



Estimation de la profondeur par évaluation du flou optique sur les contours

Christophe Simon, Frédérique Bicking, Thierry Simon

► To cite this version:

Christophe Simon, Frédérique Bicking, Thierry Simon. Estimation de la profondeur par évaluation du flou optique sur les contours. ORASIS 2001, Jun 2001, Cahors, France. pp.CDROM. hal-00331814

HAL Id: hal-00331814

<https://hal.science/hal-00331814>

Submitted on 17 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Estimation de la profondeur par évaluation du flou optique sur les contours

Christophe Simon*, Frédérique Bicking*, Thierry Simon**

*Université Henri Poincaré Nancy 1,

Ecole Supérieure des Sciences et Technologies de l'Ingénieur de Nancy

2 Rue Jean Lamour, 54519 Vandœuvre-les-Nancy Cedex

Tél : 03 83 50 16 34

Fax : 03 83 54 21 73

Christophe.Simon@esstin.uhp-nancy.fr ; Frederique.Bicking@esstin.uhp-nancy.fr

**IUT de Figeac, Avenue de Nayrac, 46100 Figeac

Tél : 05 65 50 30 68

Fax : 05 65 50 30 61

Simon.IUT.Figeac@wanadoo.fr

Mots clés : *depth from defocus*, contours, flou optique, primitives 3D.

Résumé

Nous présentons dans cet article la généralisation d'une méthode d'estimation locale de la profondeur sur les contours utilisant un couple d'images nette et floue. Après avoir positionné cette méthode de *depth from defocus*, initialement développée par T. Simon (Simon 1991), nous définissons les relations théoriques mettant en œuvre cette généralisation. De plus, nous apportons une solution à la critique de la sensibilité aux variations locales de luminance souvent formulée à l'égard de cette approche. Puis, nous quantifions l'amélioration apportée par l'application de la méthode à des images de synthèse parfaitement maîtrisées. Enfin, nous appliquons la méthode à une scène réelle et nous précisons les points à développer pour une bonne applicabilité dans un contexte réel.

1 Introduction

La profondeur, distance entre la surface visible des objets dans une scène et le capteur de la caméra, est une donnée nécessaire pour le calcul des coordonnées des points appartenant à cette surface dans un espace tridimensionnel de référence. Ces informations 3D sont des indices précieux pour la reconnaissance des objets et l'interprétation de scènes. De nombreuses méthodes ont été développées pour obtenir ces coordonnées. Elles reposent toutes sur une exploitation des variations des conditions de prises de vue d'une ou plusieurs images. Les paramètres du système d'acquisition ou de l'environnement lumineux, contrôlé ou non, sont les informations essentielles à l'établissement de la relation entre l'image et la scène réelle.

Notre méthode s'inscrit dans l'ensemble des techniques utilisant le flou optique comme source de l'information de profondeur à partir d'un seul point de vue. Ces techniques se composent de deux groupes. Le premier recherche dans l'image les zones de netteté et les lie à la profondeur par les relations classiques de l'optique géométrique. Nommées *Depth From Focus* ou DFF (Grossman 1987 ; Darell 1990 ; Subbarao 1993 ; Krotkov 1987), elles présentent l'inconvénient majeur d'un temps prohibitif d'obtention d'une carte des profondeurs. Le second groupe exploite le flou optique perceptible sur les zones non homogènes de l'image correspondant à des contours ou à des textures. Ces techniques, appelées *Depth From Defocus* ou DFD (Pentland 1987 ; Subbarao 1988 ; Ens 1991 ; Subbarao 1994 ; Pham 1999), utilisent au moins deux images d'une même scène acquises avec des réglages différents des paramètres du système optique.

Notre technique utilise uniquement deux images d'une même scène, acquises pour l'une avec un iris très fermé, donnant une profondeur de champ très grande et donc des images nettes d'objets situés à des profondeurs différentes. L'autre image, acquise avec un iris plus ouvert, donne une profondeur de champ très réduite et une quantité de flou dépendant de la profondeur des objets de la scène. Le flou optique est évalué sur les discontinuités de luminance correspondant le plus souvent à des contours d'objets.

Notre article rappelle dans un premier temps, à partir des relations de l'optique géométrique et de la modélisation du processus de formation des images, le lien entre la profondeur et le flou optique. Puis, nous précisons le modèle des discontinuités de luminance utilisé et exposons les relations issues de l'algorithme initialement posé par T. Simon (Simon 1991). Enfin, nous proposons une amélioration reposant sur une généralisation du calcul du rapport des modules des gradients des images nette et floue, à la totalité des points de la discontinuité dans la direction de son gradient. Cette amélioration conduit à une meilleure précision et à une meilleure robustesse de l'évaluation de la profondeur.

2 Le flou optique porteur de l'information de profondeur

Dans le groupe des méthodes relevant du *depth from defocus*, l'effet physique produit sur les images par la modification des paramètres d'acquisition tels que la distance focale, la mise au point et l'ouverture du diaphragme est exploité. On cherche à établir une ou plusieurs relations entre la profondeur, les paramètres de la caméra et le flou optique dans les images. L'objet de ce paragraphe est de montrer comment s'obtient la relation qui lie le flou optique et la profondeur.

2.1 Défocalisation

Pour assurer correctement la formation d'une image sur un plan, plusieurs lentilles forment généralement les objectifs photographiques. Pour des raisons de simplification, nous assimilons l'objectif de la caméra à une lentille convergente mince unique et restons dans les hypothèses de Gauss. Pour une caméra munie d'une lentille de distance focale f , la relation entre la position d'un point objet dans la scène et la position de son image focalisée est donnée par la formule de conjugaison (1) :

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f} \quad (1)$$

où s_o est la distance entre le plan principal (plan lentille) et le plan objet et s_i la distance entre le plan principal et le plan image conjugué.

La figure 1 montre le processus géométrique de formation des images. Tous les rayons lumineux émis par un objet O et interceptés par la lentille sont réfractés pour converger en un point du plan image conjugué. Chaque point dans la scène (espace objet réel) est projeté en son point conjugué sur le plan image conjugué formant sur celui-ci une image focalisée ou nette. Si le plan capteur ne coïncide pas avec le plan image conjugué, l'image formée sera une tache appelée "cercle de confusion". Ainsi, le point image devient un disque de diamètre D si l'ouverture du diaphragme de l'objectif de la caméra est circulaire.

Ce diamètre D dépend de δ (Fig. 1) et de son signe (Fig. 2). Dans le cas où $\delta \geq 0$, le plan capteur de la caméra est derrière le plan image conjugué et si $\delta < 0$, il est devant. La figure 2 montre l'évolution du diamètre D du cercle de confusion en fonction de la distance du plan objet s_o , pour une distance plan de la lentille – plan capteur s fixée à 70 mm, une longueur focale $f = 60$ mm et pour différentes ouvertures du diaphragme L .

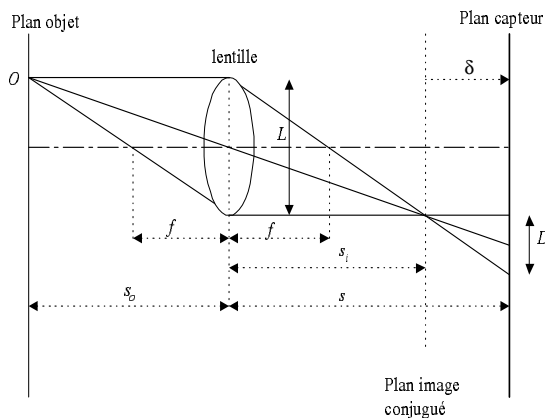


Fig. 1 : principe de formation d'une image.

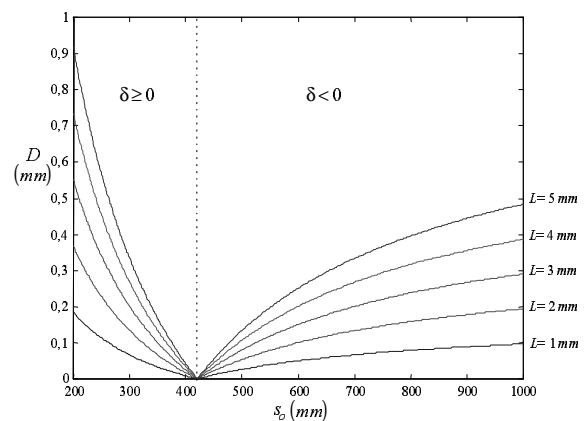


Fig. 2 : évolution du diamètre du cercle flou en fonction de s_o pour différentes valeurs de L .

A un plan image à une distance s_i fixée et pour une scène contenant plusieurs plans en profondeur, un seul plan donne une image nette, dite focalisée. Les points objets appartenant aux autres plans formeront des taches plus ou moins floues en fonction de leur distance par rapport au plan image. Le changement de signe de δ conduit à une estimation non unique du flou optique donc de la profondeur puisque à une valeur de D peuvent correspondre deux valeurs de s_o . Pour lever cette ambiguïté qui peut apparaître dans une scène comportant des objets à différentes profondeurs, il suffit de placer le plan focal objet soit en avant-plan, soit en arrière-plan de la scène. Ce choix sera guidé par un compromis à faire entre la plage de profondeurs à observer et l'évolution du flou optique. En effet, en référence à la figure 2, pour $\delta < 0$, on constate une asymptote horizontale qui limite la variation du flou mais sur une plage de profondeurs plus grande que pour $\delta \geq 0$, pour lequel la plage de profondeurs est limitée par la position de la caméra avec en revanche une évolution du flou plus rapide.

A partir des relations de l'optique géométrique, la profondeur recherchée s_o peut s'exprimer par la relation (2) :

$$\text{Pour } \delta \geq 0 : s_o = \frac{fs}{s - f - D \frac{f}{L}} \quad \text{Pour } \delta < 0 : s_o = \frac{fs}{s - f + D \frac{f}{L}} \quad (2)$$

Ces expressions sont à la base de toutes les méthodes qui utilisent le flou optique comme information de profondeur. Le lien entre le diamètre du cercle de confusion, caractéristique du flou optique, et l'information de profondeur exprimant la distance plan objet – plan principal est alors immédiat.

En général, on ne mesure pas la taille de la tache car sa distribution d'intensité lumineuse est intimement liée au comportement ondulatoire du système optique. Le modèle utilisé pour représenter cette tache est une approximation de la fonction de transfert du système optique.

2.2 Modèle de réponse d'un système optique

La formation du flou optique est liée à la fonction de transfert du système optique (Williams 1989) dans le domaine spatial représentée par sa réponse impulsionnelle dite *Point Spread Function* ou PSF. Ainsi, une relation de convolution (3) s'établit entre l'image nette $I_n(i, j)$, l'image floue $I_f(i, j)$ et la $PSF(i, j)$ où $\otimes_{2D}$ est l'opérateur de convolution bidimensionnelle et i et j représentent les coordonnées d'un pixel :

$$I_f(i, j) = I_n(i, j) \otimes_{2D} PSF(i, j) \quad (3)$$

La PSF dépend autant des propriétés des matériaux de l'optique (indice de réfraction) que de la forme géométrique de la lentille (distance focale) et des paramètres de prise de vue (distance du plan objet au plan principal, diaphragme, éclairage). Un modèle réaliste prenant en compte à la fois tous les aspects de l'optique géométrique, les effets de la diffraction et les défauts des lentilles, n'existe pas. Plusieurs modèles ont été proposés et nous avons retenu le modèle de distribution de luminance donné par une gaussienne bidimensionnelle avec un paramètre d'étalement σ_{s_o} correspondant à une profondeur s_o . Ce modèle est donné par la relation (4) :

$$PSF_{s_o}(i, j) = \frac{1}{2\pi\sigma_{s_o}^2} \exp\left(-\frac{i^2 + j^2}{2\sigma_{s_o}^2}\right) \quad (4)$$

2.3 Estimation de la profondeur

La relation entre le paramètre d'étalement σ_{s_o} et le diamètre D du cercle de confusion est donnée par la relation (5) :

$$\sigma_{s_o} = kD \quad (5)$$

où k est une constante de proportionnalité caractéristique d'une caméra donnée, déterminée expérimentalement au cours d'une phase de calibrage (Subbarao 1989). A partir de cette relation, la

profondeur peut être évaluée par la relation (2). Cependant, les paramètres s , f et L ne sont pas connus avec précision et peuvent induire des erreurs sur la détermination de s_o . Aussi, dans le cas $\delta < 0$ (cf. figure 2), la profondeur peut être obtenue par la relation (6) issue des relations (2) et (5) :

$$\frac{1}{s_o} = \frac{\sigma_{s_o}}{m} - \frac{c}{m} \quad (6)$$

où c et m sont des constantes englobant s , L , f et k représentant un réglage spécifique de la caméra. Ces constantes c et m sont déterminées au cours de la phase de calibrage.

3 Méthode initiale pour l'estimation du paramètre de flou optique

Dans les approches classiques d'estimation de la profondeur, les techniques principales exploitent soit le contenu spatial de l'image par les caractéristiques géométriques (Subbarao 1988-2 ; Pentland 1982-1987) ou la forme des objets dans la scène (Darell 1990), soit le contenu fréquentiel par une analyse de Fourier (Subbarao 1988-1).

3.1 Utilisation des contours

Nous avons montré que l'information sur la profondeur existait à travers le flou optique. Il est donc nécessaire d'identifier certaines caractéristiques de l'image pour lesquelles les effets du flou sont perceptibles comme par exemple des discontinuités de luminance qui peuvent être les contours des objets observés. Afin d'appliquer la méthode, il est nécessaire de connaître la position des points de contours.

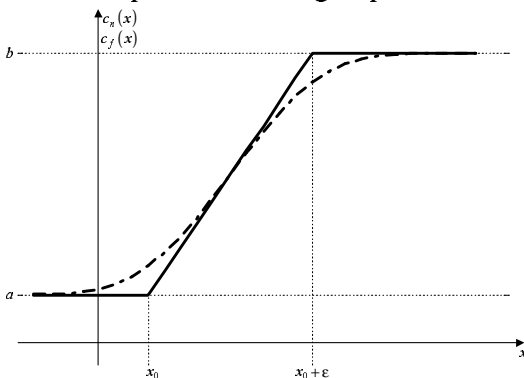
Le principe général de la méthode est similaire à celui de la méthode proposée par Pentland (1982-1987-1989). On effectue l'acquisition d'une image nette en tout point avec un iris fermé qui fournit facilement la position des contours grâce à un opérateur de gradient lisse comme celui de Prewitt. En faisant l'acquisition d'une seconde image, floue (iris ouvert), le contenu informationnel de la scène est isolé. En établissant le rapport des modules des gradients des images nette et floue qui permet de donner la même importance à chaque contour, il est possible d'identifier le paramètre d'étalement σ_{s_o} et ainsi d'estimer la profondeur par la relation (6).

3.2 Identification du paramètre d'étalement

Le rapport des modules des gradients des deux images nette et floue de la même scène s'écrit selon la relation (7) :

$$R(i, j) = \frac{|\nabla I_n(i, j)|}{|\nabla I_f(i, j)|} \quad (7)$$

Comme nous travaillons à partir des profils des contours dans l'image, le problème peut être analysé sous une forme monodimensionnelle en opérant selon une orientation horizontale ou verticale. On considère un contour en rampe, qui fait partie des formes les plus rencontrées dans une image (Monga 1990). On note $c_n(x)$ la fonction de niveau de gris dans la direction d'une normale au contour net de l'objet où x , gradué en pixels, représente une ligne ou une colonne de l'image. Son expression est donnée par la relation (8) où $(b - a)$ est l'amplitude de la rampe de longueur ϵ . Elle est représentée en ligne pleine sur la figure 3.



$$c_n(x) = \begin{cases} a & x < x_0 \\ a + \left(\frac{b-a}{\epsilon} \right) (x - x_0) & x_0 \leq x \leq x_0 + \epsilon \\ b & x > x_0 + \epsilon \end{cases} \quad (8)$$

Fig. 3 : profils des contours net et flou.

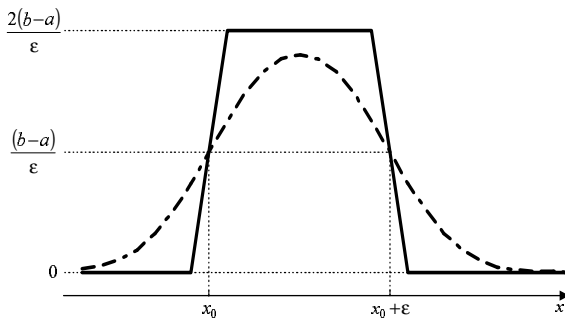
On note $c_f(x)$ la fonction de niveau de gris pour le contour flou, dans la même direction que $c_n(x)$, représentée en pointillés sur la figure 3. Elle est obtenue par convolution de $c_n(x)$ avec l'expression monodimensionnelle de la PSF. L'expression de $c_f(x)$ est donnée par la relation (9) :

$$c_f(x) = c_n(x) \otimes PSF(x) \quad \text{où } PSF(x) = \frac{1}{\sigma_{s_o} \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_{s_o}^2}\right) \quad (9)$$

Le gradient du contour net $\nabla c_n(x)$, représenté par une ligne pleine sur la figure 4, est obtenu à partir de la convolution du contour net et d'une forme monodimensionnelle de l'opérateur gradient issu de celui de Prewitt :

$$\nabla c_n(x) = c_n(x) \otimes [-1 \ 0 \ 1] \quad (10)$$

Le gradient du contour net s'exprime par la relation (11) :



$$\nabla c_n(x) = \begin{cases} 0 & x < x_0 \\ \frac{b-a}{\varepsilon} & x = x_0 \\ 2\left(\frac{b-a}{\varepsilon}\right) & x_0 < x < x_0 + \varepsilon \\ \frac{b-a}{\varepsilon} & x = x_0 + \varepsilon \\ 0 & x > x_0 + \varepsilon \end{cases} \quad (11)$$

Fig. 4 : gradients des contours net et flou.

Le gradient du contour flou, représenté en pointillés sur la figure 4, est obtenu comme précédemment par convolution avec l'expression de la PSF monodimensionnelle.

$$\nabla c_f(x) = \nabla c_n(x) \otimes PSF(x) \quad (12)$$

Le rapport des modules des gradients des images nette et floue, donné par la relation (7), prend des valeurs exploitables en fonction des valeurs de la longueur de la rampe ε . Pour des valeurs entières de ε constituant les cas étudiés dans (Simon 1991), le rapport exprimé aux seuls deux points x_0 et $x_0 + \varepsilon$ présente l'une des formes analytiques simples suivantes :

Pour $\varepsilon = 1$,

$$R(x_0) = R(x_0 + \varepsilon) = \frac{\sigma_{s_o} \sqrt{2\pi}}{1 + \exp\left(-\frac{1}{2\sigma_{s_o}^2}\right)}$$

Pour $\varepsilon = 2$,

$$R(x_0) = R(x_0 + \varepsilon) = \frac{\sigma_{s_o} \sqrt{2\pi}}{1 + 2\exp\left(-\frac{1}{2\sigma_{s_o}^2}\right) + \exp\left(-\frac{4}{2\sigma_{s_o}^2}\right)}$$

Pour $\varepsilon \geq 3$, $R(x_0) = R(x_0 + \varepsilon) = \frac{\sigma_{s_o} \sqrt{2\pi}}{1 + 2\exp\left(-\frac{1}{2\sigma_{s_o}^2}\right) + 2\exp\left(-\frac{4}{2\sigma_{s_o}^2}\right) + \exp\left(-\frac{9}{2\sigma_{s_o}^2}\right)}$ (13)

Ces relations, liant le paramètre d'étalement à la valeur du rapport des modules des gradients, sont résolues numériquement. Le rapport des modules des gradients au point x_0 permet de trouver une valeur du paramètre d'étalement σ_{s_o} et d'obtenir une profondeur s_o grâce à la relation (6).

4 Critique et amélioration de la méthode

La méthode précédemment décrite a été la base de plusieurs travaux débouchant sur l'utilisation de deux images floues (Benzougar 1995) ou de la texture des images comme élément caractéristique (Benzougar 1999). Nous avons choisi de conserver la méthode initiale et d'en lever certaines limitations.

Le premier point que nous relevons sur cette méthode est l'invalidité de l'expression du rapport $R(x_0)$ pour des longueurs de profils de discontinuités ε strictement supérieures à 3. En étudiant les gradients des profils des discontinuités avec des valeurs de ε plus importantes, on obtient une expression plus générale du module du gradient de l'image floue et donc du rapport des modules des gradients des images nette et floue. Le module du gradient de l'image floue est obtenu théoriquement par la relation (12). L'étude de la convolution du gradient du contour net en rampe entre x_0 et $x_0 + \varepsilon$ (Fig. 4) par la PSF monodimensionnelle, en chaque point, nous permet d'exprimer le gradient flou en tout point du contour.

Pour chaque point x représentant dans l'image nette un point du contour, la relation (12) s'écrit :

$$\text{Pour } \varepsilon = 1 : \quad \nabla c_f(x_0) = \left(\frac{b-a}{\varepsilon} \right) [PSF(0) + PSF(-1)]$$

$$\nabla c_f(x_0 + \varepsilon) = \left(\frac{b-a}{\varepsilon} \right) [PSF(0) + PSF(1)]$$

$$\text{Pour } \varepsilon = 2 : \quad \nabla c_f(x_0) = \left(\frac{b-a}{\varepsilon} \right) [PSF(0) + 2PSF(-1) + PSF(-2)]$$

$$\nabla c_f(x_0 + 1) = \left(\frac{b-a}{\varepsilon} \right) [PSF(1) + 2PSF(0) + PSF(-1)]$$

$$\nabla c_f(x_0 + \varepsilon) = \left(\frac{b-a}{\varepsilon} \right) [PSF(2) + 2PSF(1) + PSF(0)]$$

$$\text{Pour } \varepsilon = 3 : \quad \nabla c_f(x_0) = \left(\frac{b-a}{\varepsilon} \right) [PSF(0) + 2(PSF(-1) + PSF(-2)) + PSF(-3)]$$

$$\nabla c_f(x_0 + 1) = \left(\frac{b-a}{\varepsilon} \right) [PSF(1) + 2(PSF(0) + PSF(-1)) + PSF(-2)]$$

$$\nabla c_f(x_0 + 2) = \left(\frac{b-a}{\varepsilon} \right) [PSF(2) + 2(PSF(1) + PSF(0)) + PSF(-1)]$$

$$\nabla c_f(x_0 + \varepsilon) = \left(\frac{b-a}{\varepsilon} \right) [PSF(3) + 2(PSF(1) + PSF(2)) + PSF(0)]$$

En continuant les développements pour d'autres valeurs de ε , l'expression du module du gradient du contour flou peut être généralisée par la relation (14).

$$\forall \varepsilon \in N_*^+ \quad \nabla c_f(x) = \left(\frac{b-a}{\varepsilon} \right) \left[PSF(x - x_0) + PSF(x - (x_0 + \varepsilon)) + 2 \sum_{u=1}^{\varepsilon-1} PSF(x - (x_0 + u)) \right] \quad (14)$$

Le rapport des modules des gradients des images nette et floue s'écrit alors selon la relation (15) :

$$R(x) = \begin{cases} \frac{1}{PSF(x - x_0) + PSF(x - (x_0 + \varepsilon)) + 2 \sum_{u=1}^{\varepsilon-1} PSF(x - (x_0 + u))} & \text{pour } x = x_0 \text{ et } x = x_0 + \varepsilon \\ \frac{2}{PSF(x - x_0) + PSF(x - (x_0 + \varepsilon)) + 2 \sum_{u=1}^{\varepsilon-1} PSF(x - (x_0 + u))} & \text{pour } x_0 < x < x_0 + \varepsilon \end{cases} \quad (15)$$

Comme dans la méthode initiale, ces relations sont résolues au point x_0 pour déterminer la valeur du paramètre d'étalement σ_{s_o} . Cependant, l'estimation de σ_{s_o} est sensible à la valeur du rapport des modules des gradients au point considéré. Cela constitue le deuxième point faible de la méthode précédemment proposée.

L'idée d'amélioration est d'utiliser tous les points du contour $x \in [x_0, x_0 + \varepsilon]$ comme support de l'estimation afin de construire une carte des paramètres d'étalement et d'en déduire statistiquement une valeur représentative des estimations obtenues. La méthode se trouve un peu alourdie en temps de calcul mais l'estimation est globalement meilleure.

En résumé, le premier apport est le calcul du rapport des modules des gradients avec une forme générale du gradient de l'image floue qui permet de prendre en compte toutes les valeurs entières possibles pour ε . Le second apport est une estimation plus robuste du paramètre d'étalement σ_{s_o} car elle est effectuée sur tous les points de contours et non plus sur un seul.

5 Relation flou-éloignement : calibrage

Comme nous l'avons précisé au paragraphe 2.3, la relation flou-éloignement est obtenue par une phase de calibrage qui permet l'estimation des paramètres caractéristiques m et c de la caméra introduits dans la relation (6).

Il existe plusieurs techniques de calibrage. Nous avons opté pour l'approche classique dite "point par point" qui consiste à placer une discontinuité unidirectionnelle à plusieurs profondeurs. Les couples d'images nettes et floues obtenus sont utilisés pour l'estimation de σ_{s_o} afin d'établir les couples $(\hat{\sigma}_{s_o}, 1/s_o)$ correspondants où $\hat{\sigma}_{s_o}$ est l'estimation de σ_{s_o} pour une profondeur s_o . Par régression, nous obtenons finalement les constantes m et c spécifiques aux réglages de la caméra utilisée. Cette approche, lourde en temps de calcul, est réalisée préalablement à l'utilisation de la méthode d'évaluation de la profondeur sur des scènes réelles. Elle permet, pour les images de calibrage, de faire une estimation de σ_{s_o} sur chaque profil des contours analysé et de fournir une cartographie de $\hat{\sigma}_{s_o}$ pour chacune des profondeurs. L'imprécision des mesures, liée aux conditions d'acquisition (bruit, variation lumineuse), est prise en compte lors de l'estimation de m et c .

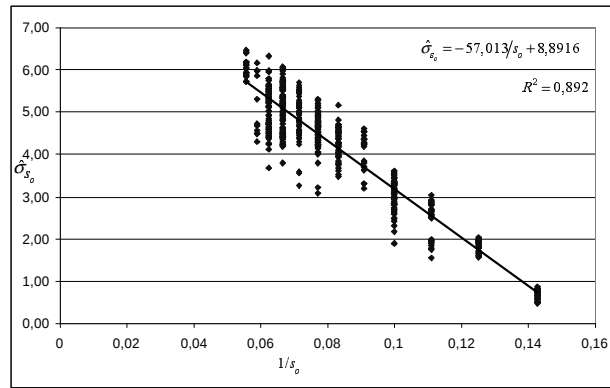


Fig. 5 : courbe de calibrage.

Le calibrage a été réalisé avec une caméra JAI CV M1 *progressive scan* et une carte d'acquisition PCI 1408 de National Instruments. Les images sont de taille 640x480 pixels codés sur 8 bits. L'optique de la caméra est un objectif industriel TAMRON 2/3 de pouces 25 mm. Les images de la discontinuité sont acquises sous une lumière blanche circulaire en environnement intérieur. Néanmoins, aucun autre contrôle de l'éclairage n'est réalisé.

Cette phase nous a permis d'obtenir, sur une plage de valeurs de la profondeur restreinte (6-17 cm), la relation flou-éloignement suivante : $\hat{s}_o^{-1} = -17,5 \cdot 10^{-3} \hat{\sigma}_{s_o} + 15,6 \cdot 10^{-1}$ (Fig. 5).

6 Essais et résultats expérimentaux

Pour juger de la validité des apports, nous avons réalisé plusieurs essais, d'une part à l'aide d'images de synthèse et d'autre part avec des images réelles. Les images de synthèse ont été générées de manière à évaluer les améliorations apportées et mettre en évidence les incapacités actuelles de la méthode de perception de profondeur.

6.1 Validation numérique pour des variations de ε et de σ_{s_o} .

Cette expérience a pour objectif de confirmer l'amélioration apportée par la généralisation de la méthode initiale. Nous avons généré une série d'images de synthèse représentant une discontinuité verticale ou horizontale avec différentes valeurs de ε pour l'image nette. Une série d'images floues est construite artificiellement avec différentes valeurs σ_{s_o} grâce aux relations (3) et (4).

La figure 6 montre les erreurs d'estimation de σ_{s_o} pour les deux méthodes. Sur la figure 6a (méthode initiale), nous constatons que les erreurs sont faibles si $\varepsilon \leq 3$ et $\sigma_{s_o} \geq 1,0$. Si $\sigma_{s_o} < 1,0$, l'estimation est d'autant moins performante que σ_{s_o} tend vers zéro. Lorsque $\varepsilon > 3$, la méthode initiale fournit des estimations d'autant plus erronées que les valeurs de σ_{s_o} et ε sont grandes.

Pour la méthode généralisée (Fig. 6b), les erreurs d'estimation se situent entre $-1,4$ et $+0,6$ pour les mêmes expériences. Cette diminution des erreurs est due à la généralisation du calcul du rapport pour ε quelconque et ceci, en chaque point $x \in [x_0, x_0 + \varepsilon]$ du contour. La moyenne des $\varepsilon + 1$ estimations ainsi fournies nous permet de diminuer la sensibilité de l'estimation à la valeur du rapport en chaque point du contour.

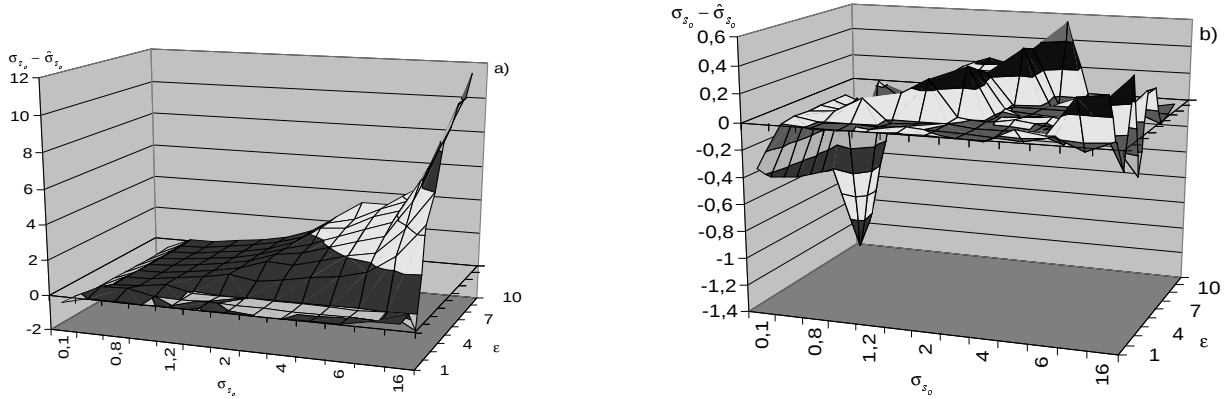


Fig. 6 : erreur d'estimation de σ_{s_o} en fonction de ε et σ_{s_o} .

a) Méthode initiale.

b) Méthode généralisée.

Cette dernière remarque soulève le problème de la précision de cette approche. Rappelons que ce n'est pas la précision théorique qui est en cause (Schechner 1998) mais la précision pratique que l'on peut améliorer à l'aide d'un capteur CCD disposant de plus de niveaux de quantification et par un changement plus significatif du jeu de réglages de la caméra.

6.2 Image de synthèse

L'application des méthodes à des images de synthèse disposant de contours horizontaux et verticaux permet de valider l'approche sur des objets simples et maîtrisés. Lorsque l'on expérimente les méthodes sur une forme carrée comme cela est montré sur les figures 7a et 7b avec $\sigma_{s_o} = 2,0$, l'estimation de σ_{s_o} présente des résultats similaires à ceux de l'expérience précédente. Néanmoins, les $\hat{\sigma}_{s_o}$ sont surévalués dès lors que les contours sont proches les uns des autres. Ceci est effectivement dû au fait que le principe de convolution opéré par l'optique tient compte de contours même éloignés et que le développement théorique sur lequel reposent les méthodes suppose que la discontinuité étudiée ne subit pas de contribution des autres contours et est homogène dans la direction orthogonale au contour. Nous observons ce phénomène sur la figure 7c où $\hat{\sigma}_{s_o}$ augmente à la proximité des coins de l'objet car l'estimation tient compte de la contribution du contour qui lui est proche perpendiculairement. Sur le pixel correspondant au coin de l'objet, l'estimation est sous-évaluée en raison de la valeur du module du gradient de Prewitt de l'image nette. Ce phénomène de contribution mutuelle des contours est d'autant moins significatif que σ_{s_o} est petit et la distance entre les contours relativement grande. La même analyse peut être faite sur des contours d'objets différents, à des distances différentes.

On peut encore noter l'intérêt du calcul de la moyenne des $\hat{\sigma}_{s_o}$ à partir de chaque point du contour, élément de $[x_0, x_0 + \varepsilon]$, puisque à proximité des coins, cette valeur est généralement inférieure à celle de l'estimation de σ_{s_o} au point x_0 .

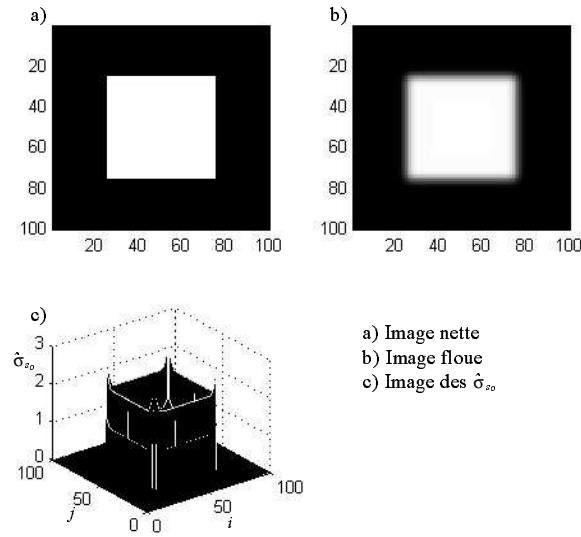


Fig. 7 : illustration de l'incidence des contours voisins sur l'estimation de σ_{s_o} .

6.3 Image réelle

Les contours sur des images réelles sont beaucoup moins homogènes que ceux d'images de synthèse, ce qui entraîne une disparité des valeurs estimées $\hat{\sigma}_{s_o}$. Le calibrage est déjà caractéristique de ce phénomène.

La méthode proposée est bien évidemment sensible au bruit notamment lorsque l'on calcule $\hat{\sigma}_{s_o}$ au point x_0 uniquement. La figure 8 montre que l'estimation de σ_{s_o} sur une scène réelle, pour des orientations horizontales et verticales uniquement, implique une forte réduction du nombre de points de calcul. Par ailleurs, la dispersion des résultats est importante. En utilisant les équations généralisées aux points de contours dans l'intervalle $[x_0, x_0 + \varepsilon]$, même s'il n'y a pas plus de points d'estimation, les résultats présentent moins de dispersion. La contrainte d'orientation reste une contrainte forte sur les conditions d'application de la méthode.

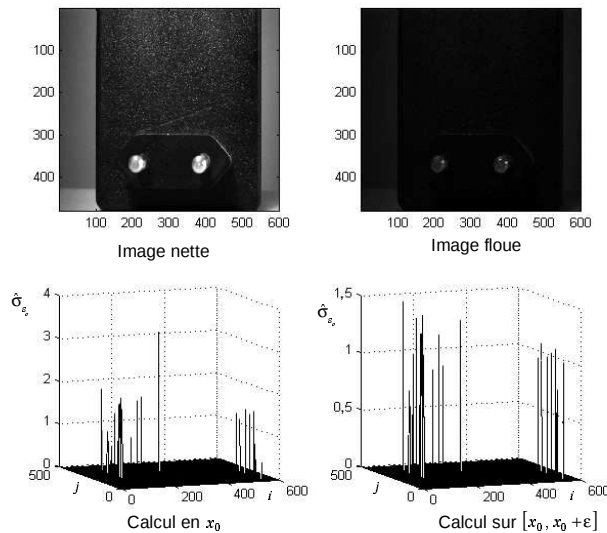


Fig. 8 : application à une scène réelle.

7 Conclusion et perspectives des travaux

Dans cet article, nous avons généralisé la méthode de perception locale de la profondeur sur les contours initialement développée par T. Simon (Simon 1991) en étendant son calcul à toutes les tailles de contours en rampe, ce qui permet de l'appliquer à des contours épais. De plus, nous avons

généralisé le calcul de σ_{s_0} en chaque point de la rampe, ce qui permet d'obtenir plusieurs estimations de la profondeur sur le même contour et ainsi de diminuer l'influence des variations locales des données images aux points de contours. L'estimation de la profondeur d'un point n'est donc plus fondée sur un seul couple de données d'entrée mais sur $\varepsilon+1$ couples où ε désigne l'épaisseur du contour en nombre de pixels.

Néanmoins, le respect des hypothèses de calcul, sur l'orientation du contour et sur l'homogénéité de la discontinuité dans le sens de son gradient, réduit fortement le nombre de points d'estimation. Ces hypothèses, formant des contraintes fortes, doivent être levées notamment en généralisant la méthode à d'autres directions.

8 Références bibliographiques

- Benzougar A., Simon T., Khoudeir M., Bernard J. (1995). Monocular stereovision: depth map from two blurred images. *Proceedings of International Workshop on stereoscopic and three dimensional imaging*, 9/6-8, Santorini Greece, pp. 169-174.
- Benzougar A., Khoudeir M., Leard M. (1999). Formulation du flou pour la perception de la profondeur sur des scènes texturées. *Vision Interface '99*, Trois-Rivières, Canada, 19-21 mai, pp. 111-116.
- Darell T., Wohn K. (1990). Depth from focus using a pyramid architecture. *Pattern Recognition Letters*, 11(12), pp. 787-796.
- Ens J., Lawrence P. (1991). A matrix based method for determining depth from focus. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.
- Grossman P. (1987). Depth from focus. *Pattern Recognition Letters*, North Holland, pp. 63-68.
- Krotkov E. (1987). Focusing. *International Journal of Computer Vision*, 1, pp. 223-237.
- Monga O. (1990). Segmentation d'images : où en sommes-nous ? Rapport de recherche INRIA n° 1216.
- Pentland A.P. (1982). Depth of a scene from depth of field. *Proceedings of DARPA image understanding workshop*, Palo Alto, pp. 253-259.
- Pentland A.P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4), pp. 523-531.
- Pentland A.P., Darell T., Turk M., Huang W. (1989). A simple real time range camera. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 256-261.
- Pham D.T., Aslants V. (1999). Depth from defocus using a neural network. *Pattern Recognition*, 32, pp. 715-727.
- Schechner Y.Y., Kiryati N. (1998). Depth from Defocus vs. Stereo: How different really are they? Technical Report EE-PUB-1155, Technion-Israel Inst. Tech.
- Simon T. (1991). Appréciation de la profondeur depuis des images nette et floue en vision 3D passive. Thèse de Doctorat de l'Université de Nancy 1.
- Subbarao M. (1988-1). Parallel depth recovery by changing camera parameters. *Proceedings of IEEE Conference on Computer Vision*, Florida, USA, pp. 149-155.
- Subbarao M., Gurumothy N. (1988-2). Depth recovery from blurred edges. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, USA, pp. 498-503.
- Subbarao M. (1989). Determining distance from defocused images of simple objects. Technical Report No 89-07-20, Computer vision laboratory, Dpt of Electrical Engineering, State University of New-York, Stony Brook, NY 11794-2350, USA.
- Subbarao M., Choi T., Nikzad A. (1993). Focusing techniques. *Journal of Optical Engineering*, 32(11), pp. 2824-2836.
- Subbarao M., Surya G. (1994). Depth from Defocus: a spatial domain approach. *International Journal of Computer Vision*, 13(3), pp. 271-294.
- Williams C.S., Beckland A.A. (1989). *Introduction to the OTF*. Wiley Series in Pure and Applied Optics.