$$\begin{aligned} F(\vec{A}_k, P_m^i) &= \frac{\partial F(\vec{A}_k, P_m^i)}{\partial a_1} \Delta a_1 + \\ &+ \frac{\partial F(\vec{A}_k, P_m^i)}{\partial a_2} \Delta a_2 + \dots + \frac{\partial F(\vec{A}_k, P_m^i)}{\partial a_8} \Delta a_8. \quad (1) \end{aligned}$$

L'équation (1) ayant huit inconnues, il est nécessaire de considérer au minimum 8 équations. Étant donné que nous obtenons une équation pour chaque extrémité d'une arête du modèle, il faut apparier au moins quatre arêtes du modèle avec quatre segments de l'image ($n \geq 4$).

Pour obtenir des résultats plus robustes il est préférable d'apparier plus de quatre arêtes. Dans ce cas le problème

est surdimensionné et nous avons choisi, afin de résoudre le système, un critère quadratique S_k de telle sorte que :

$$S_k = \sum_{i=1}^n \sum_{j=1}^2 \left[F(\vec{A}_k, P_m^{ji}) - \left(\frac{\partial F(\vec{A}_k, P_m^{ji})}{\partial a_1} \Delta a_1 + \frac{\partial F(\vec{A}_k, P_m^{ji})}{\partial a_2} \Delta a_2 + \dots + \frac{\partial F(\vec{A}_k, P_m^{ji})}{\partial a_8} \Delta a_8 \right) \right]^2$$

La minimisation de S_k fournit à chaque itération les compensations $\Delta a_1, \Delta a_2, \dots, \Delta a_8$ à partir desquelles un nouveau vecteur solution \vec{A}_k est estimé. Cette procédure est répétée jusqu'à ce que l'erreur moyenne, $e_k = \frac{1}{2n} S_k$, devienne inférieure à un seuil, fixé au cours de nos expérimentations à 1 mm. Rappelons que l'erreur calculée donne la distance moyenne, ici en millimètre, des extrémités des arêtes du modèle aux plans d'interprétations considérés.

3.3. RÉSULTATS DE CALIBRAGE

Le tableau 3.1 ci-dessous, représente les résultats de calibrage obtenus pour le zoom ANGENIEUX. Nous avons réalisé une séquence d'images, à mise au point fixe, en tournant la bague de réglage de la focale des repères 30 à 80. La mire utilisée pour cette phase de calibration est un cube de 50 mm de côté placé à environ 1 mètre de la caméra. Les valeurs de (u_0, v_0) sont celles déterminées précédemment : $(u_0 = 263,78, v_0 = 253,70)$.

L'étude détaillée du tableau 3.1 montre d'une part la stabilité des valeurs des trois rotations $R_\alpha, R_\beta, R_\gamma$ et des deux translations T_{lx}, T_{ly} , d'autre part la variation de la translation T_{lz} . Elle indique un déplacement virtuel ΔT_z d'environ 500 mm, parallèle à l'axe optique, qui correspond à un éloignement du centre optique de la caméra par rapport à l'objet.

Tableau 3.1. — Paramètres extrinsèques de la caméra.

Repères Zoom	R_α °	R_β °	R_γ °	T_{lx} mm	T_{ly} mm	T_{lz} mm	ΔT_z mm
30	315.20	313.35	6.78	0.0012	-0.0008	997.5	0.0
35	315.18	313.32	6.79	0.0012	-0.0008	1050.6	53.1
40	315.20	313.33	6.78	0.0012	-0.0008	1092.8	95.3
45	315.19	313.32	6.79	0.0012	-0.0008	1125.7	128.2
50	315.19	313.31	6.79	0.0013	-0.0008	1187.9	190.4
55	315.20	313.32	6.78	0.0013	-0.0008	1231.0	233.5
60	315.19	313.31	6.79	0.0013	-0.0008	1275.0	277.5
65	315.19	313.31	6.79	0.0012	-0.0008	1329.3	331.8
70	315.18	313.30	6.80	0.0013	-0.0008	1404.8	407.3
75	315.18	313.30	6.80	0.0013	-0.0008	1456.4	458.9
80	315.18	313.30	6.80	0.0013	-0.0008	1505.0	507.5

Le tableau 3.2 montre les valeurs des différentes focales, elles correspondent à l'ordre de grandeur des valeurs fournies par le constructeur. On peut noter que le rapport f_y/f_x reste stable pour l'ensemble des images.

La représentation d'un changement de distance focale par le modèle sténopé de la figure 3.2 montre qu'à une augmentation de distance focale doit correspondre un rapprochement de même valeur de l'objet. En revanche, la comparaison de ΔT_z et de Δf indique qu'à une augmentation de distance focale correspond une augmentation de la distance entre le centre optique de la caméra et l'objet. Ceci démontre que la représentation du changement de distance focale par le modèle sténopé, telle qu'elle est représentée figure 3.2, ne permet pas de rendre compte des résultats de calibrage. Afin d'analyser ces phénomènes nous avons été amenés à prendre en compte un modèle optique plus complet.

Tableau 3.2. — Paramètres intrinsèques de la caméra.

Zoom units	$f_x * dx$ mm	$f_y * dy$ mm	$\left \frac{f_x}{f_y} \right = \left \frac{dx}{dy} \right $	Δf_x mm
30	40.97	-40.94	1.0007	0,0
35	50.59	-50.55	1.0008	9.62
40	59.69	-59.65	1.0007	18.72
45	67.39	-67.34	1.0007	26.42
50	82.32	-82.27	1.0006	41.35
55	92.08	-92.02	1.0007	51.11
60	103.76	-103.68	1.0008	62.79
65	117.61	-117.52	1.0008	76.64
70	136.26	-136.16	1.0007	95.29
75	150.74	-150.61	1.0009	109.77
80	165.00	-164.85	1.0009	124.03

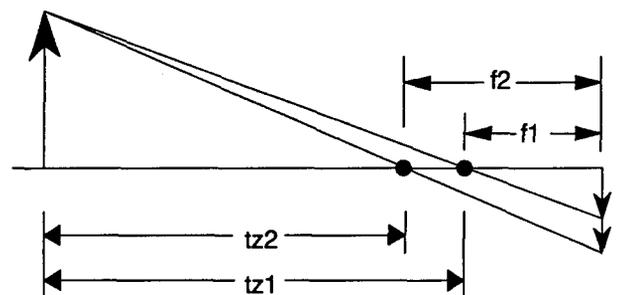


Figure 3.2. — Modèle sténopé pour deux focales différentes.

4. Utilisation du modèle épais d'un zoom

Le calibrage du zoom a montré que les translations virtuelles de l'objet, observées à travers une séquence d'images, ne correspondent pas à la seule variation de distance focale. Un tel phénomène trouve son explication

dans l'étude des systèmes optiques épais. Toutes les hypothèses développées par la suite seront faites sous l'approximation de Gauss [10], c'est-à-dire pour les rayons lumineux proches de l'axe optique : les rayons paraxiaux.

4.1. POSITION DES PLANS PRINCIPAUX DANS UN MODÈLE ÉPAIS

Lorsqu'un système optique est composé d'une succession importante de dioptries, ses dimensions ne rendent plus acceptable son approximation par le modèle sténopé sans précautions au préalable. Il faut alors prendre en compte la position des plans principaux à l'intérieur du modèle, figure 4.1.

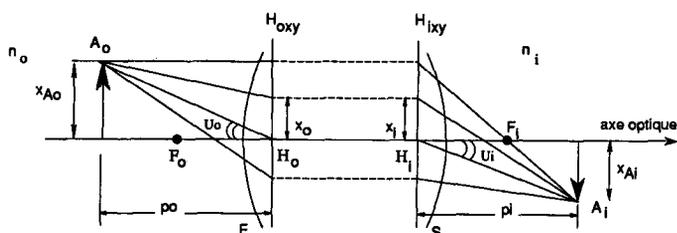


Figure 4.1. — Modèle épais.

4.1.1. Définitions

— **Plans conjugués** : Un système optique fait converger les rayons issus d'un point objet A_o vers un point A_i . On dit alors que A_i est l'image ou le conjugué de A_o .

— **Grandissement transversal G_t** : Soit x_{A_o} et x_{A_i} les distances respectives des points A_o et A_i à l'axe optique, le grandissement transversal G_t est égal au rapport x_{A_i}/x_{A_o} .

— **Plans et points principaux** : Les plans principaux objet et image sont des plans conjugués, perpendiculaires à l'axe optique, tels que le grandissement transversal $G_{tH_oH_i}$ est égal à l'unité (i.e. $x_o = x_i$). Ces plans, notés H_{oxy} et H_{ixy} , sont les plans unitaires de la littérature anglo-saxonne. H_o et H_i sont les points principaux objet et image ; ils correspondent à l'intersection des plans principaux et de l'axe optique ($x = 0, y = 0$).

D'après la définition des plans principaux, nous pouvons en déduire :

Propriété 3 : Dans un modèle optique épais, l'ensemble des rayons optiques se propageant entre les plans principaux sont parallèles à l'axe optique.

— **Grandissement angulaire** : On appelle grandissement angulaire G_a , le rapport des angles d'incidence et d'émergence (α_o, α_i) du rayon optique passant par deux points conjugués de l'axe optique

$$G_a = \left(\frac{u_i}{u_o} \right)_{x=0, y=0}$$

— **Invariant de Lagrange** : La formule de l'invariant de Lagrange indique que pour deux plans conjugués le produit (nux) est invariant. Donc,

$$n_o u_o x_{A_o} = n_i u_i x_{A_i}$$

il vient :

$$\frac{n_i u_i x_{A_i}}{n_o u_o x_{A_o}} = 1$$

où

$$G_t = \frac{x_i}{x_o} \quad \text{et} \quad G_a = \frac{u_i}{u_o},$$

ce qui conduit à la relation connue : $G_t G_a = 1$ si les indices de réfraction des différents milieux sont identiques.

En utilisant la propriété des plans principaux où $G_t = 1$, nous obtenons que $G_a = 1$ (i.e. $u_o = u_i$).

Ce qui se traduit par le fait que tout rayon passant par le point H_o ressort au point H_i avec un angle d'émergence égal à son angle d'incidence.

De cette dernière remarque et de la propriété 3, il est possible d'établir une relation entre le modèle épais et le modèle sténopé, en fusionnant les points H_{oxy} et H_{ixy} soit en déplaçant l'objet et le plan principal objet, soit en déplaçant l'image et le plan principal image. La première transformation correspond à une translation virtuelle de l'objet suivant l'axe optique, la seconde à une translation virtuelle du centre optique de la caméra (ou point d'observation) suivant le même axe.

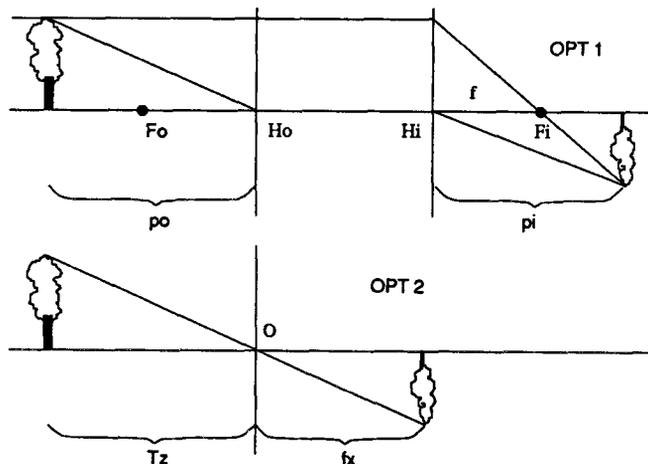


Figure 4.2. — Passage du modèle épais au modèle sténopé.

Notez que ce que nous appelons distance focale f_x prend en compte le tirage de l'objectif. Elle est représentée sur la figure ci-dessus par p_i et ne correspond pas à la distance focale définie par les opticiens $H_i F_i$.

Remarque : La distance focale des opticiens $H_i F_i$ est reliée à f_x et T_x par la relation de Descartes :

$$\frac{1}{f} = \frac{1}{f_x} + \frac{1}{T_x}$$

4.2. VARIATION DES PLANS PRINCIPAUX

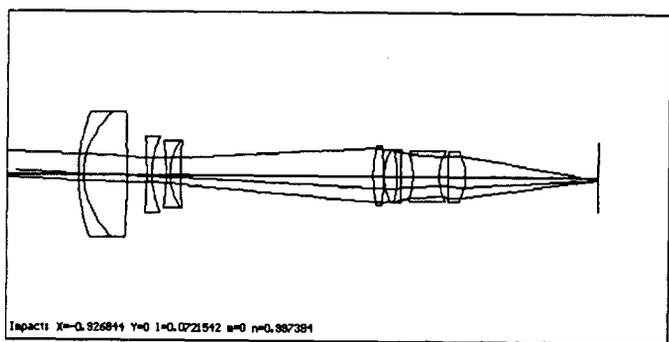
Pour un zoom, le changement de focale est réalisé par le déplacement d'au moins un bloc de lentilles. Nous avons simulé un second zoom ANGENIEUX, à partir d'informa-

tions provenant d'un brevet industriel [1], en introduisant dans le logiciel optique SOLORD [7] les caractéristiques de chacun de ses dioptries (diamètre, rayon de courbure, épaisseur et indice des verres) (tableau 4.1).

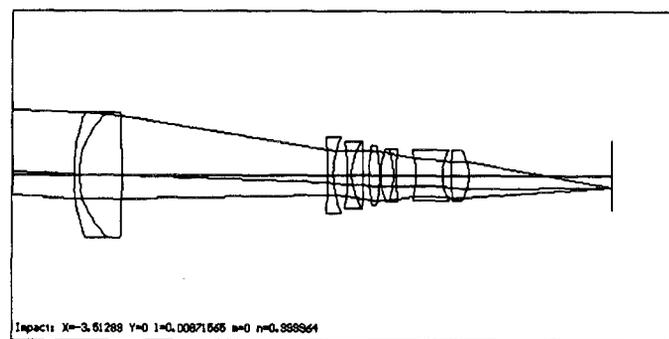
Tableau 4.1. — Caractéristiques des lentilles du Zoom [1].

numéro des lentilles	rayon mm	diamètre mm	épaisseur mm	indice	facteur ABBE
1	+ 61.47	36.00	1.22	1.6669	33.2
2	+ 28.32	36.00	9.71	1.6088	56.6
3	-504.28	36.00	4.46	air	
4	-161.60	22.00	1.22	1.6204	60.2
5	+ 32.63	21.00	3.31	air	
6	- 70.61	18.00	0.91	1.6567	57.0
7	+ 21.57	19.50	2.51	1.6992	30.2
8	+370.97	19.50	43.80	air	
9	+ 52.78	17.00	2.36	1.6588	51.1
10	- 57.68	17.00	0.15	air	
11	+ 21.96	15.00	2.90	1.6211	57.0
12	- 33.14	15.00	0.83	1.6751	32.3
13	+191.20	14.50	2.82	air	
14	- 27.62	13.50	6.00	1.6500	33.8
15	+ 19.41	14.50	2.00	air	
16	+125.25	14.50	4.00	1.6567	57.0
17	- 17.38	14.50			

La figure 4.3 montre les configurations du bloc optique mobile pour les focales minimum et maximum et les rayons extrêmes provenant d'un objet centré sur l'axe optique à 1 mètre du dioptre d'entrée et de dimension voisine de celle du cube utilisé lors de la phase de calibration.



(a) focale minimum



(b) focale maximum

Figure 4.3. — Vues en coupe du zoom simulé.

Ce logiciel nous donne pour chaque dioptre : la focale, le tirage, le tirage pupillaire, les angles et les hauteurs d'ouverture et de champ. Ceci nous permet de déterminer les positions des plans principaux par rapport aux dioptries d'entrée E et de sortie S.

A partir de ce tableau, il est facile de dessiner l'épure des deux systèmes optiques.

Tableau 4.2. — Position des plans principaux.

focale mm	EH_o mm	SH_i mm	ES mm	H_oH_i mm
-21,93	-57,76	-7,90	-88,20	-38,34
-80,39	-100,08	53,88	-89,99	63,89

La figure 4.4 représente respectivement les modèles épais et sténopés pour les focales extrêmes du zoom. Sur les schémas a-b, la distance entre l'objet A_o et l'image A_i reste fixe et les emplacements des plans principaux sont ceux calculés avec le logiciel optique (tableau 4.2). Les schémas c-d représentent les modèles sténopés équivalents obtenus en confondant H_o et H_i . Les résultats obtenus montrent,

— premièrement, un déplacement du point principal image H_i voisin de la variation de distance focale, compte tenu du petit déplacement du dioptre de sortie S,

— deuxièmement, une variation en sens inverse bien plus importante du point principal objet H_o .

Nous avons volontairement choisi de laisser le plan image fixe, l'objet subissant, entre les deux représentations, une translation virtuelle selon l'axe optique. Le raisonnement inverse (objet fixe, centre optique du modèle sténopé mobile) conduit aux mêmes conclusions :

A une augmentation de la distance focale du modèle sténopé correspond une augmentation de la distance entre l'objet et le centre optique de ce modèle.

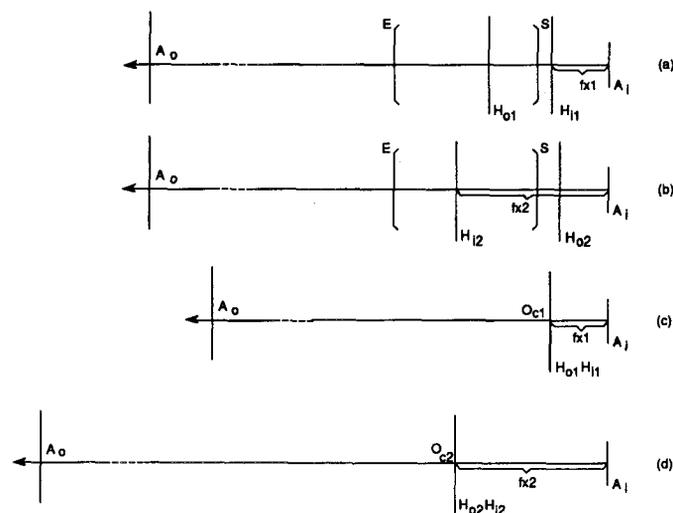


Figure 4.4. — Modèles équivalents épais et sténopés pour les deux distances focales extrêmes du zoom simulé.

Pour le zoom simulé le sens de variation de la translation virtuelle est en accord avec les résultats de calibrage obtenus avec notre objectif ANGENIEUX.

La translation virtuelle maximale est égale à :

$$\Delta T_z = T_{z2} - T_{z1},$$

elle correspond à la variation du point principal objet.

Ces résultats permettent d'expliquer les données de calibrage que nous avons obtenues au paragraphe 3.3.

Remarque : L'approximation du système optique d'un zoom, par le modèle équivalent sténopé, suppose que la distance entre l'objet et l'image ne soit plus considérée fixe.

5. Reconstruction d'un objet polyédrique

L'idée de reconstruction à partir d'images issues d'une séquence de zoom n'est pas nouvelle. En effet Ma a expérimenté cette technique en combinant la variation de focale et l'analyse des flots optiques [9] et a abordé également, dans cet article, la reconstruction des segments de droites. Les deux principales remarques que nous formulons à l'encontre de ce travail sont :

— d'une part, que la formulation mathématique proposée est basée sur une translation de l'objet égale à la variation de focale. Nous avons montré que cette approximation n'était pas acceptable dans le cas d'un zoom,

— d'autre part, que les résultats présentés ne concernent que des images de synthèse.

Les difficultés majeures que nous avons rencontrés durant nos expériences sont venues du manque de fiabilité dans la localisation des segments détectés dans les images. En effet et nous reviendrons sur ce point dans quelques paragraphes, une telle méthode de reconstruction nécessite un calcul des primitives à une précision inférieure au pixel. Or les expériences que nous avons pu réaliser nous ont montré que, pour l'instant, on ne pouvait pas obtenir de façon fiable une localisation précise d'un contour, correspondant à un changement de niveau de gris dans une image, à partir des détecteurs de contour classiques. C'est la raison pour laquelle nos efforts ont porté essentiellement sur deux points.

— pouvoir tenir compte des déformations optiques introduites par le zoom,

— utiliser des primitives de reconstruction plus fiables dans les images.

L'ensemble des expériences décrites dans cet article prend en compte la distorsion géométrique. En effet le constructeur du zoom (ANGENIEUX) a pu nous fournir, pour l'ensemble des combinaisons du réglage de la focale et de la mise au point, la distorsion géométrique théorique introduite par le système optique. Le déplacement des points s'avère important puisqu'il peut atteindre 3 %, soit un déplacement de 6 pixels pour un point situé à 200 pixels du centre de l'image.

Afin d'obtenir une localisation précise de la projection perspective d'une primitive tridimensionnelle, nous avons tracé des lignes blanches caractéristiques sur des objets polyédriques noirs. La détection de tels amers se fait directement dans l'image de luminance sans recourir au détecteur de contour. Les algorithmes utilisés calculent les paramètres de localisation et d'amplitude d'un modèle de luminance des lignes blanches dans l'image [11].

La figure 5.1 représente un détail de la superposition des points de contour à l'image de luminance. Le carré élémentaire représente un pixel et on peut noter la bonne localisation, au sous-pixel, des points détectés correspondant aux lignes blanches.

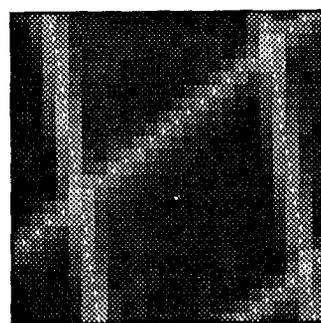


Figure 5.1. — Superposition des points détectés à l'image de luminance.

5.1. PRINCIPE DE RECONSTRUCTION

Le principe de reconstruction est fondé sur un processus de triangulation. Comme le montre la figure 5.2, le déplacement des plans principaux, au cours d'une séquence d'images prises en modifiant la distance focale, entraîne un déplacement des centres optiques des modèles sténopés équivalents et permet de mettre en œuvre la reconstruction. Cette technique est équivalente à de la stéréovision axiale. A ce titre une étude complète sur les avantages, inconvénients, mais surtout sur les erreurs de triangulation de cette technique de reconstruction a été réalisée dans [2].

Le processus de reconstruction, que nous avons développé, s'applique aux objets polyédriques. Il revient, comme le montre la figure 5.3, à calculer l'intersection des différents plans d'interprétation, relatifs à une même arête de l'objet, dans chacune des images de la séquence. Afin de faciliter la résolution, nous avons utilisé une minimisation sous contrainte.

Soit l_j^i la projection de l'arête L_i dans la j -ième image de la séquence. L'équation de son plan d'interprétation, plan passant par l'origine, dans le repère caméra R_c^j au moment de la prise de vue peut s'écrire sous la forme suivante :

$$a_j^i x + b_j^i y + c_j^i z = 0.$$

Pour résoudre le problème de la triangulation, il est nécessaire d'exprimer l'ensemble des plans d'interprétation dans un même repère. Nous avons choisi de nous placer dans le référentiel lié à la première image R_c^1 . Dans ce

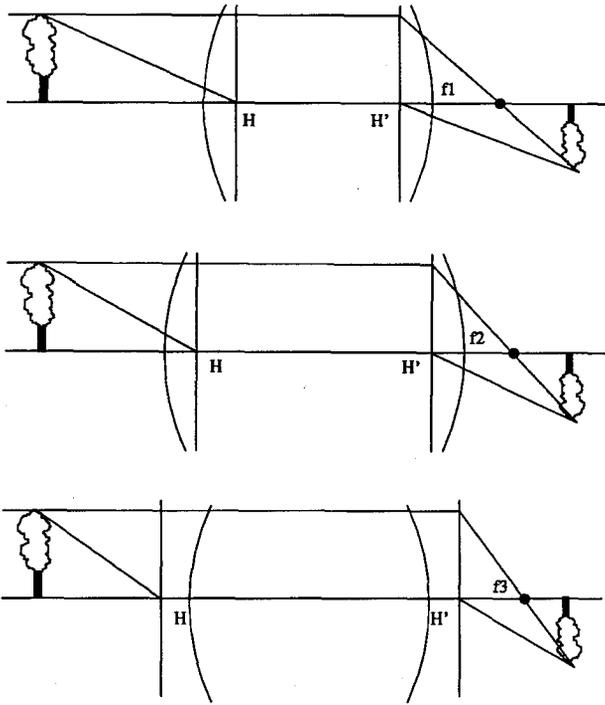


Figure 5.2. — Plans principaux d'un objectif à différentes focales.

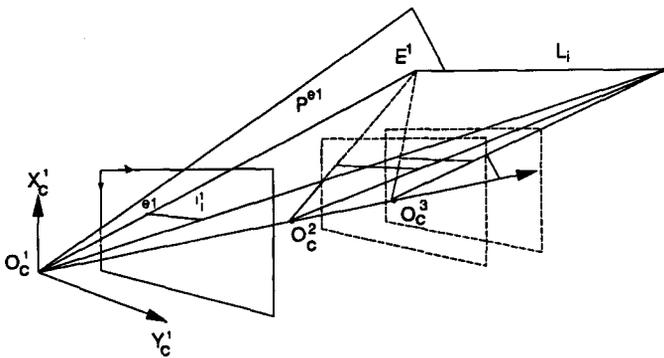


Figure 5.3. — Intersection des plans d'interprétation.

repère, le plan d'interprétation de l_i^j est défini par l'équation :

$$a_i^j x + b_i^j y + c_i^j z + d_i^j = 0 \quad \text{avec} \quad d_i^j = -c_i^j Z_0^j$$

Z_0^j définit, dans le repère R_c^j , la position du centre optique pour la j -ième image.

La reconstruction de l'arête L_i est réalisée par le calcul de deux de ses points E_1, E_2 . Soit le point e_1 de coordonnées (x^{e1}, y^{e1}) correspondant à une des extrémités du segment l_i^j . Soit le plan P^{e1} contenant l'axe optique et passant par e_1 ; nous recherchons alors le point E_1 dont les coordonnées (x, y, z) minimisent le critère suivant :

$$L = \sum_{j=1}^n [a_i^j x + b_i^j y + c_i^j z + d_i^j]^2$$

sous la contrainte :

$$-y^{e1} x + x^{e1} y = 0 \Rightarrow x = y \frac{x^{e1}}{y^{e1}}$$

où chaque élément de la somme représente la distance entre le point recherché et le plan d'interprétation considéré, sous la contrainte que ce point appartienne au plan P^{e1} , ce qui conduit, en introduisant la contrainte, à minimiser le système suivant :

$$L = \sum_{j=1}^n \left[\left(a_i^j \frac{x^{e1}}{y^{e1}} + b_i^j \right) y + c_i^j z + d_i^j \right]^2 \quad \text{pour } j \in [1, n].$$

D'une façon similaire, nous pouvons déterminer les coordonnées du point E_2 qui se projette, dans l'image, au point e_2 .

Nous obtenons ainsi les coordonnées des deux extrémités d'une arête, reconstruite dans le repère caméra et référencée dans la première image.

Le calibrage fournit, outre les paramètres de formation d'image, les paramètres extrinsèques $R_{\alpha, \beta, \gamma}$ et $T_{u, v, w}$ donnant la position spatiale de la mire compatible avec l'image, dans le repère caméra. La transformation de l'arête reconstruite L_i , par T^{-1} et R^{-1} permet de replacer le segment reconstruit dans le repère original de la mire.

5.2. RÉSULTATS EXPÉRIMENTAUX SUR DES SÉQUENCES D'IMAGES RÉELLES

Les résultats expérimentaux que nous présentons, concernent deux objets polyédriques distincts (fig. 5.4) reconstruits à partir d'images réelles. Il s'agit du cube de 5 cm de côté, qui nous a également servi d'étalon de calibrage, et d'un polyèdre quelconque dont les dimensions sont voisines de celles du cube.

5.2.1. Processus expérimental

Le processus expérimental de la reconstruction se déroule selon le schéma suivant :

- une saisie d'une séquence d'images pour des distances focales et des mises au point soigneusement repérées sur l'objectif,
- un traitement de bas niveau, qui consiste à extraire les différents amers dessinés sur les surfaces de l'objet. Dans le cas du cube, nous avons utilisé une détection d'amers de type croix et dans le cas du polyèdre une détection de type ligne [11],
- une prise en compte des déformations introduites par le système optique. La connaissance des conditions de prise de vue, distance focale et mise au point, permet de compenser chaque point de contour du facteur de distorsion radiale, à partir des tables de correction fournies par le constructeur de l'objectif,
- une phase de segmentation,
- une mise en correspondance entre les projetées d'un même amer sur les différentes images de la séquence,
- une reconstruction par triangulation de l'ensemble des segments appariés.

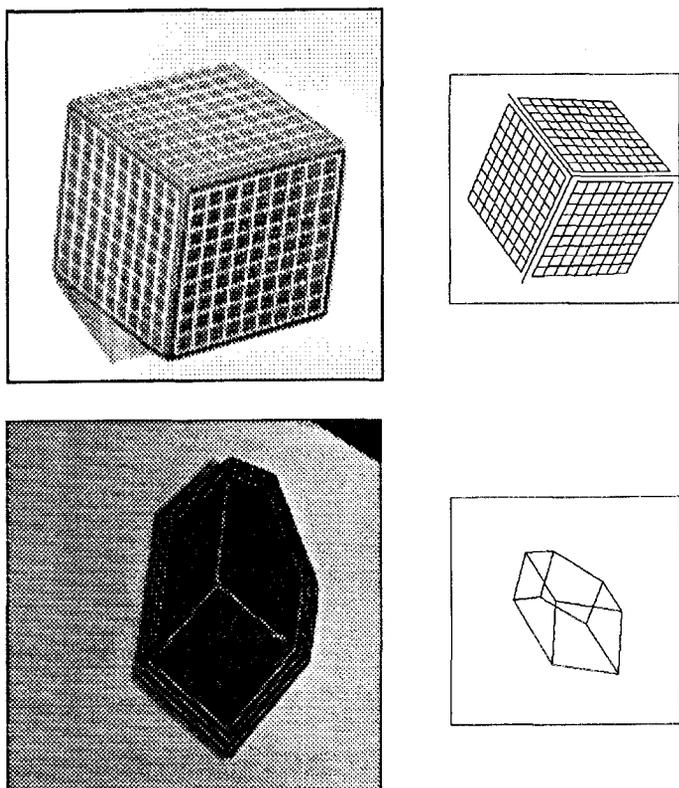


Figure 5.4. — Images de luminance de deux objets étudiés et leur modèle associé.

5.2.2. Reconstruction du cube à partir des données de calibrage

La première séquence d'images que nous présentons a été utilisée dans un premier temps pour le calibrage de la caméra et de l'objectif, puis pour la reconstruction. La figure 5.5 montre les images segmentées du cube étalon, placé à environ 0,8 m de la caméra, pour quatre distances focales différentes. Nous avons utilisé sur les lignes des grilles internes (8×8) à chaque face du cube comme primitives d'étalonnage.

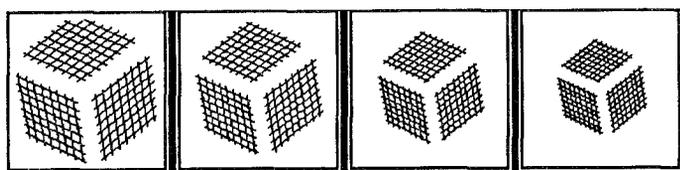


Figure 5.5. — Séquence d'images segmentées.

La reconstruction est réalisée à partir des six segments par face les plus éloignés du centre de l'image, étant donné que la stéréovision axiale ne permet pas d'inférer des informations 3D fiables pour des primitives proche de l'axe optique. La figure 5.6 montre les 18 segments reconstruits, superposés au modèle du cube pour différents points de vue.

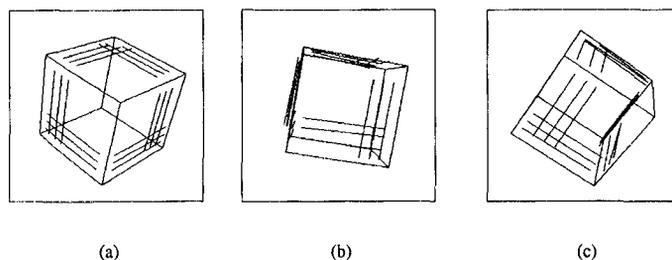


Figure 5.6. — Superposition du modèle reconstruit et du modèle théorique pour différents points de vue.

Afin d'évaluer les erreurs de reconstruction du cube, nous avons calculé la distance e_i , ($i = 1, 2$) de chaque extrémité des segments reconstruits aux arêtes associées du modèle (fig. 5.7). Le tableau 5.1 présente les résultats numériques de la reconstruction pour quelques-uns des segments. Les coordonnées (X_m, Y_m, Z_m) correspondent aux mesures de référence des deux extrémités d'une arête du modèle, les coordonnées (X_r, Y_r, Z_r) proviennent de la reconstruction.

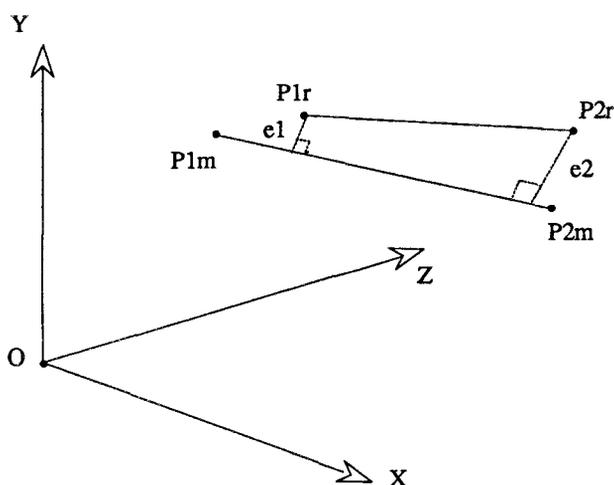


Figure 5.7. — Estimation des erreurs de reconstruction.

Tableau 5.1. — Exemples de segments reconstruits

		X_m (mm)	Y_m (mm)	Z_m (mm)	X_r (mm)	Y_r (mm)	Z_r (mm)	e_i (mm)
1	E1	-24.68	-22.40	17.08	-25.20	-18.30	16.60	0.71
	E2	-24.68	22.09	17.22	-24.80	19.70	17.10	0.17
2	E1	-24.68	-12.26	-22.31	-21.60	-15.20	-15.90	4.23
	E2	-24.68	-12.52	22.03	-21.80	-15.20	22.30	3.93
3	E1	-21.03	-07.73	-24.68	-19.20	-07.00	-25.40	1.03
	E2	23.50	-07.90	-24.68	19.20	-06.50	-26.20	2.06
4	E1	-21.04	-17.61	-24.68	-19.40	-16.70	-25.60	1.30
	E2	23.47	-17.79	-24.68	12.00	-16.50	-26.00	1.81
5	E1	13.68	21.77	-24.68	13.60	20.10	-24.70	0.03
	E2	13.62	-22.62	-24.68	12.60	-17.10	-25.80	1.55
6	E1	-22.13	24.68	12.57	-18.20	23.50	13.90	1.78
	E2	22.46	24.68	12.56	21.70	22.90	14.70	2.78

Il est possible d'évaluer également l'écart angulaire entre deux segments quelconques qui sur l'objet sont soit parallèles, soit orthogonaux. L'erreur $\bar{\alpha}$ est estimée à partir

de la somme des produits scalaires de toutes les combinaisons entre les arêtes reconstruites :

$$\bar{a} = \frac{1}{n} \sum a_{ij} \quad \text{avec,}$$

$$a_{ij} = (|\vec{L}_i \cdot \vec{L}_j|) \quad \text{si } [|\vec{L}_i \cdot \vec{L}_j| < 0,5]$$

ou bien,

$$a_{ij} = (1 - |\vec{L}_i \cdot \vec{L}_j|) \quad \text{si } [|\vec{L}_i \cdot \vec{L}_j| > 0,5].$$

Nous pouvons noter que les valeurs de reconstruction (X_r, Y_r, Z_r) du tableau 5.1 sont proches des valeurs théoriques (X_m, Y_m, Z_m) et que l'écart angulaire (\bar{a} ci-dessus) entre les segments reconstruits est faible. Ces résultats de reconstruction montrent qu'il est bel est bien possible d'inférer des informations de profondeur, à partir d'un zoom. Les résultats ne sont cependant pas parfaits, nous pouvons noter dans le deuxième exemple de reconstruction du tableau 5.1, un écart de positionnement de 4,23 mm.

Il est clair qu'utiliser les mêmes segment pour calibrer et reconstruire n'apporte rien en soi. Cette démarche permet toutefois de se rendre compte de la précision qu'il est possible d'atteindre par ce processus de reconstruction, sachant qu'une reconstruction réelle conduira inéluctablement à une dégradation de ces résultats.

L'erreur moyenne entre les lignes reconstruites et les données mesurées sur le cube, pour les 18 segments est égale à :

$$\bar{e} = 2,18 \text{ mm, avec } \sigma_e = 1,02 \text{ mm.}$$

L'erreur moyenne angulaire \bar{a} , est quand à elle égale à :

$$\bar{a} = 0,019, \quad \text{avec } \sigma_a = 0,015,$$

ce qui équivaut à un écart angulaire moyen de 1.09°.

5.2.3. Reconstruction du cube

La seconde expérience de reconstruction, figure 5.8, correspond à une nouvelle séquence d'images du cube, prise pour une attitude différente. Les données de calibrage nécessaires à l'algorithme sont celles déterminées à partir de la première séquence d'images. Cette reconstruction est donc totalement déconnectée de la phase d'étalonnage et nous nous sommes repositionnés le plus près possible des conditions de prise de vue (réglage du zoom) de la première expérience.

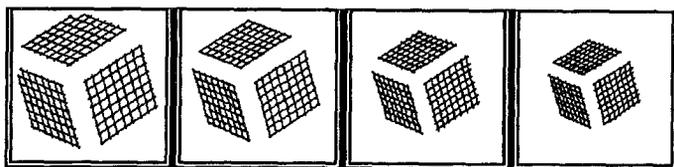


Figure 5.8. — Séquence d'images segmentées.

La figure 5.9 représente le cube reconstruit dans différentes attitudes. L'allure générale est satisfaisante puisque nous reconstruisons des segments sensiblement coplanaires dans les trois plans de référence du cube. Les valeurs angulaires entre chaque segment sont globalement respectées.

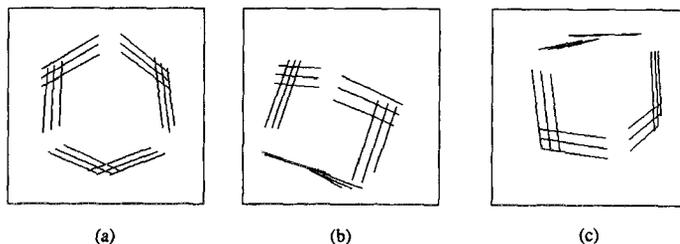


Figure 5.9. — Différents points de vue du cube reconstruit.

L'algorithme de reconstruction calcule les coordonnées des extrémités des segments reconstruits dans le repère du modèle utilisé pour la séquence de calibrage. La comparaison entre les deux modèles n'est donc pas immédiate. Nous avons de ce fait mis-au-point un processus itératif qui permet de déterminer l'attitude 3D du modèle reconstruit, minimisant un critère de distance par un rapport au modèle de référence.

Nous pouvons noter sur la figure 5.9 que le modèle reconstruit est proche de celui calculé à partir des données de calibrage. Les erreurs de reconstruction que nous avons obtenues sont les suivantes :

$$\bar{e} = 3,17 \text{ mm, avec } \sigma_e = 1,46 \text{ mm}$$

L'erreur moyenne angulaire \bar{a} est égale à :

$$\bar{a} = 0,029, \quad \text{avec } \sigma_a = 0,021$$

ce qui équivaut à un écart angulaire moyen de 1.66°.

Au vu de ces résultats, nous pouvons constater que l'erreur de positionnement des segments \bar{e} s'est légèrement dégradée et que l'écart angulaire moyen augmente modérément. Ceci est très certainement dû aux incertitudes de repositionnement du zoom, lors de l'acquisition de la nouvelle séquence d'images.

5.2.4. Reconstruction du polyèdre

Le dernier exemple de reconstruction concerne le polyèdre représenté sur la figure 5.4. De la même façon que pour la reconstruction précédente, nous avons réalisé une séquence d'images (fig. 5.10) en accord avec les données de calibrage.

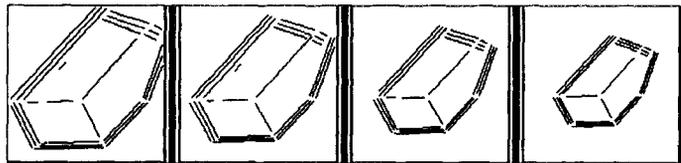


Figure 5.10. — Séquence d'images du polyèdre obtenue avec un zoom.

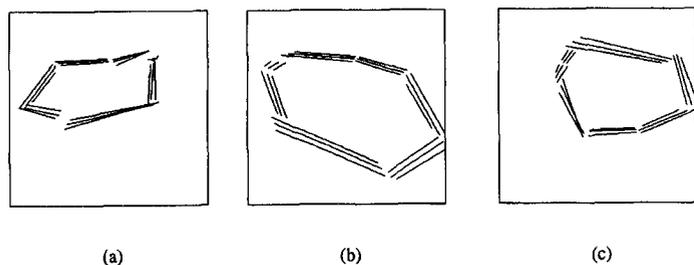


Figure 5.11. — Différents points de vue du polyèdre reconstruit.

On peut noter que ce type de reconstruction nous donne des résultats encourageant, mais nous pensons que certaines améliorations doivent être apportées. Il faut toutefois souligner qu'il n'est pas possible d'obtenir ces résultats sans tenir compte, entre autre, de la distorsion radiale.

La figure 5.12 montre qu'une faible erreur dE sur la détection des points dans les images entraîne des erreurs importantes dX , dY , dZ sur la localisation des points dans l'espace. L'erreur dZ sera d'autant plus importante que le point reconstruit sera plus proche de l'axe optique [2].

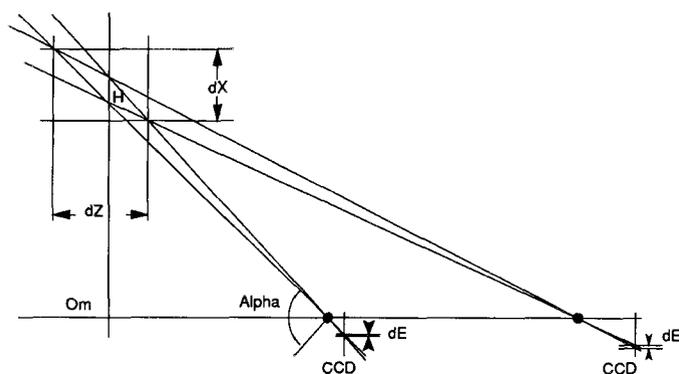


Figure 5.12. — Erreur de triangulation.

CONCLUSION

L'utilisation d'un zoom en vision monoculaire est un phénomène qui tend à se développer. De plus en plus d'équipes se penchent sur les propriétés optiques d'un tel objectif.

Son utilisation triviale permet, sans avoir à se déplacer, d'obtenir des images de tailles différentes d'un centre d'intérêt de la scène observée. A travers cet article, nous tentons d'aller plus loin en extrayant des informations tridimensionnelles à partir des propriétés optiques du zoom.

Par ailleurs, nous mettons en garde le lecteur sur les conclusions hâtives qu'entraîne une modélisation trop simpliste du changement de focale engendré par cet objectif. Afin d'analyser clairement les phénomènes optiques mis en jeu par ce changement, il est nécessaire de prendre en compte l'ensemble des paramètres du modèle optique épais.

Les résultats du calibrage, en accord avec l'analyse du modèle optique épais, montre qu'à travers une séquence d'images prises à l'aide d'un zoom, le centre optique du modèle équivalent sténopé subit une translation non négligeable selon l'axe optique. Ce déplacement n'est pas proportionnel à la variation de focale, mais à la variation du plan principal objet.

Des perspectives de reconstruction à partir d'un zoom sont alors envisageables. Cependant, il est clair que cette méthode est sensible aux erreurs de localisation des primitives dans les images.

L'objectif de nos recherches est de continuer à explorer les possibilités de reconstruction offertes par un zoom. Entre autre, nous avons développé une méthode de reconstruction, qui permet de s'affranchir d'un calibrage explicite. Le principal atout de ces nouveaux travaux est de ne plus émettre de contrainte sur la nature du déplacement à réaliser au cours de la séquence d'images, ce qui conduit, dans notre cas, à supprimer l'hypothèse du déplacement parfait du centre optique le long de l'axe optique.

BIBLIOGRAPHIE

- [1] P. ANGENIEUX, *Objectif à longueur focale variable*, Brevet d'invention n° 1.120.271, G 02 b, juillet 1956.
- [2] C. CARMONA and al., *Comparison between Two Axial Stereovision Systems*, Application of Digital Image Processing XIII, SPIE vol. 1349, 1990.
- [3] O. FAUGERAS, G. TOSCANI, *Camera Calibration for 3D Computer Vision*, Proceeding of Int. Work. On Machine Vision and Machine Intelligence, Tokyo, Japan, February 1987.
- [4] R. K. LENZ, R. Y. TSAI, *Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3D Machine Vision Metrology*, IEEE Robotics and Automation, pp. 68-75, Raleigh, April 1987.
- [5] R. K. LENZ, R. Y. TSAI, *Calibrating a Cartesian Robot with Eye-on-Hand Configuration Independent of Eye-to-Hand Relationship*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, n° 9, pp. 916-928, September 1989.
- [6] P. LIMOZIN-LONG, *Présentation et comparaison de trois méthodes de calibration*, Convention de recherche ESA INRIA, rapport intermédiaire, février 1988.
- [7] LORD INGENIERIE, *Logiciel de stimulation optique SOLORD*, 1989.
- [8] D. G. LOWE, *Perceptual Organisation and Visual Recognition*, Robotics : Vision, Manipulation and Sensors, Kluwer Academic Publishers, Boston, 1985.
- [9] J. MA, S. I. OLSEN, *Depth from Zooming*, J. Opt. Soc. Am. A, vol. 7, n° 10, pp. 1883-1890, October 1990.
- [10] J. P. PEREZ, *Optique géométrique, matricielle et ondulatoire*, Masson, Paris, 1988.
- [11] B. PEUCHOT, *Camera Virtual Equivalent Model 0.01 Pixel Detector*, 14th Int. Conf. of IEEE Engineering in Medicine and Biology, Rennes, France, pp. 41-45, November 1992.
- [12] R. Y. TSAI, *A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology using Off-the-Shelf TV Cameras and Lenses*, Research Report RC 11413 IBM T. J. Watson Research Center, Yorktown Heights, NY 10598, September 1985.
- [13] G. Q. WEI, Z. Y. HE, *Determining Vanishing Point and Camera parameters : New Approaches*, Proc. Of 9th Int. Conf. on Pattern Recognition, Italy 1988.

- [14] L. WEIQIAN, K. GUANGWEN, X. TIANNING, *Determination of the Deformation of the Bridge Model in Real Time with CCD Solid State Camera*, SPIE. Vol. 1395, pp. 1181-1188, Zurich, Switzerland, September 1990.
- [15] A. G. WILEY, K. W. WONG, *Metric Aspect of Zoom Vision*, SPIE. Vol. 1395, pp. 112-118, Zurich, Switzerland, September 1990.
- [16] K. W. WONG, M. LEW, Y. KE, *Experience with Two Vision Systems*, SPIE. Vol. 1395, pp. 3-7, Zurich, Switzerland, September 1990.

Manuscrit reçu le 4 mars 1992.