HAL open science

Détermination automatique du seuil de binarisation des modules des gradients par modélisation de leur histogramme

Bilal Karabagli, Thierry Simon, Jean-José Orteu

To cite this version:

Bilal Karabagli, Thierry Simon, Jean-José Orteu. Détermination automatique du seuil de binarisation des modules des gradients par modélisation de leur histogramme. ORASIS - Congrès des jeunes chercheurs en vision par ordinateur, INRIA Grenoble Rhône-Alpes, Jun 2011, Praz-sur-Arly, France. inria-00595260

HAL Id: inria-00595260

https://inria.hal.science/inria-00595260v1

Submitted on 24 May 2011

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.

Détermination automatique du seuil de binarisation des modules des gradients par modélisation de leur histogramme

Bilal KARABAGLI [1, 2], Thierry SIMON[1], Jean-José ORTEU [2]

[1] IUT de Figeac : UTM-LRPmip, Avenue de Nayrac, 46100 FIGEAC, France
[2] Université de Toulouse ; Mines Albi ; ICA (Institut Clément Ader) ; Campus Jarlard, F-81013 Albi, France

bkarabag@mines-albi.fr, thierry.simon@univ-tlse2.fr, jean-jose.orteu@mines-albi.fr

Résumé
Nous présentons dans cette communication une amélioration d'une méthode présentée dans [6]. Notre objectif est d'obtenir les contours dans une image à partir de la binarisation des modules des gradients. Ces contours sont obtenus à partir d'un seuil calculé automatiquement sur l'histogramme du module des gradients. La méthode doit être simple pour être rapide, implantable sur un circuit FPGA, et doit exploiter les gradients calculés par ailleurs. La méthode sera intégrée dans une tâche de perception 3D pour la vérification de montage d'usinage en vue de sécuriser l'opération d'usinage grande vitesse.

Mots Clef
Segmentation, contour, gradient, seuillage automatique.

Abstract
This paper presents an enhanced method of image thresholding presented in [6]. The goal is to obtain the edges of objects in an image using a gradient magnitude histogram to automatically determine the threshold value. The method must be simple to be fast and easy to implement on an FPGA circuit. The computed edges will be used within a 3D perception task for high-speed machining security.

Keywords
Segmentation, edge, gradient, automatic thresholding.

1 Introduction
Ce travail répond à notre besoin de disposer d'un détecteur de contours dans des images de montage d'usinage, dont un exemple est présenté Figure 1, pour vérifier la validité du montage en regard de la trajectoire de l'outil de coupe en fraisage. Ce travail s'inscrit comme un élément d'une tâche de reconstruction tridimensionnelle d'une scène de montage d'usinage pour la sécurisation de machine outil. En effet, il n'est pas rare que dans les phases d'optimisation des trajectoires de coupe, phases importantes de mise au point assurant un gain de productivité, que l'opérateur soit amené à déplacer les éléments de bridage du brut à usiner (cf. Figure 1). Ce déplacement n'est pas toujours transféré dans le système de FAO pour le calcul de la nouvelle trajectoire à cause du coût prohibitif de ce transfert en terme de temps. Ce transfert est remplacé par un essai de validation pas toujours maîtrisé. Cet essai devient impossible pour les processus d'usinage grande vitesse auxquels nous nous intéressons. En dehors de cette phase de mise au point, il est courant dans le milieu de production des pièces aéronautiques ou automobiles qu'une machine soit utilisée pour l'usinage de différentes pièces, parfois dans une même journée, demandant à l'opérateur d'assurer des montages multiples. La confusion dans le montage peut ainsi provoquer des incidents graves.

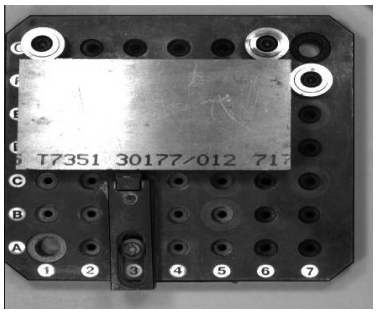


Figure 1 : Image de montage d'usinage : le brut (pièce en aluminium) est fixé par un système de bridage.

L'objectif de nos travaux, inscrits dans le cadre d'une thèse cofinancée par la région Midi-Pyrénées, le Centre Universitaire de Formation et de Recherche Jean-François Champollion d'Albi et l'IUT de Toulouse II Figeac, est de développer un système de vision capable de vérifier que le montage d'usinage est conforme, avant de lancer l'opération d'usinage, et ainsi éviter des incidents (rupture de broche due à des collisions par exemple)

Nous avons choisi de réaliser la perception tridimensionnelle à partir des contours. La scène est composée d'objets polyédriques, cylindriques ou sphériques comme le montre la Figure 2 présentant le montage d'usinage saisi en CAO.

La réponse de validité du montage doit être proposée à l'opérateur dans un temps court compatible avec les

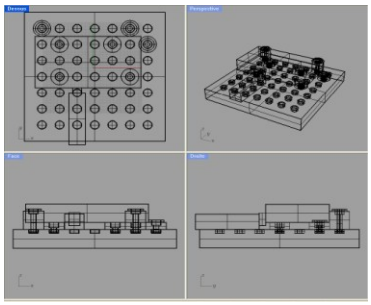


Figure 2 : Modèle CAO de montage d'usinage

contraintes de production. Cette approche contour comme élément de segmentation reste d'actualité comme le montrent les derniers articles de la communauté [1] [2] et le travail du « computer vision group » de l'université Berkeley [3,19].

Nous avons choisi dans un premier temps une approche par binarisation des modules des gradients. Le calcul des gradients par l'opérateur de Prewitt peut être confié à un circuit FPGA de composition simple permettant d'obtenir un résultat à la cadence vidéo. De plus les gradients seront utilisés par la première méthode de perception 3D que nous envisageons de tester : une méthode basée sur l'exploitation du flou optique [4] [5].

La méthode que nous présentons ici est une amélioration d'une méthode que nous avons développée et présentée dans [6]. Cette méthode peut être classée dans le premier groupe d'après la structuration des méthodes proposée par G. Papari et N. Petkov dans leur état de l'art [1]. Les trois groupes constitués sont caractérisés par l'information image utilisée :

1- Méthodes basées sur des caractéristiques locale ; on trouve dans ce groupe les méthodes dérivatives suivi d'un seuillage plus ou moins sophistiqué.

2- Méthodes globales ou contextuelles parmi lesquelles on trouve les contours actifs [16].

3-Méthodes multi-résolutions, avec par exemple les filtres de Gabor [15].

Dans le premier groupe nous pouvons identifier plusieurs sous groupes définissant plusieurs type de méthodes.

1- Méthodes utilisant explicitement la modélisation de l'histogramme [10,11], dont notre méthode fait partie.

2- Méthodes utilisant implicitement la forme de l'histogramme : algorithme proposé par Yi Wan, Jiangchang Wang, Xingbo Sun et Ming Hao [18] qui présente une amélioration de la méthode d'OTSU en tenant compte de la forme asymétrique de l'histogramme des modules des gradients, ou sur l'entropie [12], algorithmes proposés par KAPUR, YEN, etc.

3- Méthodes qui utilisent une information locale, passage par zéro du Laplacien, dérivée du filtre de CANNY ou DOG [13, 14].

La section 2 présente succinctement notre méthode initiale et la section 3 l'amélioration proposée. Les résultats de la section 4 comparent les performances de notre algorithme par rapport à d'autres méthodes dans l'esprit des « benchmarks » de l'université Berkeley, pour la méthode initiale ainsi que pour les méthodes de seuillage automatique très utilisées comme celle proposée par [18], et une méthode exploitant l'entropie [12]. Enfin, la discussion des résultats est abordée dans la section 5, suivi d'une conclusion dans la section 6.

2 Méthode initiale

La méthode initiale part du constat que la forme de l'histogramme du module des gradients est constante et peut être modélisée par l'expression (1), somme de deux fonctions de Rayleigh où $x \in [0,255]$ le niveau de gris, α, β, λ, n, m des paramètres à identifier :

$$H(x) = \frac{(1-\lambda)}{\Gamma(n+1)} x^n e^{-\alpha x} \alpha^{n+1} + \frac{\lambda}{\Gamma(m+1)} x^m e^{-\beta x} \beta^{m+1} \quad (1)$$

La première partie représente les gradients peu intenses de zones homogènes bruitées ou de petites zones texturées, la seconde partie représente les fortes discontinuités de luminance marquant les contours et le bruit « poivre et sel ».

Si la valeur de λ est petite, cela conduit à une forme unimodale très marquée. Le modèle est réduit alors à l'expression (2).

$$H(x) = \frac{1}{\Gamma(n+1)} x^n e^{-\alpha x} \alpha^{n+1} \quad (2)$$

Placer un seuil de binarisation lié à la largeur du mode a été la solution envisagée dans [6]. L'intersection de la tangente au point d'inflexion de plus grande abscisse avec l'axe des abscisses donne une bonne évaluation de cette largeur.

Cette intersection peut être calculée à partir des paramètres du modèle par l'expression (3).

$$W = \frac{\left(n + \sqrt{n}\right)^2}{n\alpha} \quad (3)$$

Les paramètres sont calculés à partir des moments d'ordre 0, 1, 2 par :

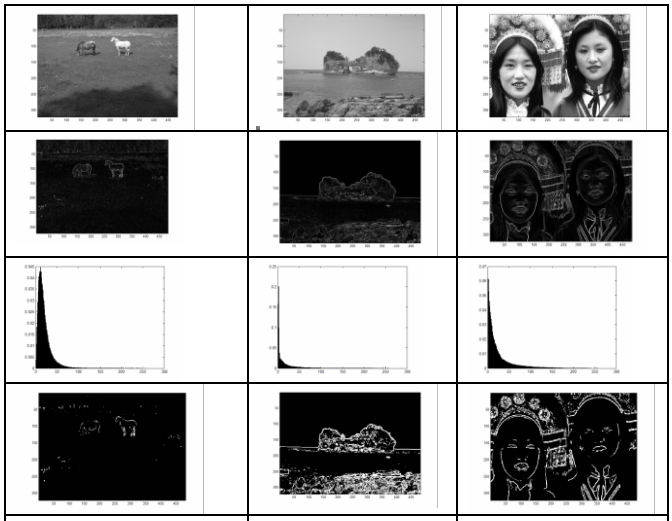
$$n = \frac{2m_1^2 - m_0 m_2}{m_0 m_2 - m_1^2}, \ \alpha = \frac{m_0}{m_1}(n+1)$$

L'histogramme est normalisé, $m_0 = 1$, le calcul des paramètres se réduit aux équations suivantes :

$$n = \frac{2m_1^2 - m_2}{m_2 - m_1^2}, \ \alpha = \frac{1}{m_1}(n+1)$$

Dans le cas où n est nul, pour des images comportant de grandes zones saturées au noir ou au blanc, le modèle devient une exponentielle décroissante :

$H(x) = e^{-\alpha x}\alpha$.



S=78	S=33	S=62
n=0,8882, α=0,1352	n=0, α=0.0840	n=0, α=0,0451

Figure 3 : Images de la base [8], résultat du seuillage par la méthode initiale.

Le seuil calculé est obtenu par l'intersection de la tangente à l'origine avec l'axe des niveaux de gris, sa valeur est :

$$W = \frac{1}{\alpha}$$

Le seuil W ainsi appliqué donne une image des contours très bruitée. Une comparaison avec des seuils proposés par des experts a conduit à l'augmentation d'une proportion moyenne de 2,8 pour un intervalle de variation comprise entre 2 et 3.

La méthode donne de bons résultats mais ne tient pas compte de l'étalement du deuxième mode. La Figure 3 (ci-dessus) présente des exemples de résultats sur trois images de la base « Berkeley Segmentation Dataset and Benchmarks » [8]

3 Amélioration proposée

L'amélioration proposée consiste à tenir compte du deuxième mode et d'estimer sa largeur. Si on applique la méthode de modélisation initiale sur l'histogramme diminué de la première modélisation, le résultat n'est pas satisfaisant. La forme du mode étant plus bruitée, les paramètres estimés ne donnent pas un bon modèle. Nous choisissons de prendre un développement limité de la forme initiale, soit l'expression (4) pour privilégier une forme parabolique :

$$H(x) = \frac{(1-\lambda)}{\Gamma(n+1)} x^n e^{-\alpha x} \alpha^{n+1} + \lambda.(ax^2 + bx + c) \qquad (4)$$

Nous considérons l'histogramme comme une densité de probabilité (probabilité d'apparition d'un niveau de gris par rapport à l'ensemble des pixels) ce qui implique que les paramètres *a, b, c* et λ seront calculés pour respecter :

$$\int_0^{255} H(x)dx = 1 .$$

Le triplet (*a, b, c*) contient les coefficients du polynôme de la forme parabolique. Ils sont déterminés par approximation aux moindres carrés régressive (MCR) [7]. La Figure 4 présente l'histogramme des modules des gradients d'une image et le modèle associé.

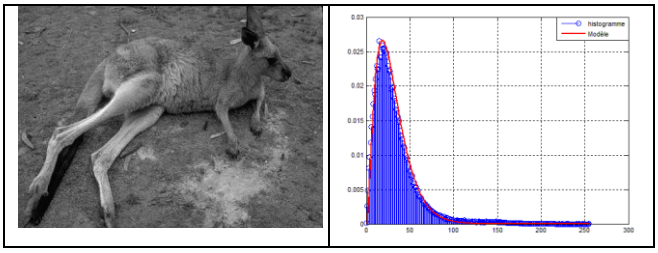


Figure 4 : Modélisation de l'histogramme. Image de la Base BSDS300 [8].

La détermination des coefficients du modèle parabolique est réalisée sur un intervalle de niveau de gris $x \in [\varepsilon_1, \varepsilon_2]$.

La valeur ε_1 est calculée comme la première valeur de la densité réduite (dans le sens des niveaux de gris croissants), valeur absolue de la différence entre l'histogramme expérimental complet et son modèle unimodal, supérieur à un seuil significatif, seuil fixé heuristiquement à 10^{-4}. La borne supérieure ε_2 est obtenue par la même technique mais dans le sens des niveaux de gris décroissants.

Le seuil est calculé en prenant l'abscisse du point d'intersection des deux tangentes *D1* et *D2* (Figure 5a), respectivement au point d'inflexion du modèle de Rayleigh, et au point d'abscisse la plus petite du modèle polynomial (Figure 5.b).

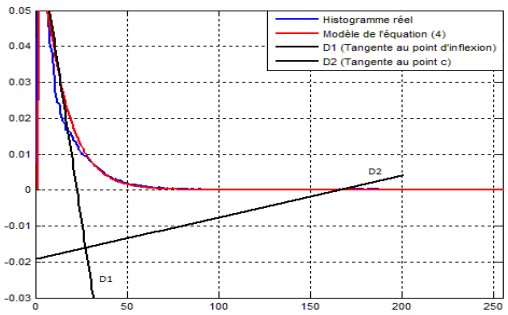


Figure 5.a : Seuil de binarisation : Point de croisement des tangentes (*D1* et *D2*).

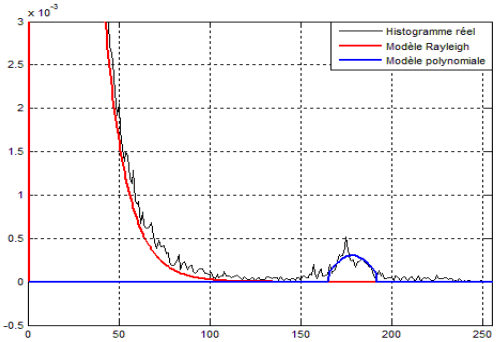


Figure 5.b : Histogramme du module des gradients : Zoom avant sur le deuxième mode.

Cette technique de binarisation donne des meilleurs résultats, comme nous le montrons dans la section suivante, par rapport aux méthodes existantes de même catégorie.

4 Résultats et comparaisons

Afin de comparer notre méthode nous proposons de reprendre les indicateurs de performance proposés par le groupe vision de l'Université de Berkeley. La binarisation est considérée comme une classification des pixels en deux classes. Celle des « contours » (classe A) et celle des « non contours » (classe $\overline{A}$), cela suppose qu'il existe une référence pour contrôler la classification. Cette référence a été produite à partir de proposition d'experts. Les Figures 6.a et 6.b présentent des exemples extraits de la base.

Figure 6.a : Extrait 1 de la base.

Figure 6.b : Extrait 2 de la base.

Ces indicateurs sont résumés par deux graphes. La courbe *ROC* (*Receiver Operating Characteristics*) qui caractérise la qualité de discrimination du processus de classification et la courbe de la précision (*P*) en fonction du rappel (*R*). Enfin la mesure F ou F-score proposée par Van Rijsbergen [9] et [17] évalue l'efficacité de la méthode à trouver les points du contour ; elle peut être complétée par l'exactitude (E), fraction des pixels bien classés. La mesure F exploite la précision et le « rappel » combiné par l'expression suivante :

$$F = \frac{2 * P * R}{P + R}.$$

Le rappel est le nombre de pixels correctement attribués aux contours par rapport au nombre total de pixels des contours.

La précision est le nombre de pixels correctement attribués aux contours rapporté au nombre de pixels attribués aux contours.

La courbe *ROC* exploite la sensibilité (*Sen*) qui mesure la capacité de trouver les pixels des contours et la spécificité (*Sep*), capacité d'écarter les pixels des « non contours ».

Ces cinq quantités, *P*, *R*, *E*, *Sen*, *Sep* sont calculées à l'aide de la matrice de confusion, donnée par le tableau suivant :

		Référence	
		Classe A	Classe $\overline{A}$
Prédiction	Classe A	*VP*	*FP*
	Classe $\overline{A}$	*FN*	*VN*

VP est le nombre de pixels de l'image correctement attribués aux contours, *FP* le nombre de pixels de l'image incorrectement attribués aux contours, *FN* le nombre de pixels de « non contours » attribués aux contours et *VN* les pixels de « non contours » correctement attribués.

$$P = \frac{VP}{VP + FP}, \quad R = \frac{VP}{VP + FN}, \quad E = \frac{VP + VN}{VP + FP + VN + FN},$$

$$Sen = \frac{VP}{VP + FN}, \quad Sep = \frac{VN}{FP + VN}$$

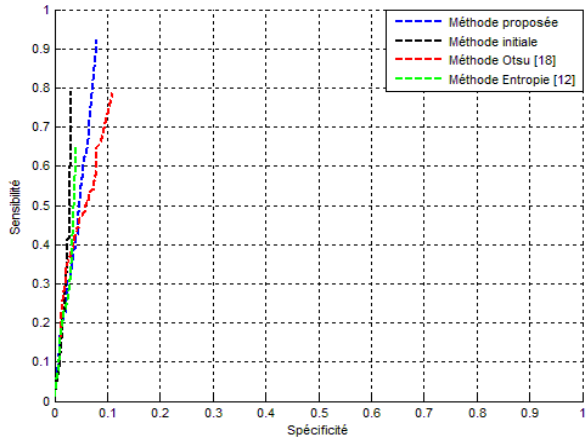


Figure 7 : Courbes ROC.

La Figure 7 présente la courbe *ROC* pour chaque méthode de calcul automatique du seuil de binarisation : la méthode d'OTSU [18], une méthode à base d'entropie [12], la méthode initiale [6] et la nouvelle méthode.
La Figure 8 présente les courbes *P-R* de ces méthodes placées dans le graphe des méthodes testées par le groupe vision de l'Université de Berkeley, le F-score de notre méthode est précisé. Le calcul est réalisé sur une centaine d'images tirées aléatoirement de la base BSDS300. Le point vert indique la mesure F des experts.

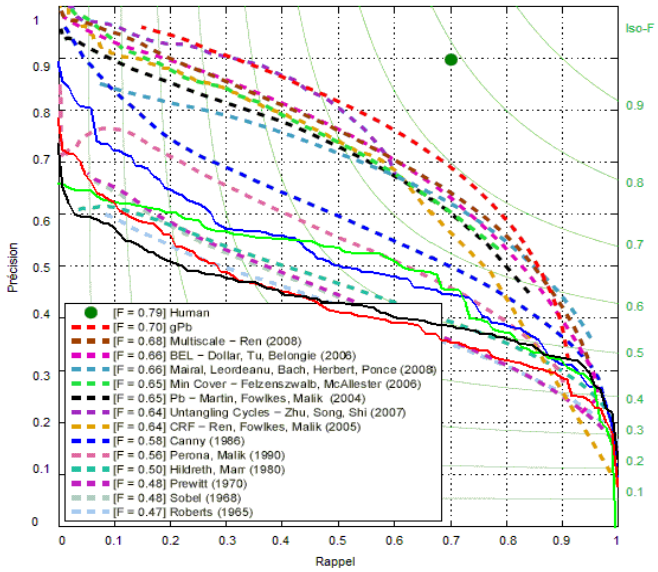


Figure 8 : Courbes P-R : en rouge continu la méthode Otsu [18] avec [F-mesure = 0.50], en bleu continu, la méthode que nous proposons avec [F-mesure= 0.56], en vert continu la méthode d'entropie [12] avec [F-mesure = 0.57], et en noir continu la méthode initiale avec [F-mesure = 0.48]

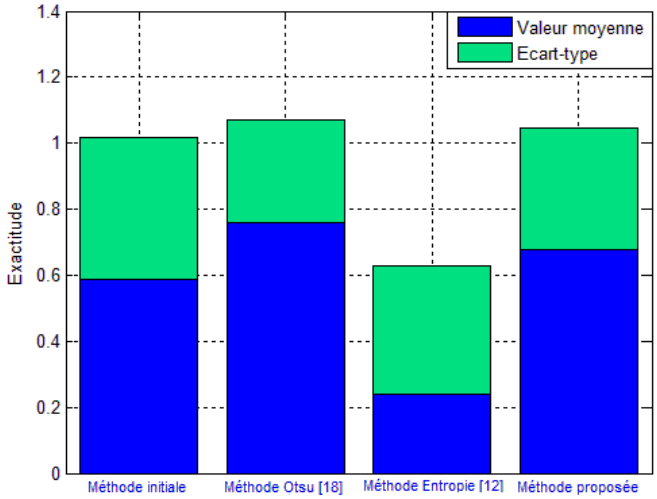


Figure 9 : Diagramme d'Exactitude : en bleu la valeur moyenne d'exactitude, et en vert l'écart-type correspondant. Le calcul a été fait sur les centaines d'images de la base BSDS300.

La Figure 9 présente l'évaluation de l'exactitude pour les quatre méthodes de binarisation. On retrouve en abscisse la méthode utilisée et en ordonnée la valeur moyenne de (E) et l'écart-type du processus de calcul d'exactitude sur la même centaine d'images de la base BSDS300.

Les figures 10 et 11 donnent des exemples de résultats avec dans la première ligne les contours de référence, la deuxième ligne les contours obtenus par notre méthode, la troisième par l'entropie et la quatrième par la méthode [18].

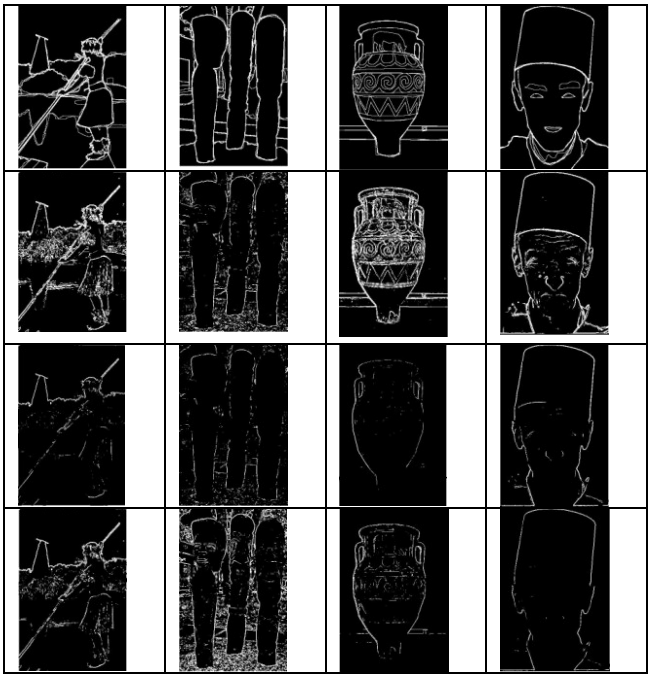


Figure 10 : Résultat sur l'extrait 1.

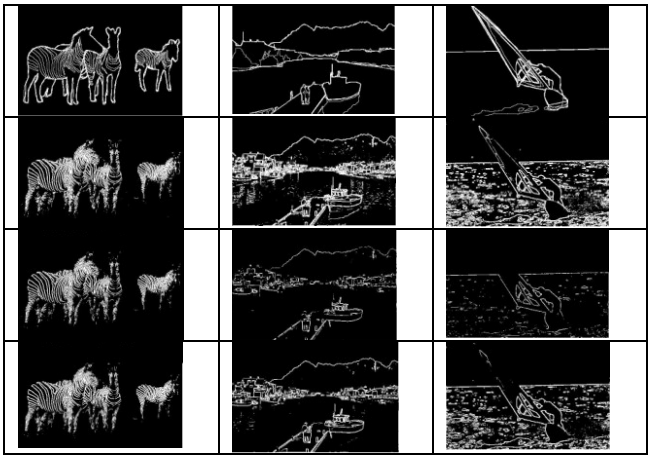


Figure 11 : Résultat sur l'extrait 2.

Enfin, nous avons étudié l'impact de la complexité temporelle (CPU Time) toujours pour la même banque d'images de la base BSDS300 sur une station de calcul

possédant un processeur Xeon 2.53Ghz (avec 4 Go de mémoire vive).
Pour l'ensemble des images nous évaluons le temps moyen de calcul du seuil de binarisation qui est présenté sur la Figure 12 sous forme d'un diagramme à barres. L'axe des abscisses porte le nom de l'algorithme utilisé et l'axe des ordonnées le temps moyen de calcul

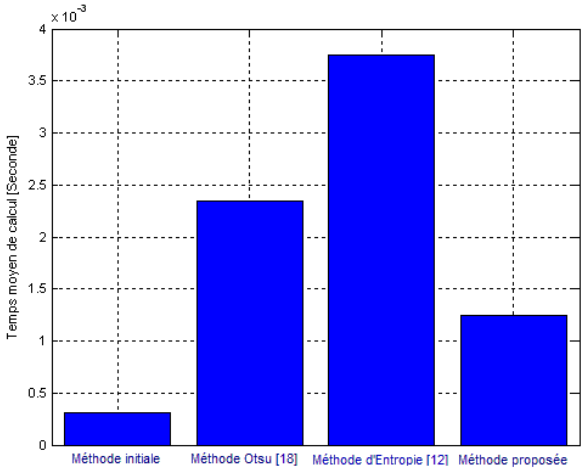


Figure 12 : L'évaluation du temps de calcul de seuil de binarisation. Notre algorithme proposé est le plus rapide par rapport a celui d'OTSU [18] ainsi d'entropie [12].

du seuil de binarisation correspondant exprimé en secondes.

La Figure 13 donne les résultats de binarisation obtenus par l'application de la nouvelle méthode sur l'image de notre montage d'usinage.

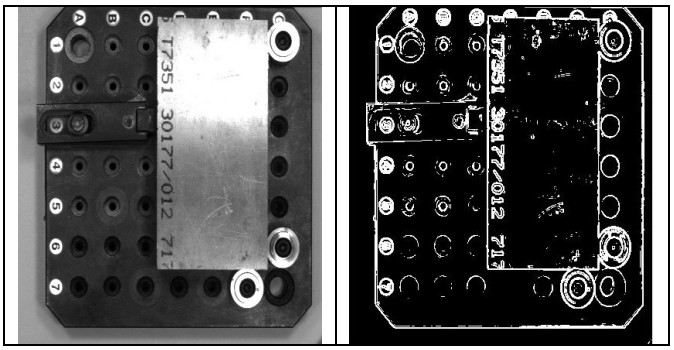


Figure 13 : Seuillage automatique : Application de la méthode proposée sur l'image de montage d'usinage.

5 Discussion

Nous proposons dans cette section de comparer le résultat de binarisation des quatre méthodes utilisées. Sur la Figure 7 nous constatons que la courbe ROC de notre méthode produit son maximum pour une valeur plus petite de spécificité que la méthode d'OTSU mais plus grande que la méthode initiale et d'entropie, et se situe bien au dessus en sensibilité. La courbe ROC idéale est à 1 pour toute spécificité. Sur la Figure 8, courbe P-R, nous constatons que notre méthode se situe au niveau de la méthode d'entropie et légèrement supérieure à celle d'OTSU avec un F-score honorable de 0,56. Enfin, l'étude de l'exactitude confirme le bon classement de notre méthode par rapport aux autres, avec une valeur moyenne autour de 68% et un écart-type de 37%, alors que la méthode d'entropie et la méthode initiale ont leur valeur moyenne autour de 25% et 59% et un écart-type de 39% et 43% respectivement. Cette expérience montre la capacité de chaque méthode a bien détecter les pixels des contours en écartant les pixels des non-contours. Enfin nous avons étudié l'impact de facteur du temps de calcul, dont les résultats présentés sur la Figure 13 montrent que la méthode proposés est beaucoup plus rapide que celle d'OTSU [18] et d'Entropie [12].

6 Conclusion

Notre méthode de seuillage automatique des modules des gradients par modélisation de l'histogramme donne des résultats aussi bons voire meilleurs que les méthodes existantes selon le critère de performance utilisé.
En outre, notre algorithme n'étant pas itératif comme les deux autres méthodes évaluées, il est facilement implantable sur des circuits FPGA. Les résultats de la Figure 13 restent bruités, et nous devons maintenant évaluer l'impact de ce bruit sur la qualité de la reconstruction 3D qui sera obtenue à partir des images de contours générées.

Bibliographie

[1] Giuseppe Papari, Nicolai Petkov, "Edge and line oriented contour detection: State of the art". Image and Vision Computing, n°29, pp 79–103, 2011.

[2] Debashis Sen, Sankar K. Pal, "Gradient histogram: thresholding in region of interest for edge detection", Image and vision computing journal, n° 28, pp 677-695, 2010.

[3]http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html#bsds500

[4] J.V. Leroy, Th. Simon, F. Deschesnes, "Real time monocular depth from defocus", ICISP, Cherbourg, juillet 2008.

[5] M. Léard, M. khoudeir, J. Brochard, J. Bernard, “Depth measurement of textured surface in a 3D scene”, IEEE International Conference on Image Processing, Rochester (New York), pp. II-469- II-472 vol.2, 22-25 September. 2002.

[6] T. Simon, M. Khoudeir, B. Heit, J. Bernard, “Image segmentation by setting thresholds for the magnitude of gradient”, ICECS'94-IEEE, Electronics, Circuits and Systems, Egypte, pp 40-45, 1994.

[7] "Méthodes adaptatives pour le signal (outils mathématiques et mise en œuvre des algorithmes", François Michaut, Traité des nouvelles technologies, série traitement du signal, Hermès, Paris 1992, pp 82-108.

[8] P. Arbelàez, M. Maire, C. Fowlkes and J. Malik, “Contour Detection and Hierarchical Image Segmentation”, Transactions on PAMI, IEEE, Août 2010.

[9] Swets, J.A. (1996). Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers. Mahwah, NJ: Erlbaum.

[10] M. Sezgin, B. Sankur**, “**Survey over image thresholding techniques and quantitative performance evaluation”, Journal of Electronic Imaging, January 2004.

[11] Henstock,, Peter V. and Chelberg, David M., “Automatic Gradient Threshold Determination for Edge Detection Using a Statistical Model A Description of the Model and Comparison of Algorithms”, ECE Technical Reports, Paper 95, 1996.

[12] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, ‘‘A new method for gray-level picture thresholding using the entropy of the histogram”, Graph. Models Image Process. 29, pp 273–285, 1985.

[13] D. Marr, E. Hildreth. Theory of Edge Detection. Proc. R. Soc. London, Ser. B, Biol. Sci., Vol. 207, pp 187-217, 1980.

[14] John Canny. “A computational approach to edge detection. Pattern Analysis and Machine Intelligence, IEEE Transactions on, PAMI-8(6):679–698, Nov. 1986.

[15] K. R. Namuduri, R. Mehrotra and N. Ranganathan, “Edge Detection Models Based on Gabor Filters,” International Conference on Image, Speech and Signal Analysis, vol.3, pp.729-732, 1992.

[16] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes : Active contour models,” International Journal of Computer Vision 1, pp 321–331,1988.

[17] D.P. Huijsmans and N. Sebe, “Extended performance graphs for cluster retrieval, institute of advanced computer science”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01) - Volume 1, Kauai, Hawaii, December 08-December 14, 2001

[18] Yi Wan, Jiangchang Wang, Xingbo Sun and Ming Hao: A Modified Otsu Image Segment Method Based On th Rayleigh Distribution. Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference.

[19] Michael Randolph Maire, PhD thesis “Contour Detection and Image Segmentation”, UNIVERSITY OF CALIFORNIA, BERKELEY, 2009.