



Segmentation d'images par multi-seuillage et fusion de régions labellisées minimisant un critère de similarité

Chafik KERMAD, Kacem CHEHDI et Claude CARIOU

Groupe Image LASTI-ENSSAT, 6, rue de Kerampont, 22300 Lannion Cedex, FRANCE.
E-mail: kermad@enssat.fr, chehdi@enssat.fr, cariou_c@enssat.fr

RÉSUMÉ

Une méthode de segmentation fondée sur une combinaison de deux processus de transformation de l'image est présentée. La première transformation résulte d'une classification automatique par multi-seuillage des niveaux de gris de l'image suite à une analyse d'histogrammes locaux. La seconde transformation est basée sur une procédure de fusion de régions minimisant un critère de similarité. Celui-ci tient compte de la dynamique des histogrammes et des statistiques locales calculées à la fois sur l'image originale et sur celle des étiquettes résultantes de la première transformation. Cette méthode a été testée sur différentes catégories d'images et les résultats obtenus sont satisfaisants.

1. Introduction

La segmentation est l'une des phases les plus importantes dans la chaîne d'analyse d'images. La qualité des résultats obtenus à l'issue de cette opération détermine la qualité globale de tout le système d'interprétation.

Le problème de la segmentation d'image est largement abordé dans la littérature mais les résultats dans le cas général restent non concluants [1] [2] [3] [4] [5]. Les résultats peu satisfaisants sont, en grande partie, dus aux problèmes d'extraction d'informations caractéristiques permettant d'augmenter la séparation entre les différents modèles de texture et à la fluctuation des niveaux de gris de zones homogènes.

La grande diversité des algorithmes de segmentation rend nécessaire la recherche d'une méthode combinant plusieurs techniques et fusionnant différents types d'informations afin d'aboutir à une segmentation fiable. Dans ce sens, nous présentons une approche de segmentation qui traite les images à plusieurs niveaux. Elle se base dans un premier temps sur l'utilisation d'informations issues d'histogrammes locaux de l'image pour engendrer une pré-segmentation au sens des niveaux de gris. L'aspect textuel est ensuite pris en compte et la qualité de la segmentation est améliorée par une analyse plus étendue des régions issues de la pré-segmentation. Enfin, les frontières séparant les différentes régions sont affinées et localisées avec précision.

ABSTRACT

A segmentation method which combines two transformation processes of the image is presented. The first transformation allows a pre-segmentation by an automatic classification of image intensities. This procedure is based on a multi-thresholding technique of a global histogram formed from local histograms. The second transformation is a region merging operation which minimizes a similarity criterion. The fusion criterion takes into account grey levels of original image and labels resulting from the pre-segmentation stage. This method has been tested on several images and results are satisfactory.

2. Méthode développée

2.1 Pré-segmentation par multi-seuillage

Les techniques de segmentation par seuillage d'histogramme sont très utilisées pour isoler les différentes zones de l'image [6] [7].

L'approche que nous adoptons ici consiste à classer les points de l'image en comparant leurs niveaux de gris aux seuils déterminés par l'analyse d'un histogramme global construit à partir d'informations particulières issues d'histogrammes locaux. L'objectif de cette étape est d'éliminer l'information non pertinente afin de mettre en évidence les principaux éléments constituant les textures et d'homogénéiser les régions en niveaux de gris. Cette étape permet de faciliter la procédure de segmentation finale [8].

La procédure de pré-segmentation est effectuée en deux phases: *construction de l'histogramme global et classification des niveaux de gris de l'image.*

2.1.1 Construction de l'histogramme global

Les histogrammes locaux sont transformés de manière à mettre en évidence leurs principaux modes [9]. Les modes significatifs des histogrammes permettent de donner une indication sur le nombre de classes de niveaux de gris qui seront retenues. Les modes validés serviront alors pour la construction de l'histogramme global des maxima locaux. Cette série d'opérations est effectuée en utilisant plusieurs tailles de



fenêtres (16x16, 32x32, 64x64) afin de tenir compte de la résolution spatiale de chaque image et de rendre ainsi le traitement adaptatif. Les deux étapes pour la construction de l'histogramme global sont détaillées ci-dessous.

Etape 1: Transformation des histogrammes locaux

L'image originale notée I , de dimension $N \times N$ est divisée en blocs B_c chacun de taille $d \times d$.

Les histogrammes locaux de l'image originale présentent des distributions rarement régulières. Les modes de l'histogramme, se recouvrent trop pour laisser apparaître des pics significatifs bien séparés. Les limites de chaque mode sont rarement bien définies, les extrema sont alors difficiles à déterminer. Afin d'accentuer de manière significative la séparation des modes les plus importants, une procédure de transformation est employée. Pour cela, quatre critères basés sur la mesure des fluctuations des niveaux de gris sont utilisés. Ces critères notés $C1$, $C2$, $C3$ et $C4$, donnent une indication sur l'homogénéité d'un point par rapport à ses quatre voisins les plus proches:

$$\begin{aligned} C1 &= |U(x,y) - I(x-1,y)| \leq 1 \\ C2 &= |U(x,y) - I(x,y-1)| \leq 1 \\ C3 &= |U(x,y) - I(x,y+1)| \leq 1 \\ C4 &= |U(x,y) - I(x+1,y)| \leq 1 \end{aligned}$$

où $I(x,y)$ représente le niveau de gris du point (x,y) appartenant à B_c .

En fonction de ces critères, les fréquences d'apparition $h_c(i)$ des niveaux de gris $I(x,y)$ sont modifiées de la manière suivante:

$$\begin{aligned} h_c(i) &= h_c(i)+4 && \text{si } (C1 \text{ et } C2 \text{ et } C3 \text{ et } C4) \text{ est vrai;} \\ h_c(i) &= h_c(i)+3 && \text{si } (C1 \text{ et } C2 \text{ et } C3) \text{ ou } (C1 \text{ et } C2 \text{ et } C4) \text{ ou} \\ &&& (C1 \text{ et } C3 \text{ et } C4) \text{ ou } (C2 \text{ et } C3 \text{ et } C4) \text{ est vrai;} \\ h_c(i) &= h_c(i)+2 && \text{si } (C1 \text{ et } C2) \text{ ou } (C1 \text{ et } C3) \text{ ou } (C1 \text{ et } C4) \text{ ou} \\ &&& (C2 \text{ et } C3) \text{ ou } (C2 \text{ et } C4) \text{ ou } (C3 \text{ et } C4) \text{ est vrai;} \\ h_c(i) &= h_c(i)+1 && \text{si } C1 \text{ ou } C2 \text{ ou } C3 \text{ ou } C4 \text{ est vrai;} \\ h_c(i) &= h_c(i) && \text{sinon.} \end{aligned}$$

En tenant compte des relations spatiales, cette transformation d'histogramme permet de mettre en évidence les principaux modes de l'histogramme tout en diminuant les fluctuations des niveaux de gris.

Avant de déterminer les principaux pics de l'histogramme, une opération de lissage lui est appliquée. Celle-ci est nécessaire pour éliminer les irrégularités de l'histogramme. On utilise à cet effet un noyau gaussien de taille 5. Les niveaux de gris de l'histogramme lissé dont les fréquences d'apparition sont inférieures à la moyenne des $h_c(i)$ sont ensuite éliminés.

Etape 2: Recherche des pics significatifs

Les pics significatifs des histogrammes résultant des traitements précédents doivent vérifier les trois critères suivants:

- Un niveau de gris i est un pic de l'histogramme s'il présente une fréquence d'occurrence locale maximale. Ceci revient à vérifier le critère suivant:

$$h_c(i) > h_c(i-1) \text{ et } h_c(i) > h_c(i+1).$$

- La distance séparant deux pics significatifs doit être supérieure à un seuil S_d . Dans le cas contraire, on retient comme pic significatif celui dont la fréquence d'occurrence est la plus grande.
- La hauteur d'un pic doit être supérieure à une fraction S_h du maximum de l'histogramme. Ceci permet d'éliminer les pics de taille relativement faible.

Les valeurs de seuils S_d et S_h sont calculées de manière adaptative et représentent respectivement la moyenne des distances des couples de pics successifs et la moyenne des fréquences d'apparition des pics.

La liste des pics valides servira pour la construction de l'histogramme global des maxima locaux. Celui-ci, noté H_{glob} , présente une forme multi-modale bien prononcée (Fig. 1 (c)).

2.1.2 Classification des niveaux de gris

La classification est obtenue après une procédure de regroupement itérative des niveaux de gris opérant sur la base des fréquences de H_{glob} . Le but est de détecter les seuils s_k les plus significatifs en déplaçant itérativement les valeurs de niveaux de gris vers leurs centres de masse jusqu'à ce qu'il n'y ait plus de changement dans l'histogramme. Cette procédure est décrite ci-dessous.

Soit $H_0(i) = H_{glob}(i)$, $i \in [i_{min}, i_{max}]$, l'histogramme initial à traiter, et soit H_k , l'histogramme obtenu à l'itération k . Les fréquences d'occurrence des niveaux de gris à l'itération $k+1$ sont déterminées selon les relations suivantes:

$$H_{k+1}(i) = \sum_{s \in V_i} H_k(s) \cdot \delta(i - M_k(s)) \quad (\text{EQ 1})$$

où $\delta(i)$ représente le symbole de Kronecker.

et $M_k(s)$ désigne la moyenne glissante à l'itération k .

$$M^k(i) = \frac{\sum_{s \in V_i} s \cdot H^k(s)}{\sum_{s \in V_i} H^k(s)} \quad (\text{EQ 2})$$

V_i représente l'ensemble des $2m+1$ niveaux de gris voisins de i , il est défini par:

$$V_i = \{\max((i-m), i_{min}) \dots \min((i+m), i_{max})\}$$

La relation (EQ 1) traduit le déplacement d'une partie des masses de $H_k(i)$ vers leurs centres de gravité locaux. Le paramètre de taille, m , varie graduellement jusqu'à ce que des tailles successives fournissent des résultats similaires.

L'histogramme final obtenu est constitué de masses ponctuelles. A chaque masse correspond une classe et la position de chaque masse indique le niveau de gris qui sera affecté en tant qu'étiquette à chaque classe. Les seuils de segmentation s_k sont alors donnés par les moyennes des niveaux de gris déterminées par la position de ces masses. Dans la suite de l'article nous

utiliserons le terme image des étiquettes pour désigner le résultat de la pré-segmentation.

Cette première segmentation se révèle correcte dans le cas des images composées d'un certain nombre de régions bien contrastées et avec une faible variation des luminances à l'intérieur d'une même région. Mais les résultats s'avèrent peu intéressants dès que les éléments de l'image sont variés et nombreux, notamment dans le cas des images avec des conditions d'éclairage non uniformes ainsi qu'en présence de zones fortement texturées. Dans ce dernier cas, l'intérêt de cette première analyse réside dans l'extraction de l'information susceptible de ne contenir, compte tenu de l'aspect textuel de l'image, que les éléments dont l'influence est la plus significative pour la caractérisation des différentes régions de l'image. Ceci permet une première simplification des données avant de procéder à un traitement complémentaire.

Notons enfin, que la technique de multi-seuillage tient compte des propriétés locales d'homogénéité de chaque histogramme et ne demande aucune connaissance a priori sur l'image à segmenter.

2.2 Segmentation par fusion de régions

La technique de segmentation développée entre dans le cadre des méthodes opérant par fusion d'ensemble de points. Les difficultés à résoudre seront de deux ordres: d'une part, le choix d'un critère sur lequel sera basé le processus de regroupement des points de l'image, et d'autre part, la définition d'une stratégie efficace de regroupement. En considérant qu'une texture est une structure disposant de certaines propriétés spatiales homogènes et invariantes par translation, nous avons choisi pour la caractériser les étiquettes présentes dans le bloc et leurs fréquences d'occurrence.

Le problème de segmentation est ramené à la localisation des différentes régions de l'image labellisée ayant un comportement semblable au sens des étiquettes et de leurs fréquences d'apparition.

Soit $p_c(i)$ (resp. $p_n(i)$) la fréquence d'occurrence du niveau de gris i dans le bloc courant B_c (resp. bloc voisin B_n)

Soit V un voisinage constitué des b blocs précédents et des $2b+1$ blocs des q lignes supérieures. b et q sont des variables calculées en fonction de la taille du bloc d'observation (par exemple $b=40$ et $q=10$ pour $d=16$ et $b=80$ et $q=15$ pour $d=8$).

Soient NE le nombre de labels et NG le nombre de niveaux de gris de l'image.

Le principe du processus de fusion est le suivant: les statistiques locales sont calculées à l'intérieur des blocs d'observation de l'image des étiquettes et de l'image originale correspondante. Ensuite, l'algorithme examine les deux images bloc par bloc selon un schéma séquentiel de parcours. Pour chaque bloc on cherche dans le voisinage V le bloc qui minimise un critère de similarité.

Pour mesurer le degré de similarité entre blocs nous avons défini un indice faisant intervenir la somme des différences absolues

des moyennes des niveaux de gris des étiquettes (m_c et m_n) et les deux premiers moments (μ_x et σ_x , ($x=c,n$)) calculés sur les points de l'image originale. Cet indice de similarité est donné par l'expression suivante:

$$d(B_c, B_n) = |m_c - m_n| + |\mu_c - \mu_n| + |\sigma_c - \sigma_n| \quad (\text{EQ 3})$$

Deux blocs sont considérés comme similaires si leur indice de similarité est inférieur à un seuil δ . Ce seuil est réglé empiriquement et est de l'ordre de 20% à 40% la surface du bloc. Dans le cas où aucun bloc adjacent n'est similaire au bloc traité on entreprend la construction d'une nouvelle région dont celui-ci est le premier élément. Ce processus est réitéré jusqu'à balayage complet de l'image des étiquettes.

L'utilisation des histogrammes des valeurs des étiquettes et des deux premiers moments des niveaux de gris s'avère fiable dans certains cas de figures mais ils ne sont pas suffisants pour caractériser entièrement la texture. Cette technique présente l'inconvénient de ne pas tenir compte de l'arrangement des étiquettes et des relations spatiales qui les gèrent. Cette limitation engendre une faiblesse de leur pouvoir de discrimination notamment dans le cas de textures à caractère régulier. Pour remédier à cet handicap une analyse avec chevauchement spatial des fenêtres d'observation doit être effectuée. Un recours à des statistiques tenant compte des dépendances spatiales des points de l'image peut également s'avérer nécessaire pour une meilleure discrimination. Dans notre cas nous avons opté pour la première solution en faisant une analyse avec un déplacement de la fenêtre d'observation égale à la moitié de la taille du bloc dans les directions horizontale et verticale.

Pour déterminer avec plus de précision les contours séparant les différentes régions, les blocs situés sur les frontières des régions issues de la segmentation sont localisés. Ces blocs frontières sont ensuite divisés en sous-blocs puis pour chaque sous-bloc SB_c on cherche dans un voisinage en 8-connexité le sous-bloc SB_n qui minimise la quantité

$$d_{\text{hist}}(h(SB_c), h(SB_n)) = \sum_{i=0}^{NG} |p_c(i) - p_n(i)| \quad (\text{EQ 4})$$

et qui vérifie le critère

$$d_{\text{hist}}(h(SB_c), h(SB_n)) < s$$

où $h(SB_c)$, $h(SB_n)$ représentent respectivement les histogrammes des blocs SB_c et SB_n de l'image originale. s est un seuil de tolérance calculé en fonction de la taille du bloc. Cette opération est répétée jusqu'à un certain niveau de résolution.

3. Résultats

Nous présentons dans ce paragraphe quelques résultats expérimentaux obtenus sur une image synthétique composée des régions de faibles et de fortes textures extraites de l'album de



Brodatz [10] (D24, D29, D38, D68). La figure 1(d) donne un exemple de pré-segmentation après rehaussement de contraste. Une illustration de la méthode de segmentation par fusion de régions homogènes appliquée aux images des figures 1(a) et 1(d)

est donnée figure 1(e). La figure 1(f) illustre le résultat de la procédure de localisation des frontières sur la carte des régions de la figure 1(e).

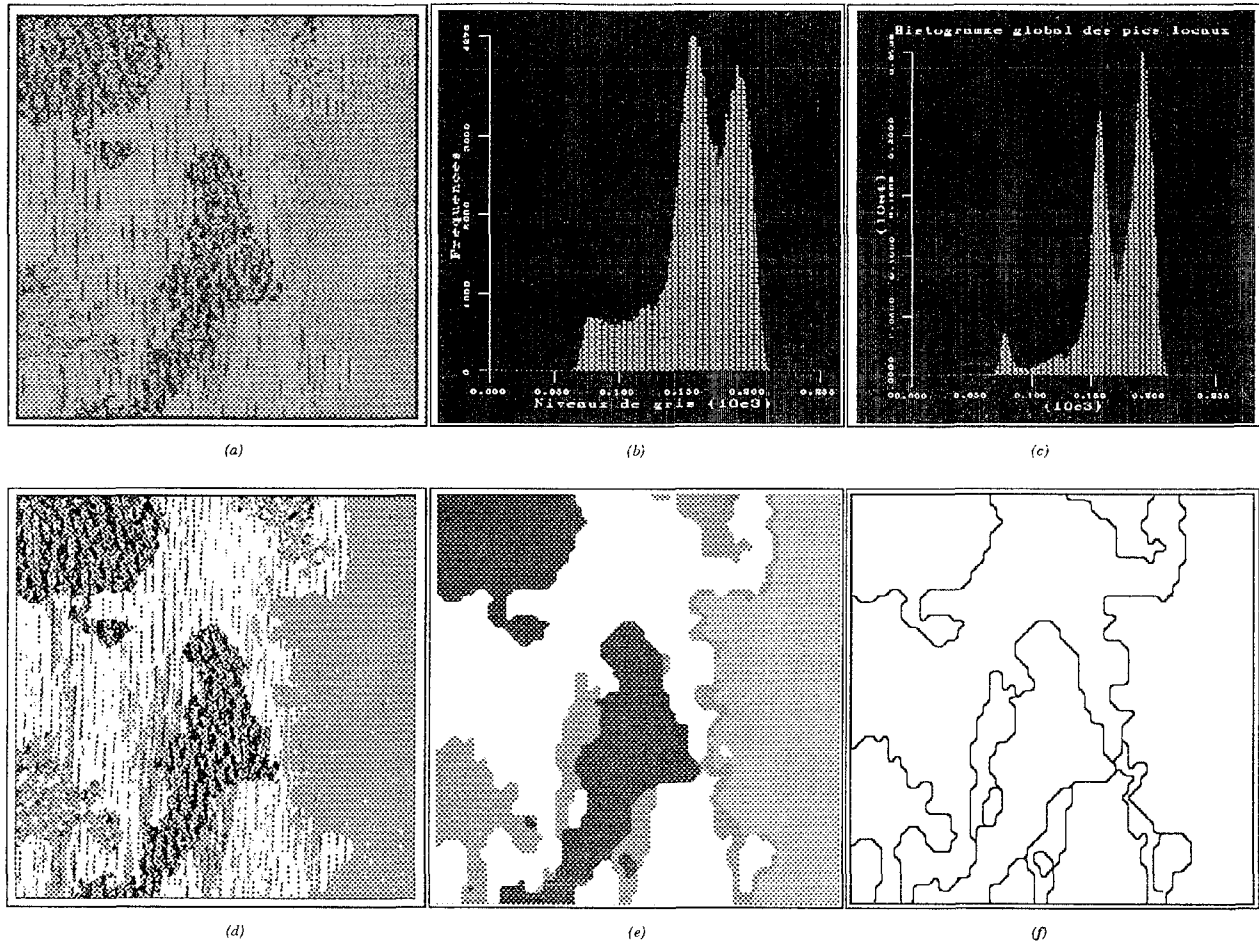


Figure 1. (a) Image originale composée des textures D24, D29, D38 et D68. (b) Son histogramme. (c) Histogramme global des maxima locaux (seuils de segmentation: 93, 119, 145, 175). (d) Résultat de la pré-segmentation (image rehaussée des étiquettes). (e) Résultat de la segmentation par fusion de régions après un filtrage médian, avec un seuil $\delta = 80$ et une fenêtre de taille $dxd = 16 \times 16$. (f) Frontières des régions.

4. Conclusion

La méthode présentée dans cet article a été expérimentée sur plusieurs catégories d'images. Elle donne des résultats satisfaisants, car les différentes régions d'une image sont localisées correctement, ce qui contribue à une meilleure modélisation et par conséquent à un pouvoir d'identification fiable. Ces résultats montrent également l'intérêt de la combinaison de différentes techniques dans le domaine de la segmentation. Notons enfin qu'elle est simple à mettre en œuvre et permet de segmenter les images en tenant compte de la non homogénéité des régions.

Références

- [1] Beveridge J.R, Griffith J.S, Kohler R.R, Hanson A.R., and Riseman E.M., Segmenting Images Using Localized Histograms and Region Merging, IJCV(2), No. 3, January 1989, pp. 311-352.
- [2] Chu C.C., and Aggarwal J.K., The Integration of Image Segmentation Maps Using Region and Edge Information, PAMI(15), No. 12, December 1993, pp. 1241-1254.
- [3] Manjunath B.S., and Chellappa R., Unsupervised Texture Segmentation Using Markov Random Field Models, PAMI(13), No. 5, May 1991, pp. 478-482.
- [4] Pal N.R., and Pal S.K., A Review on Image Segmentation Techniques, PR(26), No. 9, September 1993, pp. 1277-1294.
- [5] Reed T.R., and duBuf J.M.H., A Review of Recent Texture Segmentation and Feature Extraction Techniques, CVGIP(57), No. 3, May 1993, pp. 359-372.
- [6] Papamarkos N., Gatos B., A New Approach For Multilevel Threshold Selection, CVGIP-IP(56), No. 5, September 1994, pp. 357-370.
- [7] Sahoo P.K., Soltani S., Wong A.K.C., and Chen Y.C., A Survey of Thresholding Techniques, CVGIP(41), No. 2, February 1988, pp. 233-260.
- [8] Chehdi K., Cariou C., and Kermad Ch., Image segmentation and texture classification using local thresholds and 2D AR modelling, EUSIPCO'94, UK, September 1994, pp. 30-33.
- [9] Chehdi K., and Coquin D., Binarisation of various images by detecting local thresholds with a validation test, Proc. IEEE CCSP'91, Canada, pp. 611-614.
- [10] Brodatz P., Textures Album, New York, Dover, 1966.