

Evaluation sans référence de la segmentation en contours

A.CHERFA-BOUZOUAD¹, Y.CHERFA¹, S. AMIR², A. TAMERA²

Laboratoire LARIM, Département d'électronique / Université de Blida, Rte de Soumaa, BP.270,
Blida, Algérie

Tel.: (213) 73 02 01 91, Fax: (213) (25) 43 11 64,

Email: ¹{assia_bz, cherfa_yazid}@yahoo.fr ²{samir.amir, abd_tam}@yahoo.fr

RESUME

La segmentation des images est une étape clé en traitement d'images et différents algorithmes de segmentation ont été proposés. Mais généralement une technique de segmentation est plus ou moins adaptée à un certain type d'image. Pour cela il est intéressant de choisir la technique de segmentation la plus appropriée à notre application. Dans cet article nous allons proposer une méthode d'évaluation sans référence de la segmentation en contours d'images basée sur le calcul d'un indice tenant compte des caractéristiques de l'image segmentée et une méthode d'évaluation par comparaison de deux images segmentées. Les performances des deux méthodes sont évaluées en utilisant des images synthétiques et des images réelles.

1. INTRODUCTION

Le traitement d'image a pour vocation l'étude, la conception et la réalisation de système d'exploitation d'images considérées comme véhicule de l'information. Son but ultime est d'extraire le contenu informationnel (information pertinente) des images. Plusieurs méthodes de segmentation se sont développées ces dernières décennies et les chercheurs ont ressenti le besoin d'avoir une mesure de qualité de la segmentation pour l'évaluation des résultats de la segmentation.

Le problème de l'évaluation de la segmentation est crucial car il permet au chercheur de :

- comparer une méthode de segmentation par rapport à d'autres ;
- choisir l'algorithme le plus adéquat à son application et de régler ses paramètres en fonction du problème à résoudre car le paramétrage des algorithmes est un problème bien connu des traiteurs d'images.

Pour comparer de façon objective des méthodes de segmentation, il est préférable d'utiliser des images de synthèses dont les contours et les régions sont parfaitement localisées, mais de telles images ont l'inconvénient de ne pas reproduire toutes les situations possibles, de ne pas tenir suffisamment compte de la réalité, et de ne pas être toujours disponibles .

Les méthodes de mesure de la qualité de la segmentation qui existent actuellement peuvent être regroupées en trois grandes catégories :

- les techniques d'évaluation par un expert humain (visuelles);
- les techniques d'évaluation avec référence où l'on fait appel à des images de synthèses construites à la main par un expert du domaine d'application ;
- les techniques d'évaluation sans référence, dans le cas où l'on ne dispose pas de la vérité terrain et qu'il n'y a pas suffisamment d'informations a priori sur l'images.

Certaines techniques d'évaluation mesure un indice de « bonne qualité » de la segmentation, cet indice est d'autant plus élevé que la segmentation est bonne, d'autres mesurent la "différence " entre l'image segmentée et l'image de référence et dans ce cas cette différence est d'autant plus faible que la segmentation est bonne (proche de la référence). Mais en ce qui concerne la référence, elle ne peut être définie de façon unique que si l'image initiale est simple.

2. LES DIFFERENTES TYPES DE TECHNIQUES D'EVALUATION

2. 1. Evaluation visuelle [1]

L'évaluation visuelle est effectuée par un expert humain qui est amené à regarder l'image initiale et l'image segmentée. Cette évaluation est très subjective car elle dépend des critères utilisés par l'expert humain. Ces critères vont dépendre du domaine d'intérêt de l'expert, du but précis qu'il cherche à atteindre, de plus celui-ci a souvent une idée globale de l'image et approximative de ce qui doit être détecté. Cette évaluation dépendra de ce à quoi s'intéresse l'expert : objets, zones homogènes ou texturées, discontinuité. L'évaluation visuelle est qualitative, elle permet de dire si un algorithme fonctionne, mais ne permet pas de comparer plusieurs algorithmes. Pour améliorer la qualité de cette comparer

plusieurs algorithmes. Pour améliorer la qualité de cette évaluation, on peut faire appel à plusieurs experts et ne considérer que les résultats cohérents. Certains auteurs ont faits une évaluation hybride analytique et visuelle [2][3]

2. 2. Evaluation avec référence [4] [5] [6] [7] [8]

2.2.1. Evaluation de la segmentation en contours

Un grand nombre d'algorithmes d'analyse d'images s'intéressent à la détection des contours. Un certain nombre d'algorithmes d'évaluation ont été faits dans ce sens. Généralement, ces méthodes sont basées sur des calculs de distance entre les pixels contour de l'image segmentée et l'image de référence.

2.2.2. Evaluation de la segmentation en régions

Les algorithmes d'évaluation de la segmentation sont basés soit sur des calculs de distances (entre pixels mal segmentés et les régions auxquelles ils appartiennent ou distance entre) soit sur les couples de régions ayant un recouvrement maximal.

2. 3. Evaluation sans référence [8]

Ce type de mesure d'évaluation est fondé uniquement sur l'image initiale et sur les caractéristiques du résultat, sans avoir accès à une segmentation de référence. Il se base sur des critères d'évaluation non supervisés permettant d'estimer la qualité de résultat de segmentation à partir de statistiques calculées sur chaque région ou contour détecté. Il est possible de trouver une liste exhaustive de tels critères. L'aptitude de ces méthodes reste assez méconnue surtout dans le cas où l'image à segmenter contient des zones texturées, comme la plupart des images naturelles.

2.3.1. Evaluation de la segmentation en contours :

Levin et Nazif ont proposé une mesure basée sur le contraste, Kitchen et Rosenfield ont utilisé la cohérence locale des contours. La mesure de Tan, Gelfand et delp ont est plus complexe que la précédente, elle essaye de combiner des critères liés à l'image initiale et à la structure locale des contours.

2.3.2. Evaluation de la segmentation en régions

Les mesurent proposées sont basés sur l'uniformité des régions, sur le contraste, des travaux ont été faits pour évaluer la segmentation des images couleurs.

2. 4. Evaluation des images seuillées [8]

Parfois les images que l'on utilise ne sont pas segmentées en contours ou en régions, mais seuillées (images binarisées). Il serait intéressant d'avoir des algorithmes

d'évaluation qui permettront de choisir les seuils les plus adéquats. On trouve différentes mesures permettant d'évaluer de telles images, basées sur la forme, l'entropie, les probabilités de mauvaise classification des pixels et les caractéristiques des objets segmentés.

3. LES METHODES D'EVALUATION PROPOSEES

Nous allons présenter dans cette partie les deux méthodes que nous proposons pour l'évaluation de la segmentation sans référence : la première consiste à combiner entre la continuité, la fermeture et la longueur des contours et la seconde fait une comparaison entre deux résultats de segmentation en contours afin de classer ces deux cartes contours par ordre de qualité.

3.1. Première méthode

Cette mesure est fondée sur la cohérence locale des contours qui repose sur la méthode de Kitchen et Rosenfield. Elle utilise deux critères de caractérisation qui sont la continuité, la fermeture et la longueur des contours car nous savons qu'une segmentation idéale donnera uniquement des contours fermés, de plus les petits contours sont souvent dus à une mauvaise segmentation (segmentation d'images texturées). La carte contour est évaluée par un score qui est la moyenne des scores locaux calculés en chaque point contour de l'image. Le score local E est obtenu par une combinaison linéaire entre un score pour la continuité (C) et un score pour la fermeture et la longueur des contours (F).

$$E = \gamma C + (1 - \gamma)F \quad (1)$$

C : Critère de continuité

γ : Coefficient compris entre 0 et 1

$$F = \frac{(\text{Nombre contours} + \text{Nombre contours} > L)}{(\text{nombre total de contours})} \quad (2)$$

Le critère de continuité est parfaitement satisfait (C=1) (critère de Kitchen et Rosenfeld). Le critère F est totalement satisfait (F=1) si tous les contours sont fermés ou ont une longueur supérieure à L. Selon que ces deux conditions sont plus ou moins bien remplies, les deux scores prennent des valeurs dans l'intervalle [0,1].

3.2. Deuxième méthode

Nous savons que pour une même image nous pouvons avoir des images segmentées différentes selon l'algorithme utilisé. L'idée consiste à faire une comparaison entre deux cartes contours qui sont des résultats de deux méthodes de segmentation différentes appliquées sur la même image afin de classer ces deux

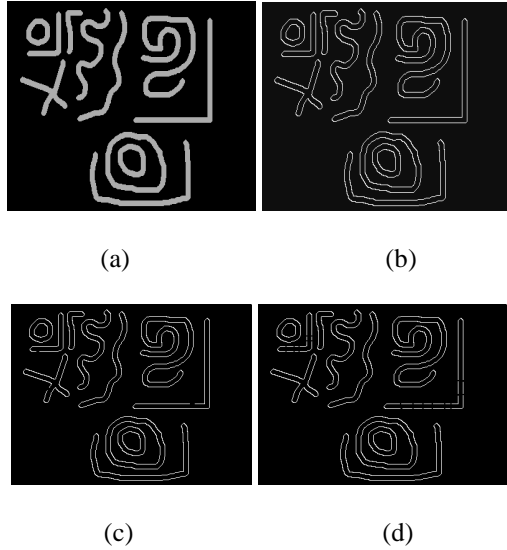


Figure 1. Images tests utilisées

(a) Image originale, (b), (c) et (d) Images segmentées en contours avec respectivement pas de discontinuité, quelques discontinuités, plusieurs discontinuités

cartes par ordre de qualité (meilleure segmentation). Ces deux images segmentées présentent surement des points communs correspondants probablement à une bonne segmentation (nous supposons que la probabilité pour que des pixels soient mal segmentés par les deux algorithmes est faibles), et des points litigieux correctement segmentés soit par la première technique de segmentation soit par la seconde. Pour cela nous ferons la soustraction des deux cartes, et le résultat donnera une carte qui ne contient que les pixels segmentés différemment. Un pixel est considéré comme bien segmenté s'il obéit au critère de continuité. Soient $S1$ et $S2$ les deux scores tirés des deux cartes contours et qui nous permettent de classer ces cartes par ordre de qualité.

$$S1 = \frac{NC + N1 - Nf1}{NC + N1 + N2 - Nf2} \quad (3)$$

$$S2 = \frac{NC + N2 - Nf2}{NC + N1 + N2 - Nf1} \quad (4)$$

NC : nombre de pixels communs aux deux cartes

N1 : nombre de pixels contours détectés dans la première carte et non détectés dans la seconde.

N2 : nombre de pixels contours détectés dans la seconde carte et non détectés dans la première.

Nf1: nombre de pixels faux (isolés) dans la première carte.

Nf2: nombre de pixels faux (isolés) dans la deuxième carte

L'image segmentée qui donne le score le plus élevé correspondra à l'image la mieux segmentée. L'image segmentée idéale a un score de 1.

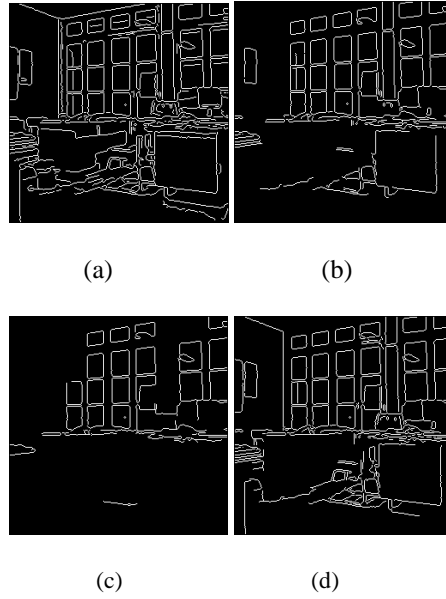


Figure 2. Image bureau segmentée par Canny avec différents jeux de paramètres

3. RESULTAT DE L'EVALUATION

Nous avons testé notre premier algorithme sur des images synthétiques, réalisées manuellement. L'image (a) est l'image originale, les images (b), (c) et (d) correspondent à des images segmentées en contours avec différents degrés de discontinuités, pour modifier le nombre de contours fermés et leurs longueurs. La table 1 donne les scores obtenus pour chaque image.

Images	NF	NSL	NIL	C	L	F
(a)	14	0	0	0,896	0	0,963
(a)	11	3	0	0,895	0	0,963
(b)	11	2	1	0,895	259	0,916
(b)	11	1	2	0,895	375	0,870
(c)	12	26	0	0,884	0	0,959
(c)	12	14	12	0,884	14	0,754

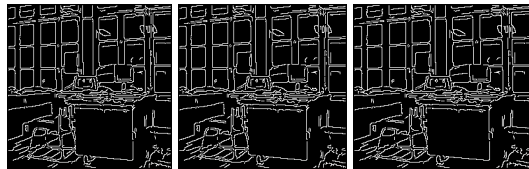
Table 1. Tableau des scores d'évaluation de la segmentation

C : Score de continuité.

F : Score final par la méthode proposée.

L : Longueur minimale d'un contour pour qu'il soit pris en considération.

Nous remarquons que le résultat de l'évaluation est élevé pour les images (b) et (c), car ces images ne contiennent que des contours fermés ou des contours longs (supérieurs à L), puis le résultat diminue lorsque le nombre de contours non fermés inférieurs à L augmente, ce qui est tout à fait normal puisque les contours non fermés ou petits sont dus à une mauvaise segmentation ou à la présence de zones fortement texturées dans l'image. L'avantage de la méthode est que l'utilisateur peut agir sur la valeur de L et donc donner de faibles valeurs à L en cas où il dispose d'une image présentant des détails fins.



(a) Canny (b) Deriche (c) Shen-Castan



(d) Prewitt (e) Sobel (d) Roberts

Figure 3. Résultats d'images segmentées par différents algorithmes

Pour tester notre deuxième algorithme nous avons fait deux manipulations : la première consiste à appliquer l'algorithme sur des images segmentées par la méthode de Canny mais avec un jeu de paramètres différents (Table 2) afin de choisir ceux qui donnent la meilleure segmentation, la seconde consiste à appliquer l'algorithme sur des images segmentées par des méthodes différentes (Table 3), afin de classer ces méthodes par ordre de qualité.

Images	σ	Sh	Sb	Classement
Image (a)	0,74	0,1	0,28	1
Image (b)	0,60	0,3	0,68	3
Image (c)	0,50	0,50	0,50	4
Image (d)	0,53	0,20	0,30	2

Table 2. Résultat de l'évaluation de l'image bureau segmentée par Canny avec différents paramètres

σ : L'intensité du filtrage gaussien.

Sh: Le seuil haut d'hystérésis.

Sb: Le seuil bas d'hystérésis.

D'après le classement, l'image (a) correspond à la meilleure segmentation donc au meilleur jeu de paramètres. Les images (b) et (3) sont les moins bien classées car la segmentation a fait perdre beaucoup d'informations, nous avons une sous-segmentation bien visible.

Algorithmes de segmentation	Classement
Canny	1
Deriche	2
Shen-Castan	3
Prewitt	4
Sobel	5
Roberts	6

Table 3. Classement des six algorithmes de détection de contours

Nous avons classés par notre méthode six algorithmes de segmentation, selon les scores obtenus. Le classement est confirmé par la bibliographie. Nous remarquons une nette dégradation pour Prewitt, Sobel et Roberts.

4. DISCUSSION ET CONCLUSION

Les deux algorithmes proposés ont donné des résultats cohérents, ils sont relativement simples à mettre en oeuvre. Le premier algorithme s'applique bien aux images n'ayant pas de contours très complexes. Le deuxième est particulièrement bien adapté pour choisir les meilleures paramètres d'une segmentation, il pénalise la sous-segmentation, mais aussi la présence de pixels isolés dus à une sur-segmentation, de plus il ne pénalise pas la localisation des contours qui peuvent être mis d'un côté de la région pour un algorithme de segmentation et de l'autre par un autre algorithme.

La recherche de critères permettant une bonne évaluation reste encore un domaine ouvert, il faut penser à introduire des critères qui tiennent compte du problème de croisement des contours.

5. BIBLIOGRAPHIE

- [1] I.H. Jermyn, C.W. Shaffrey, N.G. Kingsbury, "Evaluation Methodologies for images retrieval", Proceedings of ACIVS 2002, Ghent, Belgium, September 9-11,2002
- [2] A. Cavallaro, E. Drelie Gelasca, T. Ebrahimi, " Objective evaluation of segmentation quality using spatio-temporal context", Proceedings of IEEE, International Conference on Image Processing, Rochester (New-York), 22-25,2002.
- [3] Barranco-Lopez,R Roman- Roldan, R.& G. Lopera, J. Francisco, J. Martinez-Arosa, " A mesure of quality for Evaluating methods of segmentation and edge detection", Pattern Recognition, Vol. 34, No.5, PP. 969-980, 2001.
- [4] V. Chalana and Y. Kim, " A methodology for evaluation of boundary detection algorithms on medical images", IEEE Trans. On Medical Imaging, October 1997.
- [5] C. Alberola-Lopez, M.Martin-Fernandez and J. Luiz-Alzola, " Comments on: A methodology for evaluation of boundary detection algorithms on medical images", IEEE Trans. on Medical Imaging, Vol. 23, no 5, may 2004
- [6] P. Bamford, " Automating Cell Segmentation with annotated Examples", WDIC 2003, Friday 7th February, University of Queensland, S^t Lucia, Brisbane.
- [7] G. Bueno, M. Fisher, K. Burnham, J. Mills and Olivier Hass, "Automatic segmentation of clinical structures for RTP: Evaluation of a morphological approach" Conf. on Medical Image Understanding and Analysis 2001, 16-17 July 2001, Birmingham
- [8] Y.J. Zhang, «A survey on evaluation methods for images segmentation", Pattern Recognition, Vol. 29, No. 8.