

STÉRÉOVISION MONOCULAIRE PASSIVE DEPUIS DEUX IMAGES FLOUES

A. Benzougar, T. Simon, M. Koudeir, J. Bernard

Equipe Signal, Image et Communication (SIC) de l'IRCOM, URA n°356
Université de Poitiers, UFR Sciences, 40, av. du Recteur Pineau, 86022 Poitiers Cedex
Tel : 49 45 37 12, email : simon@ingrid.univ-poitiers.fr

RÉSUMÉ

La méthode de stéréovision monoculaire passive proposée exploite deux images floues d'une même scène 3D, acquises d'un même point de vue mais avec une ouverture de l'iris différente. Le système optique ne possédant qu'un seul plan net à une profondeur donnée, les images présentent un flou lié à la profondeur. La modélisation de la réponse impulsionnelle du système optique et la modélisation des discontinuités de luminance dans l'image sont nécessaires à l'évaluation du flou. Cette évaluation après calcul du moment d'ordre deux du gradient normé des images, est précédée d'une phase de calibration permettant d'identifier les caractéristiques globales de la caméra. Les lois de l'optique géométrique utilisant ces caractéristiques donnent alors la carte des profondeurs de la scène observée.

1. Introduction

La perception des trois dimensions dans les applications de visionique est une nécessité qui n'est plus à démontrer. La perte de l'information de profondeur et du facteur d'échelle, provoquée par la projection sur le capteur de la scène 3D observée, est compensée par l'exploitation de plusieurs images obtenues avec des caractéristiques d'acquisition différentes. Parmi l'ensemble des techniques de stéréovision, nous avons porté notre attention sur les méthodes monoculaires passives utilisant le flou optique comme information de profondeur. Deux images floues prises d'un même point de vue, et avec des ouvertures de l'iris différentes sont comparées afin de déterminer la carte des profondeurs.

La méthode présentée nécessite au préalable la modélisation de la réponse impulsionnelle du système optique et des discontinuités de luminance. La connaissance de ces deux modèles permet la détermination du flou optique sur les contours d'objets, puis des profondeurs au travers des relations d'optique géométrique

Nous présentons en premier lieu, les modèles retenus puis nous décrivons notre méthode de détermination du flou optique. Dans un second lieu, nous précisons la phase de calibration permettant de connaître les caractéristiques de la caméra pour chaque profondeur et calculons la carte des profondeurs de la scène. Enfin nous terminons par un exemple de résultat.

ABSTRACT

The passive monocular stereovision method proposed, is based on the exploitation of two blurred images from the same 3D scene, acquired by a same point of view and with the same optical parameters except aperture of iris, in order to determine a depth map. As the optical system get only one sharp plan, contours in the images present a blur related to the depth. The blur degree's assessment needs a preliminary patternmaking of the pulse response of optical system and also of luminance discontinuities. The blur parameter is deduced from the second moment of the image normalized gradient. The geometric optic laws allow to establish a relation between depth, blur and global characteristics of the camera which are evaluated at the end of the calibration phase.

2. Le flou optique : Une information de profondeur

2.1. Relation entre la profondeur et la tache image

Nous schématisons l'objectif de la caméra par une lentille mince (figure 1). La relation (1) classique de l'optique géométrique lie la position du plan image focalisé, image nette, à la position du plan objet contenant un point source :

$$\left| \frac{1}{Z_i} + \frac{1}{d} - \frac{1}{f} \right| = 0 \quad (1)$$

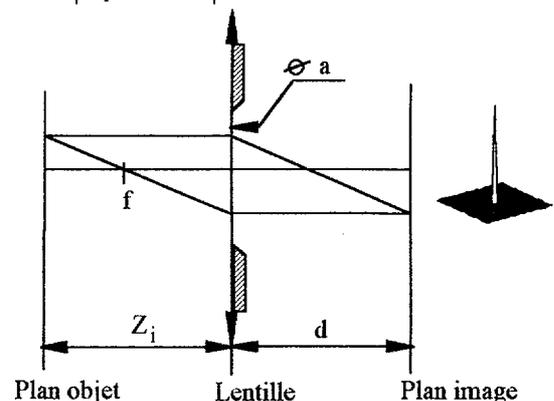
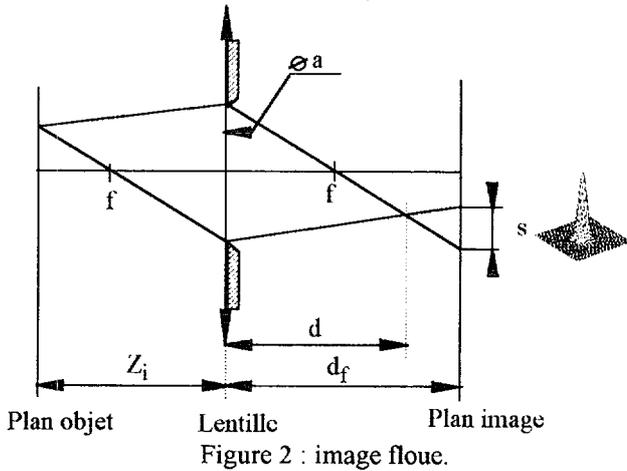


Figure 1 : image nette focalisée.



Dans le cas d'un plan image défocalisé, image floue, l'image du point source est une tache dont le diamètre, pour un iris circulaire, est lié à la distance entre la lentille et l'objet par la relation :

$$s = a \left| \frac{d_f - f}{d_f} \right| \left| \frac{d_f}{d_f - f} \frac{Z_i - f}{Z_i} \right| \quad (2)$$



Le diamètre s de la tache dépend des caractéristiques du système optique, la focale f, l'ouverture a, la position du plan image d_f et de la profondeur Z_i . L'ensemble des techniques utilisant le flou optique comme information de profondeur exploite cette relation.

2.2. Modélisation de la réponse impulsionnelle du système optique

Pour mesurer le diamètre de la tache image, il faut connaître la répartition de l'énergie de luminance dans cette tache. On utilise habituellement un modèle gaussien tenant compte des effets d'aberrations géométriques et chromatiques de la lentille :

$$h(x,y) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \quad (3)$$

Le modèle gaussien permet de considérer que le diamètre de la tache est proportionnel à l'écart-type de la gaussienne : $s=k\sigma$, avec $k=6$ nous prenons en compte 99% de l'énergie recueillie dans cette tache par le capteur. L'écart-type de ce modèle est représentatif du flou optique et permet de calculer la profondeur par la relation (4) :

$$\frac{1}{Z_i} = \frac{\sigma_i}{m} - \frac{c}{m} \quad (4)$$

où m et c englobe les paramètres du système optique, et sont déterminés dans la phase de calibration. Il existe pour chaque profondeur un modèle gaussien d'écart-type différent. La figure 3 présente l'évolution du flou en fonction de la profondeur pour différentes ouvertures de l'iris.

L'image floue $if_i(x,y)$, pour une profondeur Z_i , est déduite de l'image nette $in_i(x,y)$ par la relation (5) :

$$if_i(x,y) = h_i(x,y) \otimes_{2D} in_i(x,y) \quad (5)$$

L'image nette est la projection sur le plan image de la scène observée, sans tenir compte des effets de la défocalisation.

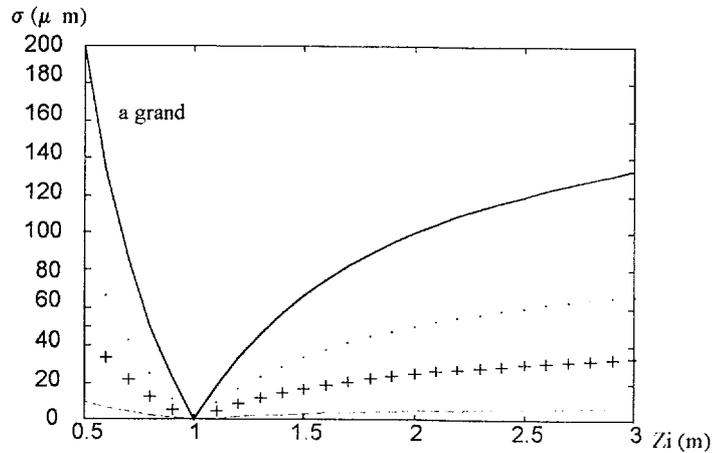


Figure 3 : Relation flou-éloignement.

2.3. Modèle de discontinuité de luminance

Dans des images réelles, il n'existe généralement pas de source ponctuelle isolée, pour chaque plan de profondeur, l'évaluation directe de la réponse impulsionnelle du système optique est impossible. Cependant, une scène éclairée réagit en réflexion comme un ensemble structuré de sources ponctuelles réémettant dans toutes les directions si l'on admet un modèle lambertien. Les discontinuités de luminance, compte tenu de la géométrie de la scène, représentent le plus souvent des contours d'objets. Ces discontinuités, dans une seule direction, sont modélisées comme le montre la figure 4 où ϵ est lié à la géométrie des objets composants la scène et à leur éloignement par rapport à la caméra.

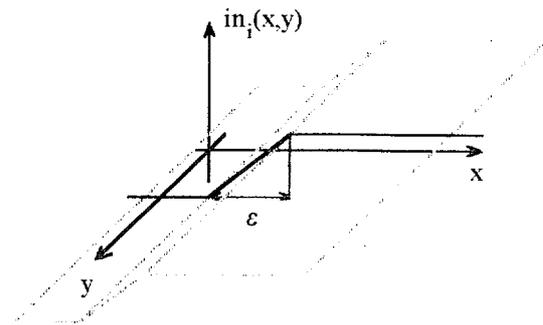


Figure 4 : Modèle des discontinuités de luminance.

Nous pouvons remarquer que quand ϵ tend vers 0, l'image floue représente la réponse indicielle du système optique, et sa dérivée donne la réponse impulsionnelle. De façon générale, les gradients constituent une mesure du flou très utilisée [PIE-95][BEN-94][SIM-93][PEN-87][KHO-95]. Notre méthode s'appuie sur la recherche des discontinuités unidirectionnelles, l'équation de convolution se réduit alors à l'expression suivante (6) par la séparabilité linéaire du modèle

de la réponse impulsionnelle, $h_i(x,y) = \theta_i(x)\theta_i(y)$; pour les discontinuités en x :

$$if_i(x,y) = \theta_i(x) \otimes_{1D} in_i(x,y) \quad (6)$$

La détermination du flou optique impose que la dérivée du modèle de la discontinuité de luminance soit paire et de limites nulles aux extrémités de son support. Le modèle retenu vérifie ces conditions, cependant la technique proposée peut utiliser tout autre modèle respectant ces contraintes.

3. Détermination du flou optique

La méthode consiste à calculer sur chaque discontinuité le gradient normé selon x :

$$g(x) = \frac{\frac{\partial if_i(x,y)}{\partial x}}{\int_{-\infty}^{+\infty} \frac{\partial if_i(x,y)}{\partial x} dx} \quad (7)$$

avec un support d'intégration limité à une fenêtre d'analyse de largeur minimale $6\sigma_{\max}$ pour le plus grand flou détectable. La variance σ_g^2 de $g(x)$ vaut alors $\sigma_g^2 = \sigma_i^2 + \sigma_n^2$ avec

$$\sigma_n^2 = \frac{1}{b} \int_{-\infty}^{+\infty} x^2 \frac{\partial in_i(x,y)}{\partial x} dx, \quad (8)$$

où b est l'amplitude de la discontinuité et σ_i l'écart-type de $\theta_i(x)$ que nous recherchons. En utilisant une deuxième image floue prise avec une ouverture différente, on obtient une deuxième équation pour former un système (9) à trois inconnues.

$$\begin{cases} \sigma_{g,1}^2 = \sigma_{i,1}^2 + \sigma_n^2 \\ \sigma_{g,2}^2 = \sigma_{i,2}^2 + \sigma_n^2 \end{cases} \quad (9)$$

Pour résoudre ce système il faut générer une autre équation sans faire apparaître d'autres inconnues. La phase de calibration permet d'obtenir deux équations (10) liées aux deux ouvertures utilisées.

$$\begin{cases} \sigma_{i,1} = m_1 Z_i^{-1} + c_1 \\ \sigma_{i,2} = m_2 Z_i^{-1} + c_2 \end{cases} \quad (10)$$

$$\text{d'où : } \sigma_{i,2}^2 = \frac{\beta}{\alpha^2 - 1} \text{ et } \sigma_{i,1}^2 = \frac{\alpha^2}{\alpha^2 - 1} \beta \quad (11)$$

$$\text{avec } \beta = \sigma_{g,1}^2 - \sigma_{g,2}^2 \text{ et } \alpha = \frac{m_1}{m_2} \quad (12)$$

Pour obtenir les profondeurs des contours d'objets de la scène, la méthode procède en trois étapes. Après la phase de

calibration et l'acquisition de deux images floues, la première étape détermine la position des contours, la seconde calcule les gradients normés des images, enfin la troisième utilise les relations (10)(11)(12) pour calculer les profondeurs. Ces opérations simples offrent la possibilité d'une réalisation à la cadence vidéo sur du matériel spécialisé de traitement d'images.

4. Calibration

La phase de calibration consiste en l'acquisition de deux images d'un objet présentant une seule discontinuité et placé à différentes profondeurs connues. Une image floue correspondant à une ouverture choisie, et une image nette en tout point, acquise en fermant au plus petit l'iris, permettent d'évaluer σ_g^2 et σ_n^2 [BEN-94]. Pour l'image nette, seuls les rayons paraxiaux forment l'image comme le montre la figure 5; la quantité de lumière reçue sur la capteur étant très diminuée on effectue un sur-éclairage ou une augmentation du temps d'intégration. L'ensemble des couples (σ_i, Z_i) permet de calculer par régression linéaire les constantes m et c . L'application de cette phase de calibration pour les deux ouvertures utilisées précédemment permet d'évaluer les constantes m_1, c_1, m_2 et c_2 .

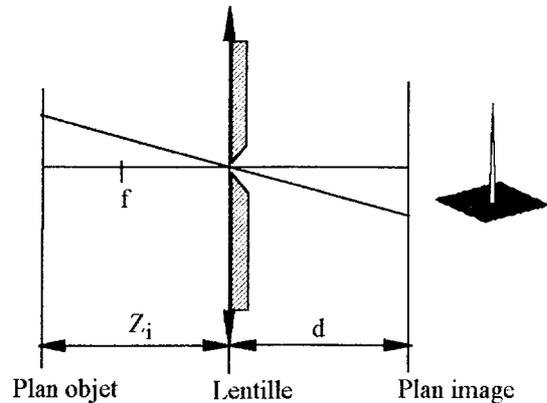


Figure 5 : Image nette en tout point.

5. Résultats

Les figures 6 et 7 présentent les deux images floues d'une scène composée de cinq cibles placées à des profondeurs connues. Le plan net pour ces images est placé à 250 cm afin d'exploiter la plus grande sensibilité flou-éloignement. La phase de calibration a conduit aux valeurs suivantes des constantes : $m_1 = 923 \text{ cm} \cdot \text{pixel}$, $c_1 = -3.71 \text{ pixel}$, $m_2 = 421 \text{ cm} \cdot \text{pixel}$, et $c_2 = -1.73 \text{ pixel}$, donnant un plan net à 245 cm. La figure 8 montre les résultats obtenus sur des contours issus d'une segmentation par filtre optimal à réponse impulsionnelle infinie opérant une dérivation. La carte des profondeurs obtenue sur les contours d'objets est exploitable en robotique mobile ou d'assemblage, et en reconnaissance de forme avec une précision inférieure à 5%.

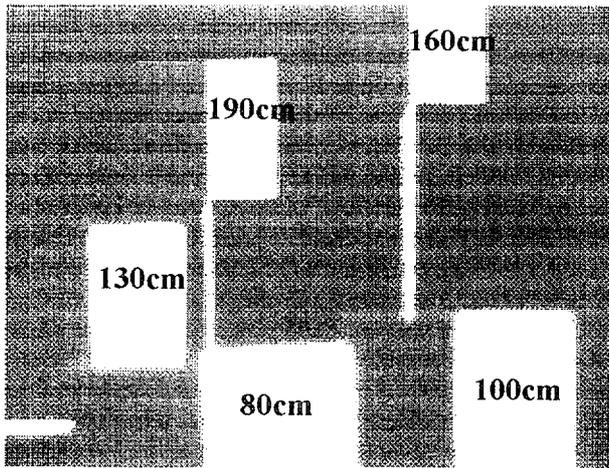


Figure 6 : Image floue 1 : grande ouverture de l'iris, profondeurs réelles.

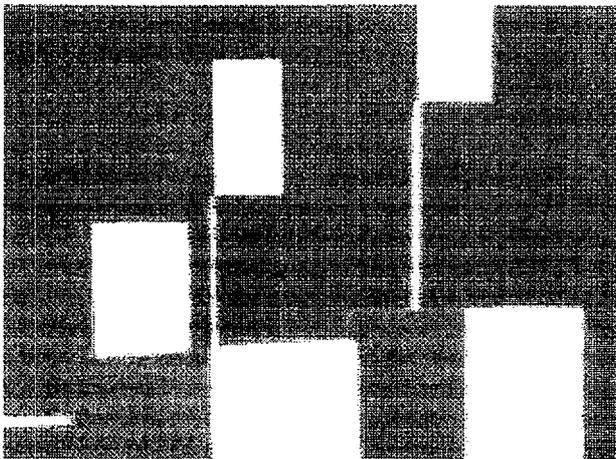


Figure 7 : Image floue 2 : petite ouverture de l'iris.

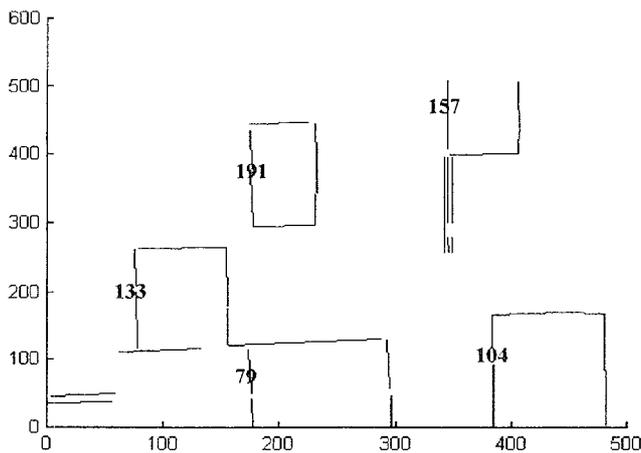


Figure 8 : Image des contours : profondeurs calculées.

[PEN-87] A.P.PENTLAND, "A New Sense for Depth of Field", IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, pp 523-531.

[SIM-91] T.SIMON, "Appréciation de la profondeur depuis des images nette et floue en vision 3D passive", Colloque l'Interface des mondes réels et virtuels, Montpellier 23-27 mars 1993.

[PIE-95] F. PIETRZAK, Th. SIMON, J. BERNARD, B. HEIT, "Stéréovision monoculaire passive: Amélioration de la méthode Jarvis", 2ième Conférence Internationale sur l'Automatisation Industrielle, Nancy 7-9 juin 1995.

[KHO-95] M. KHOUDEIR, Th. SIMON, A. BENZOUGAR, J. BERNARD, "Stéréovision passive : coopération entre une approche monoculaire et une approche binoculaire", 1er colloque de Productique- Robotique SEA, Bourges 1-2 juin 1995, pp 193-198.

Bibliographie :

[BEN-94] A. BENZOUGAR, Th. SIMON, M. KHOUDEIR, J. BERNARD "Passive Monocular Stereovision, Extension of the Subbarao Method", IEEE Proceeding MECHATRONICS, The 2nd Japan-France Congress on Mechatronics, Takamatsu, Kagawa, Japan, 1994, pp693-696