



**HAL**  
open science

## Approche Génétique pour la Perception de la Profondeurs dans les Images

Frédérique Bicking, Christophe Simon, Thierry Simon

► **To cite this version:**

Frédérique Bicking, Christophe Simon, Thierry Simon. Approche Génétique pour la Perception de la Profondeurs dans les Images. 5ème Conférence internationale sur l'automatisation industrielle, Jun 2007, Montréal, Canada. 4 p. hal-00162712

**HAL Id: hal-00162712**

**<https://hal.science/hal-00162712>**

Submitted on 16 Jul 2007

**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.

# APPROCHE GENETIQUE POUR LA PERCEPTION DE LA PROFONDEUR DANS LES IMAGES

Frédérique Bicking<sup>1</sup>, Christophe Simon<sup>1</sup>, Thierry Simon<sup>2</sup>

<sup>1</sup> CRAN UMR 7039 UHP-INPL-CNRS,

ESSTIN 2 Rue Jean Lamour, 54519 Vandœuvre, France.

E-mail : {Frederique.Bicking, Christophe.simon}@esstin.uhp-nancy.fr

<sup>2</sup> IUT de Figeac, Département GMP

Avenue de Nayrac 46100 Figeac, France. E-mail :Thierry.Simon@univ-tlse2.fr

**Résumé :** L'estimation de la profondeur par l'analyse du flou optique sur les contours d'images peut être vue comme un problème d'estimation de paramètres de modèles de contours et de défocalisation grâce à l'acquisition d'un couple d'images nette et floue. Nous proposons de résoudre ce problème par une méthode génétique appliquée simultanément à l'image nette et à l'image floue. La technique génétique est ensuite utilisée uniquement avec l'image floue en estimant l'ensemble des paramètres de contours et la quantité de flou optique en évitant l'acquisition de l'image nette. Les performances de cette approche sont évaluées avec deux puis une image et comparées à une de nos méthodes précédemment développées.

**Mots Clef :** Profondeur d'images, méthode génétique, vision monoculaire passive.

**Abstract :** This article deals with the application of a genetic method to depth perception in images in a passive monocular

vision system. Depth is estimated using the optical blur on the edges. Parameters of edge models and optical model are determined using the genetic method. The genetic approach is then applied with the blurred image only to estimate edge parameters and the amount of blur avoiding the sharp image acquisition. Performance of this genetic approach with two images, then with one, is evaluated and compared to one of our method previously developed.

**Keywords :** Depth perception, genetic method, passive monocular vision.

## 1 Introduction

De nombreuses méthodes ont été développées pour obtenir les coordonnées 3D des objets en utilisant des images. Les approches monoculaires exploitent le flou optique, introduit par la défocalisation, qui est principalement perceptible sur des zones hétérogènes de l'image comme les bords des objets ou les textures. Ces techniques, appelées Depth from Defocus ou DFD, emploient au moins deux images acquises avec des configurations différentes des paramètres de la caméra. Elles exploitent le contenu spatial de l'image au travers des caractéristiques géométriques ([6], [10]), de la forme des objets [2] dans la scène. Dans [8], nous avons défini une approche dans le domaine spatial basée sur des caractéristiques d'objets utilisant la réponse impulsionnelle du système optique d'acquisition. Cette méthode emploie une image nette en tout point et sa correspondante floue obtenue avec le même point de vue. Des opérateurs traditionnels de traitement d'images sont mis en œuvre pour estimer la quantité de flou optique présente localement et pour évaluer la profondeur des objets dans une scène. Les détails sont présentés dans la prochaine section.

Les méthodes génétiques ont trouvé beaucoup d'applications en visionique notamment dans les domaines de la détection de contours, de la segmentation, de la classification d'images ou de la reconnaissance de formes. Certaines de ces applications sont décrites dans les références ([5], [1], [9], [3]). Pour notre part, l'approche génétique est em-

ployée comme outil d'optimisation afin de calculer la profondeur d'objets à partir de leurs images floues et nettes, comme une méthode DFD classique, puis en utilisant uniquement l'image floue de la scène en supposant un modèle classique de contours.

Dans cet article, nous précisons rapidement l'aspect théorique de notre méthode de perception de profondeur en définissant d'une part, la relation entre la profondeur et la réponse impulsionnelle du système optique et d'autre part, les modèles mathématiques employés. Dans la seconde partie, la méthode génétique et ses applications à notre technique de perception de la profondeur sont décrites. Enfin, nous évaluons les performances de ces approches par l'analyse de la précision des mesures obtenues notamment dans un contexte bruité.

## 2 La méthode de perception de profondeur

Les méthodes de DFD permettent la détermination de la profondeur des objets d'une scène en utilisant au moins deux images acquises avec un seul système optique mais avec des conditions différentes d'acquisition pour chaque image. L'effet physique produit par la modification de la longueur focale ou l'ouverture du diaphragme sur les caractéristiques des images est exploité. Ainsi, la relation entre la profondeur, les paramètres de caméra et la quan-

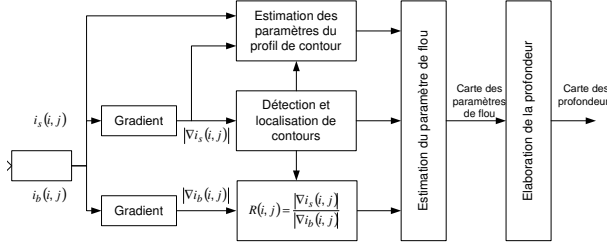


FIG. 1 – Synopsis de la méthode

tité de flou optique peut être établie.

Chaque point dans une scène est projeté en un point image conjugué sur le plan focal induisant la numérisation d'une image nette si le plan capteur est confondu avec ce plan focal. Si le plan capteur ne coïncide pas avec le plan focal, l'image d'un point devient une tache circulaire floue de rayon  $r_b$  en supposant que l'ouverture du diaphragme est circulaire. Les modèles employés pour représenter cette tache image sont caractéristiques de la fonction de transfert du système optique (OTF) liant  $r_b$  au paramètre d'étalement de la tache. Cette OTF est ainsi caractéristique d'une profondeur et permet de définir la relation entre un plan net et un plan flou. Dans le domaine spatial, cette relation s'écrit  $i_b(i,j) = i_s(i,j) \otimes_{2D} h(i,j)$  où  $\otimes_{2D}$  est l'opérateur de convolution bidimensionnelle,  $h(i,j)$  la réponse impulsionnelle du système optique,  $i_s(i,j)$  l'image nette et  $i_b(i,j)$  sa version floue. La connaissance de la réponse impulsionnelle du système optique appelée fonction d'étalement de point (PSF) permet d'obtenir les paramètres d'étalement pour des profondeurs différentes. Le lien entre le paramètre d'étalement  $Pe_{s_o}$  et la profondeur  $s_o$  est alors immédiat par la relation suivante :

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

où les constantes  $c$  et  $m$  sont caractéristiques d'un jeu de paramètres caméra. Ils sont déterminés par une procédure appropriée de calibrage.

Le principe général de la méthode est présenté sur la figure 1. L'acquisition d'une image nette avec un diaphragme en trou d'épingle et d'une image floue avec un diaphragme ouvert est réalisée. La position des contours des objets de l'image est détectée avec un opérateur de gradient. Une estimation du paramètre d'étalement de la PSF est obtenue à partir du rapport des modules des gradients des images nette et floue. A partir de l'estimation du paramètre d'étalement, on obtient la profondeur par la relation (1). Si ce paramètre est déterminé pour chaque pixel appartenant au profil de chaque contour des objets, nous formons une carte de contours épais de paramètres d'étalement et donc une carte de contours épais en profondeur.

Considérons un profil net de discontinuité d'ordre 1, dans la direction  $x$  selon la normale au contour  $c_s(x)$  de l'objet, exprimé en niveau de gris. Il est défini en (2) comme une

pende d'amplitude  $b - a$  et de longueur  $\varepsilon$  avec  $x$  gradué en pixel et représentant une ligne ou une colonne de l'image nette :

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

Le profil du contour flou est obtenu par la convolution monodimensionnelle du profil du contour net par la PSF selon la relation :

$$c_b(x) = c_s(x) \otimes h(x) \quad (3)$$

Cette relation exploite la fonction d'étalement de la ligne (LSF) définie par :

$$h(x) = \int_{-\infty}^{\infty} h(x,y) dy = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \quad (4)$$

En utilisant l'opérateur de Prewitt comme opérateur gradient, le rapport des modules des gradients des images devient :

$$R_\varepsilon(x) = \begin{cases} \frac{1}{h(x-x_0)+h(x-(x_0+\varepsilon))+2\sum_{u=1}^{\varepsilon-1} h(x-(x_0+u))} & \text{pour } x = x_0 \text{ et } x = x_0 + \varepsilon \\ \frac{2}{h(x-x_0)+h(x-(x_0+\varepsilon))+2\sum_{u=1}^{\varepsilon-1} h(x-(x_0+u))} & \text{pour } x \in ]x_0, x_0 + \varepsilon[ \end{cases} \quad (5)$$

où  $x_0$  est la position du premier point de contour selon l'axe  $x$  et  $\varepsilon$  la longueur de la discontinuité. L'indice  $\varepsilon$  signifie que  $R(x)$  dépend de la valeur de  $\varepsilon$  que l'on estime sur le profil du contour net.

Le paramètre d'étalement est estimé à partir de (5) pour chaque point  $x \in [x_0, x_0 + \varepsilon]$ , ce qui conduit à l'élaboration d'une carte de contours épais de profondeurs.

### 3 L'approche génétique

La méthode présentée est basée sur les algorithmes génétiques [4] et les stratégies d'évolution [7]. Elle combine le principe de survie de l'individu le plus fort et la combinaison structurée d'information utilisant des opérateurs génétiques afin d'élaborer un mécanisme de sélection élitiste et efficace. Ce mécanisme peut être esquissé globalement comme suit :

1. Initialiser aléatoirement une population dans le respect des contraintes énumérées
2. Evaluer la performance de chaque individu dans la population
3. Tester le critère d'arrêt. S'il est satisfait alors fin
4. Sinon sélectionner une proportion des meilleurs individus (parents pour la production de nouveaux individus)
5. Combiner le matériel génétique des parents choisis pour produire un nouvel individu

6. Tester le critère de coût pour ce nouvel individu. S'il est satisfait alors ce nouvel individu est accepté, sinon il est détruit et un autre individu est généré aléatoirement (mutation)
7. Répéter les étapes 4 à 6 jusqu'à la reconstruction de la population
8. Retourner à l'étape 3

Selon la définition du problème, une solution est représentée par un vecteur  $X$  appelé individu et doit satisfaire toutes les contraintes définies. Chaque individu représenté par un chromosome est une chaîne composée de  $m$  gènes où  $m$  est le nombre de paramètres à trouver. La valeur de chaque gène (appelé allèle) peut être entière ou réelle. Chaque individu est représenté par :  $I_p = [a_1, \dots, a_l, \dots, a_m]$  avec  $p = 1, \dots, N$  où  $N$  est la taille de la population.

Initialement, une population de  $N$  individus est créée aléatoirement. La justesse (ou fitness) de chaque individu est évaluée grâce à la fonction objectif. Contrairement aux algorithmes génétiques, la méthode employée ici est conçue pour minimiser et non pour maximiser. Ainsi, plus la valeur de la fonction objectif est faible, plus la justesse de l'individu est bonne.

Le processus de reproduction doit accentuer la survie des individus les plus aptes. Dans la méthode employée, une proportion d'individus possédant la meilleure aptitude (la moindre valeur de la fonction objectif) est sélectionnée pour la phase de reproduction faisant intervenir les opérateurs génétiques.

Après tirage aléatoire de deux parents parmi les individus sélectionnés, la recombinaison est exécutée : les gènes des deux parents  $P_1$  et  $P_2$  sont combinés pour générer un nouvel individu  $I_k$  selon le principe suivant :  $I_k = I_{P_1} + (I_{P_2} - I_{P_1}) \times \Delta$  où  $\Delta$  est un vecteur de nombres aléatoires. La combinaison de gènes peut être formalisée par :  $\{a_l\}_k = \{a_l\}_{P_1} + (\{a_l\}_{P_2} - \{a_l\}_{P_1}) \times \delta_1$  où  $\delta_1$  est une variable aléatoire uniforme avec  $k = N \times G + 1, \dots, N$  et  $l = 1, \dots, m$ .

Le nouvel individu créé  $I_k$  est alors évalué. Si sa justesse est meilleure que la pire des individus sélectionnés, il est intégré dans la population. Si ce n'est pas le cas, l'opérateur de mutation peut intervenir. La mutation consiste en la modification aléatoire d'un ou plusieurs gènes de ce nouvel individu. Ce schéma de reproduction est répété jusqu'à la reconstruction entière de la population. Cette nouvelle génération devient la population actuelle et le critère d'arrêt est évalué. S'il est satisfait, on considère que la population entière a convergé vers la solution optimale sinon la phase de reproduction est répétée. Le critère d'arrêt employé dans cette méthode exprime que tous les individus ont convergé vers la même solution et suppose qu'aucune évolution n'intervient, c'est-à-dire qu'aucun meilleur individu n'a été conçu. La solution finale représente la solution optimale selon les critères définis par la fonction objectif sous les contraintes imposées.

## 4 Evaluation expérimentale

Pour étudier les performances de la méthode génétique d'estimation de profondeur, nous considérons différentes images en environnement bruité. Les images bruitées sont générées en ajoutant un bruit blanc gaussien. Le rapport signal sur bruit (SNR) est imposé sur chaque image. On suppose que les images nette et floue sont bruitées avec le même SNR.

La comparaison des performances est basée sur une image nette de synthèse qui présente des discontinuités verticales. La méthode d'estimation est appliquée pour différentes valeurs de SNR, de longueur de la discontinuité  $\varepsilon$  et de paramètre d'étalement  $\sigma$ . Pour chaque test, 10 expériences sont réalisées et pour chacune d'elle, les images bruitées sont régénérées afin d'établir une statistique générale sur l'erreur d'estimation.

Afin de déterminer le paramètre d'étalement de la PSF, l'approche génétique est appliquée aux deux images dans un processus séquentiel. L'application de la méthode génétique sur l'image nette vise à trouver les quatre paramètres du profil de contour qui sont  $a, b, \varepsilon$  et  $x_0$  dans la relation (2). La justesse de l'individu est la valeur de la fonction objectif à minimiser définie par la somme quadratique des écarts entre le profil estimé et le profil réel du contour extrait de l'image nette :  $f(a, b, \varepsilon, x_0) = \sum_N (f_{est} - f)^2$ . La méthode génétique utilisant deux images nette et floue est évaluée sur des images de  $100 \times 100$  pixels en 256 niveaux de gris, dégradées avec des SNR de 30 à 50 dB.

Les résultats obtenus montrent une certaine efficacité de la méthode. Les performances sont relativement homogènes quelque soit le rapport signal sur bruit. On constate cependant sur l'erreur RMS pour un SNR donné, l'influence des valeurs du paramètre d'étalement  $\sigma$  et de la longueur du profil du contour  $\varepsilon$ . En effet, pour un  $SNR = 30dB$ , pour  $\sigma < 5$ , l'erreur RMS augmente avec  $\varepsilon$ , c'est-à-dire que plus le contour est épais, plus l'estimation est délicate. En revanche, pour  $\sigma \geq 5$  l'erreur RMS reste du même ordre de grandeur et ce quelque soit l'épaisseur du contour  $\varepsilon$ . Cette observation est également valable pour les autres valeurs de SNR.

Afin de quantifier l'intérêt de l'approche génétique pour la perception de profondeur, nous avons comparé ces performances à celles obtenues avec une méthode de DFD présentée dans [8].

La figure 2 illustre les erreurs RMS obtenues par la méthode génétique (notée AG), présentée sur la gauche, et celles obtenues par notre méthode de DFD (notée DFD), présentée sur la droite, pour des valeurs de  $\varepsilon = 1, 3, 5, 8, 10$  et des paramètres d'étalement  $\sigma = 1, 2, 3, 5, 8$ . Les résultats présentés montrent clairement l'intérêt de la méthode génétique notamment sur les contours épais  $\varepsilon \geq 5$  et pour des  $SNR \leq 40dB$ .

La méthode génétique utilisant uniquement l'image floue est évaluée sur des images de  $100 \times 100$  pixels en 256 niveaux de gris, dégradées avec des SNR de 30 à 50 dB.

Afin de déterminer le paramètre d'étalement de la PSF,

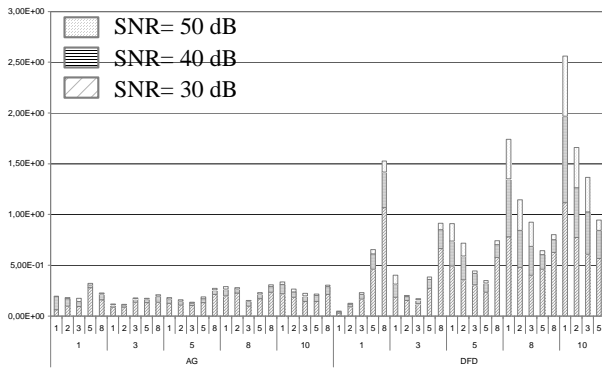


FIG. 2 – Comparaison des erreurs RMS entre la méthode génétique et la méthode DFD

l'approche génétique est appliquée en un processus unique. Selon le formalisme génétique, un individu est composé par un chromosome de 5 gènes. Les quatre premiers gènes représentent les paramètres du profil de contour net inconnu ( $a$ ,  $b$ ,  $\varepsilon$  et  $x_0$ ) et le cinquième le paramètre d'étalement  $\sigma$ . La justesse est la valeur de la fonction objectif définie par la somme quadratique des écarts entre un profil du contour flou estimé et le profil du contour flou extrait de l'image floue. Le contour flou estimé est calculé en générant un contour net grâce aux quatre premiers gènes selon le modèle de contour adopté (2). Le profil du contour flou estimé est calculé avec la relation (3).

Les résultats obtenus avec une image floue présentent des erreurs légèrement plus élevées que lors de l'utilisation des deux images mais les estimations sont globalement correctes en contexte bruité. l'influence de l'épaisseur du contour  $\varepsilon$  est ici moins perceptible tout comme celle du paramètre de flou  $\sigma$ . Les variations sur les valeurs des erreurs RMS semblent être dues majoritairement à l'influence du bruit.

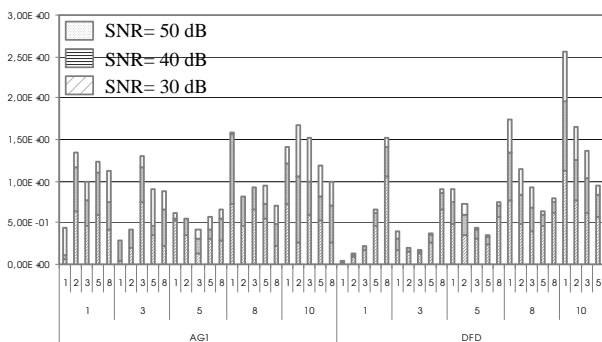


FIG. 3 – Comparaison des erreurs RMS entre la méthode génétique à une image floue et la méthode DFD

La figure 3 illustre les erreurs RMS obtenues par la méthode génétique n'utilisant que l'image floue (notée AG1), présentée sur la gauche, et celles obtenues par notre

méthode de DFD (notée DFD), présentée sur la droite. Les erreurs d'estimation sont globalement plus importantes en restant acceptables. Toutefois, l'approche génétique présente une moins grande sensibilité au bruit de par son procédé d'estimation.

## 5 Conclusion

Nous avons présenté une approche génétique pour la perception de profondeur dans les images. La première méthode génétique présentée utilisait, comme certaines méthodes de DFD classiques, une image nette et une image floue de la même scène. Les performances montrent un gain intéressant en terme d'erreur d'estimation par rapport à nos travaux précédents. La seconde méthode génétique présentée n'utilise que l'image floue. Ses performances sont moindres bien qu'acceptables en contexte bruité mais elle reste cependant une voie intéressante car seule l'image floue est utile. L'intérêt majeur d'appliquer une approche génétique à la détermination de profondeur est la recherche simultanée des paramètres de caractéristiques issues d'images nette et floue.

## Références

- [1] P. Andrey. Selectionist relaxation : genetic algorithms applied to images segmentation. *Image and Vision Computing*, 14(3-4) :175–187, 1999.
- [2] T. Darell and K. Wohn. Depth from focus using a pyramid architecture. *Pattern Recognition Letters*, 11 (12) :787–796, 1990.
- [3] B. C. G. Garai. A cascaded genetic algorithm for efficient optimization and pattern matching. *Image and Vision Computing*, 20(4) :265–277, 2002.
- [4] D. Goldberg. *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, New York, 1989.
- [5] J. K. M. Mirmehdi, P.L. Palmer. Genetic optimization of the image feature extraction process. *Pattern Recognition Letters*, 18 :355–365, 1997.
- [6] A. Pentland. A new sense of depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4) :523–531, July 1987.
- [7] H. Schewefel. *Numerical optimization of computer models*. Wiley edition, 1981.
- [8] C. Simon, F. Bicking, and T. Simon. Estimation of depth on thick edges from sharp and blurred images. In *IEEE Instrumentation and Measurement Technology Conference*, pages 323–328, Anchorage, USA, May 21-23 2002. IEEE Instrumentation and Measurement Society.
- [9] H. Z. S.M. Bhandarkar. Image segmentation using evolutionary computation. *IEEE Transactions on Evolutionary Computing*, 3(1) :1–21, 1999.
- [10] D. Ziou and F. Deschenes. Depth from defocus in spatial domain. *Computer Vision and Image Understanding*, 81 :143–165, 2001.