

Mesure de netteté basée sur les descripteurs généralisés de Fourier appliquée à la reconstruction 3D par Shape from Focus

BASTIEN BILLIOT¹, FREDERIC COINTAULT², PIERRE GOUTON¹

¹ Laboratoire Electronique, Informatique et Image

9 avenue Alain Savary, BP 47870, 21078 Dijon Cedex, France

² AgroSup Dijon

26 boulevard Docteur Petitjean, BP 87999, 21079 Dijon Cedex, France

¹bastien.billiot@u-bourgogne.fr

Résumé – L'étape principale de la méthode de reconstruction 3D « Shape from Focus » est l'utilisation d'un opérateur de mesure de netteté de chaque pixel de la séquence d'image. Le choix de l'opérateur de mesure de netteté est une étape cruciale pour une reconstruction 3D de qualité. La précision de la mesure de netteté dépend de la taille du voisinage autour du pixel choisi et de la présence ou non de bruit additif dans la séquence d'images. Dans cet article, nous présentons deux nouveaux opérateurs de mesure de netteté basés sur les Descripteurs Généralisés de Fourier. Une nouvelle étude comparative des différents opérateurs est présentée. Cette comparaison est basée sur un plan d'expériences mettant en évidence les effets de la variation du voisinage utilisé et le bruit additif sur la précision de la reconstruction. Les résultats font office de recommandation pour le choix de l'opérateur à utiliser en fonction du voisinage choisi et du bruit présent. Les résultats de nos opérateurs montrent qu'ils sont à recommander lorsqu'on est en présence d'une scène bruitée et que l'on travaille sur un petit voisinage.

Abstract – The principal step of the “Shape from Focus”, a 3D reconstruction method, is the use of a focus measure operator to compute the sharpness for each pixel of the acquired image sequence. The choice of focus measure operator is a crucial step for the quality of 3D reconstruction. The accuracy of the focus measure depends on the local size of the neighborhood of each pixel used for the focus measure computation and the presence or not of noise in the image sequence. In this paper, we present two new focus measure operators based on Generalized Fourier Descriptors. A new comparative study of the different operators is presented. This comparison is based on a factorial experiment identifying the effects of Region Of Interest (ROI) size variation and additive noise on the reconstruction accuracy. The results provide recommendation of which operator to use based on defined case (i.e. different ROI size and noise). The results of our operators have shown that they are recommended for noisy scenes and small ROI.

1 Introduction

L'obtention de l'information tridimensionnelle d'une scène est une problématique récurrente en vision par ordinateur. L'état de l'art dans ce domaine est vaste mais nous pouvons classer ces méthodes de reconstruction 3D en trois catégories : les approches géométriques, photométriques et celles basées sur les propriétés physiques du système d'acquisition utilisé. La méthode communément appelée « Shape from Focus » [1] fait partie de cette troisième catégorie et est une méthode de reconstruction 3D passive et monoculaire. Son principe général est divisé en quatre étapes. La première est le procédé d'acquisition particulier qui consiste à acquérir une séquence d'images de la même scène en faisant varier la position du plan focal afin de balayer cette scène. Le point déterminant de ce procédé d'acquisition est d'obtenir une profondeur de champ relativement faible. La notion de profondeur de champ peut être vue comme la zone de la scène visualisée avant et après laquelle on estime que l'image est floue. Il en résulte une zone d'une certaine profondeur où les pixels sont nets. L'épaisseur de cette zone conditionnera le pas de variation du plan focal entre deux acquisitions.

Une séquence d'images couvrant toute la profondeur de la scène désirée est obtenue. La deuxième étape est la mesure de la netteté de chaque pixel pour toutes les images. Un même pixel de la séquence représentant le même point spatial dans la scène, une courbe de valeur de netteté est obtenue pour chacun des points de la scène. Le maximum de cette courbe renseigne donc directement sur la position du pixel le plus net dans la séquence d'images. La troisième étape est souvent une approximation de cette courbe afin d'améliorer la précision de la mesure. La dernière étape est la création de la carte de profondeur de la scène où un niveau de gris est associé à une position d'image dans la séquence qui est elle-même associée à une valeur de distance par rapport au système d'acquisition. Ainsi, on obtient une représentation de l'information tridimensionnelle sous forme de niveaux de gris où, par exemple, un pixel sombre représentera un point éloigné alors qu'un clair représentera un point proche.

La qualité de la carte de profondeur est donc fortement dépendante de l'efficacité de la mesure de netteté utilisée. On retrouve dans la littérature beaucoup de mesures de netteté différentes réparties en plusieurs

catégories : les mesures différentielles, statistiques, de corrélation ou encore fréquentielles.

Dans ce papier, nous proposons deux nouveaux opérateurs de netteté dont l'efficacité sera comparée à celle d'opérateurs existants au moyen d'une méthode adaptée basée sur un plan d'expérience factoriel.

2 Opérateur de mesure de netteté

Les mesures différentielles regroupent l'ensemble des opérateurs se basant sur des filtres gradients ou Laplaciens. En effet, les contours seront plus marqués dans une image nette que dans une image floue. Ainsi, [2] propose d'évaluer la netteté d'un pixel par la somme de l'amplitude du gradient des pixels voisins (FM_{ten}). [3] proposent une variante en utilisant la variance des amplitudes de chaque pixel du voisinage comme mesure de netteté (FM_{tenvar}). L'utilisation d'un filtre Laplacien modifié est présenté par [1] comme mesure de netteté (FM_{smi}).

Les mesures de netteté statistiques se basent sur le fait que l'énergie lumineuse provenant d'un point de la scène est distribuée sur un ensemble de pixels adjacents lorsque l'image est floue. Ainsi, la mesure de la variance locale des niveaux de gris autour d'un pixel a été proposée par [4] (FM_{var}). L'analyse de l'histogramme (FM_{range}) ou de l'entropie (FM_{ent}) de chaque voisinage permet également une évaluation de la netteté locale.

[5] utilise l'autocorrélation de chaque voisinage comme mesure de netteté. En effet, pour un voisinage flou, l'autocorrélation sera plus élevée du fait des niveaux de gris moyennés par l'étalement de l'énergie lumineuse sur le capteur (FM_{corr}).

Enfin, partant du principe qu'une zone nette comporte plus de composantes hautes fréquences qu'une zone floue, on retrouve beaucoup de mesures de netteté basées sur une analyse fréquentielle du voisinage. C'est le cas de [6] qui suggèrent d'utiliser le ratio entre le moment du second ordre des hautes fréquences d'une transformée en ondelette et le coefficient de base fréquence ($FM_{waveratio}$). Cela s'explique par le fait que lorsqu'une image devient floue, l'énergie des hautes fréquences décroît simultanément à l'augmentation de l'énergie des bases fréquences.

La méthode de reconstruction 3D étant principalement fonctionnelle sur des zones texturées, nous avons définis deux mesures de netteté basée sur les Descripteurs Généralisés de Fourier [7] habituellement utilisés en analyse de texture. Ils sont définis par :

$$D_f(\lambda) = \int_0^{2\pi} |\hat{f}(\lambda, \theta)|^2 d\theta \quad (1)$$

$$\hat{f} = \int_{\mathbb{R}^2} f(x) e^{-j\xi x} dx \quad (2)$$

Avec \hat{f} la transformée de Fourier de f et (λ, θ) les coordonnées polaires du point ξ .

Le vecteur D_f est calculé pour le voisinage de chaque pixel de l'ensemble des images de la séquence. On obtient donc un nombre de vecteurs pour chaque point de la scène égal au nombre d'images acquises. Ainsi, deux manières d'utiliser ces informations pour créer des mesures de netteté sont proposées. La première est d'additionner les valeurs de chaque vecteur car, travaillant sur des zones texturées, cette somme sera plus élevée pour un pixel net que pour un flou. On exprimera la mesure de netteté comme suit :

$$FM_{gfdsun} = \sum_{\lambda=1}^N D_f(\lambda) \quad (3)$$

Avec N la taille du voisinage.

Pour la seconde mesure de netteté, on utilisera une analyse en composante principale sur la matrice composée des différents vecteurs concaténés (FM_{gfdacp}). La position du score factoriel maximum du premier axe informera sur la position du pixel le plus net.

$$FM_{gfdacp} = position(\max(\text{Scores factoriels})) \quad (4)$$

3 Etude comparative

Afin de choisir l'opérateur de mesure de netteté à utiliser en fonction de la séquence d'images acquise, il est nécessaire d'effectuer une comparaison de ces opérateurs. Deux paramètres impactent l'efficacité d'une mesure de netteté. Le premier est la taille du voisinage. En effet, une taille de voisinage large induira un nombre d'échantillons plus important pour la mesure et donc une meilleure précision. Cependant, le fait de choisir une zone trop large entraîne beaucoup d'erreurs et un effet de lissage de la carte de profondeur dans le cas d'une scène présentant beaucoup de discontinuités. A contrario, une zone très petite induira une carte de profondeur plus précise en cas de mesure correcte mais plus sensible aux erreurs liées au nombre d'échantillons plus faible. Le deuxième paramètre influant sur l'efficacité de la mesure de netteté est la présence de bruit dans les images. La méthode classique de comparaison de l'efficacité des opérateurs de netteté est le calcul de l'erreur quadratique entre la carte de profondeur obtenue avec un opérateur et celle correspondant à la vérité terrain [8]. Cependant, nous souhaitons mesurer la robustesse des opérateurs au bruit additif ainsi qu'au changement de taille de voisinage. Pour cela, nous avons mis au point une méthode de comparaison adaptée à notre problématique qui est la reconstruction 3D de scène texturée. L'image de référence sur laquelle se base cette comparaison est une mosaïque constituée de trente images issues de la base texturale de Brodatz [9]. Afin d'obtenir une séquence d'images semblable à celles utilisées en « Shape from

Focus », une seconde bande correspond à l'image mosaïque à laquelle un flou de défocalisation est appliqué. Une troisième bande sera créée par un flou deux fois plus important. Ainsi on obtient une carte de profondeur considérant la première bande comme la plus proche du système d'acquisition pour une mesure de netteté sans erreur. Dans notre cas, cela correspond à une image entièrement noire.

Afin d'évaluer la robustesse de chaque opérateur au bruit et à la taille du voisinage, un plan d'expériences factoriel est utilisé. Un plan d'expériences permet de mettre en évidence le comportement d'une réponse en fonction de la variation des paramètres choisis. Dans notre cas, on utilise un plan factoriel à deux niveaux car les paramètres d'entrée sont le bruit additif ainsi que le changement de la taille du voisinage considéré pour la mesure de netteté. La variation du bruit additif sera faite sur les trois bandes de l'image de référence et sera de type gaussien. Ce type de bruit simulant un bruit thermique affectant le capteur. Ainsi, on aura deux types d'image de référence : une sans bruit et une avec bruit. La variation de la taille du voisinage sera effectuée lors de l'application des différentes mesures de netteté. Nous obtenons ainsi quatre cartes de profondeur pour chaque opérateur à analyser. L'erreur quadratique est ensuite calculée entre chaque carte de profondeur obtenue et celle correspondant à la vérité terrain. Le plan d'expérience est construit à partir de ces informations et permet de calculer les effets de la variation des différents paramètres sur l'erreur quadratique ainsi que l'interaction existante induite par la variation simultanée des deux paramètres. L'interaction renseigne sur la sensibilité d'un opérateur à une variation simultanée de deux paramètres. Une forte interaction signifie que l'erreur induite est démultipliée du fait de la variation simultanée des deux paramètres influant sur la mesure de netteté. Une faible interaction, quant à elle, signifie que la variation de paramètres simultanée n'entraîne pas plus d'erreur que si un seul paramètre varie.

4 Résultats

On retrouve dans le tableau 1 les différents effets et interactions pour plusieurs opérateurs de mesure de netteté. La variation de la valeur d'erreur quadratique en fonction des différentes valeurs de paramètres utilisés est présentée figure 1.

L'interprétation des résultats doit être réalisée non seulement d'après les différents effets et interactions pour chaque opérateur mais aussi conjointement à l'analyse de l'évolution de l'erreur.

Tab 1 : Effets et interactions

Opérateurs	Effet taille	Effet bruit	Interaction
FM_{ten}	0.655	0.52	0.224
FM_{tenvar}	0.719	0.342	0.342
FM_{sml}	0.126	0.85	-0.106
FM_{var}	0.721	0.339	0.339
FM_{range}	0.734	0.324	0.324
FM_{ent}	0.251	0.716	-0.417
FM_{corr}	0.832	0.223	-0.087
$FM_{waveratio}$	0.225	0.834	-0.066
FM_{gfdacp}	0.408	0.764	-0.01
FM_{gfdsum}	0.737	0.322	0.322

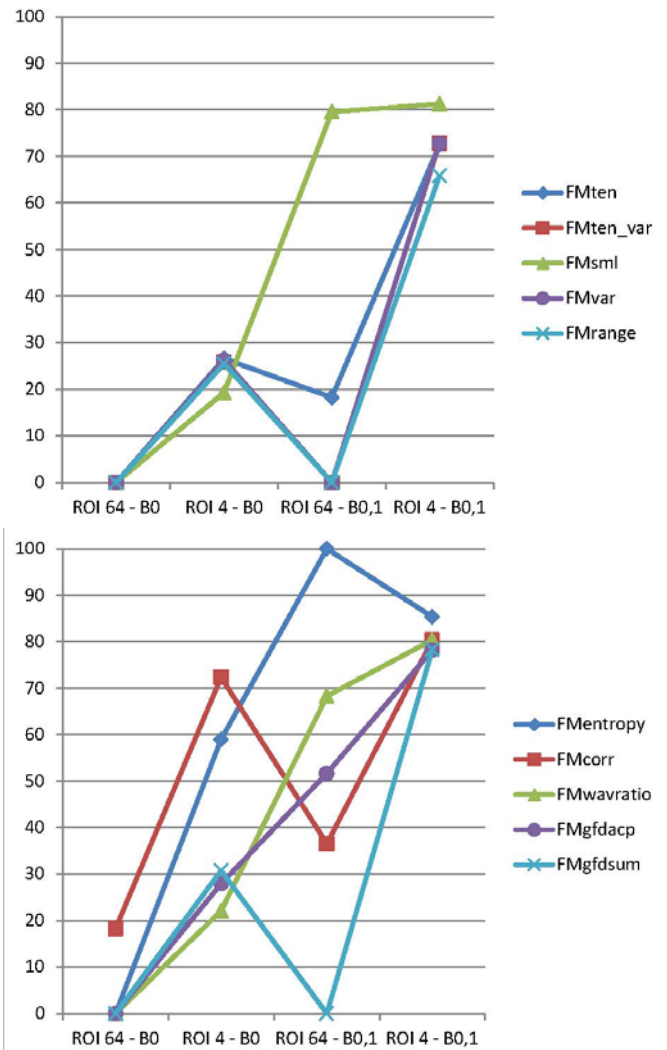


Figure 1 : Erreurs quadratiques en fonction des valeurs des paramètres

Ainsi, on constate que les opérateurs les moins affectés par la diminution de la taille du voisinage sont $FM_{waveratio}$ et FM_{sml} . Cependant, ces opérateurs sont fortement impactés par un bruit additif. Inversement, les opérateurs FM_{tenvar} , FM_{var} , FM_{range} et FM_{gfdsum} sont moins sensibles au bruit additif mais fortement à la diminution du voisinage. Ces opérateurs souffrent d'une interaction relativement forte impliquant que ces opérateurs sont à déconseiller dans le cas de présence de bruit et la nécessité d'un petit voisinage.

FM_{ent} fait état d'une forte sensibilité au bruit et d'une interaction élevée négative. Ceci a pour effet de diminuer l'erreur globale lors de la variation des deux paramètres impactant l'efficacité des mesures.

Enfin, on remarque que l'opérateur FM_{gfdacp} voit son erreur globale augmenter linéairement en fonction de la variation des paramètres choisis. Ceci en fait un opérateur adapté à la plupart des cas.

On peut synthétiser les résultats en quatre catégories. La première est l'utilisation d'un grand voisinage autour de chaque pixel et aucun bruit additif. Dans cette configuration optimale pour une mesure de netteté correcte, seul l'opérateur FM_{corr} entraîne quelques erreurs. L'ajout de bruit à cette précédente configuration met en évidence quatre opérateurs fonctionnant efficacement en présence de bruit dont l'un de nos opérateurs (FM_{gfdsum}). Une petite taille de voisinage et aucun bruit additionnel induit des erreurs pour tous les opérateurs sauf pour deux qui sont fortement impactés par ce petit voisinage. Enfin, dans le cas extrême où le voisinage est de petite taille et du bruit est présent dans la séquence d'images, l'efficacité de tous les opérateurs est fortement impactée. Le paramètre à prendre en compte dans ce cas lors du choix de l'opérateur à utiliser est la plus faible interaction possible entre la diminution de la taille du voisinage et la présence de bruit. Dans notre cas, l'opérateur présentant la plus faible interaction est l'opérateur que nous avons proposé, FM_{gfdacp} .

5 Conclusion

Dans ce papier, deux nouveaux opérateurs permettant la mesure de la netteté sont présentés ainsi qu'une méthode de comparaison adaptée à la problématique de reconstruction 3D par « Shape from Focus ». En effet, cette comparaison est basée sur des images de référence texturées et permet de mettre en avant le comportement des différentes mesures par rapport à la taille du voisinage considéré ainsi que la présence ou non de bruit. L'objectif de ce type de comparaison est de déterminer les forces et faiblesses des différents opérateurs de mesure de netteté et pas seulement de les classer. Ceci permet à l'utilisateur connaissant les caractéristiques de sa scène de choisir l'opérateur le plus adapté dans son cas. Les deux opérateurs présentés sont à recommander au regard de leurs résultats comparés aux opérateurs traditionnels.

6 Remerciements

Nous remercions le Conseil Régional de Bourgogne et le pôle de développement Nicéphore Cité pour leurs soutiens financiers.

7 Références

- [1] S. K. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, pp. 824–831, August 1994.
- [2] J. Tenenbaum, *Accommodation in computer vision*. PhD thesis, 1970.
- [3] J. Pech-Pacheco, G. Cristobal, J. Chamorro-Martinez, and J. Fernandez-Valdivia, "Diatom autofocusing in brightfield microscopy: a comparative study," in *icpr*, p. 3318, Published by the IEEE Computer Society, 2000.
- [4] F. Groen, I. Young, and G. Ligthart, "A comparison of different focus functions for use in autofocus algorithms," *Cytometry*, vol. 6, no. 2, pp. 81–91, 1985.
- [5] D. Vollath, "Automatic focusing by correlative methods," *J. Microsc.*, vol. 147, no. 3, pp. 279–288, 1987.
- [6] H. Xie, W. Rong, and L. Sun, "Wavelet-based focus measure and 3-d surface reconstruction method for microscopy images," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pp. 229–234, IEEE, 2006.
- [7] F. Smach, C. Lemaître, J. Gauthier, J. Miteran, and M. Atri, "Generalized fourier descriptors with applications to objects recognition in svm context," *Journal of Mathematical Imaging and Vision*, vol. 30, no. 1, pp. 43–71, 2008.
- [8] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion," *Pattern Recognition Letters*, vol. 28, no. 4, pp. 493–500, 2007.
- [9] P. Brodatz, *Textures: a photographic album for artists and designers*, vol. 66. Dover New York, 1966.