

Guidage de contour actif par classificateur binaire supervisé pour la segmentation d'images texturées

Julien OLIVIER¹, Romuald BONÉ^{2,1}, Jean-Jacques ROUSSELLE¹, Hubert CARDOT¹

¹Université François Rabelais Tours
Laboratoire d'Informatique
64 avenue Jean Portalis, 37200 Tours, France
{julien.olivier, romuald.bone, rouselle, hubert.cardot}@univ-tours.fr

²École Nationale d'Ingénieurs
du Val de Loire
Rue de la Chocolaterie, BP 3410, 41034 Blois cedex, France

Résumé – Cette article propose un nouveau modèle de contour actif intégrant un classificateur binaire supervisé dans son équation d'évolution. Les caractéristiques de textures d'Haralick, extraites d'une image d'apprentissage comprenant une segmentation experte, sont utilisées pour réaliser l'apprentissage du classificateur. Une implémentation récente des ensembles de niveaux permet alors d'introduire le classificateur directement dans l'équation d'évolution du contour actif. L'approche proposée est appliquée à la segmentation d'images texturées 2D et 3D de différents types. Trois classificateurs sont testés et leur précision est comparée à celle du modèle de Chan et Vese à l'aide de plusieurs critères d'évaluation.

Abstract – This paper proposes a new active contour model driven by a supervised binary classifier. Haralick texture features are extracted from a learning image with an expert segmentation and used to carry out the learning task of the classifier. Then, a recent level set implementation allows the classifier to be introduced in the motion equation of the active contour. The proposed approach is applied to different kind of 2D and 3D textured image segmentation problems. Three classifiers are tested and they are compared to the Chan and Vese model in terms of precision using several evaluation measures.

1 Introduction

Introduit par Kass *et al* dans [1] les contours actifs sont de puissants outils de segmentation utilisés fréquemment dans la communauté de l'analyse d'image. Un contour actif est une courbe paramétrée $C(s, t)$, de paramètre s , initialisée dans une image Ω , évoluant dans le temps t de manière à minimiser une fonctionnelle d'énergie de la forme :

$$E(C(s, t)) = \int f(s, t) ds \quad (1)$$

f est une fonctionnelle composée de termes contrôlant la géométrie de la courbe et de critères d'adéquation du modèle à l'image. Les fonctionnelles peuvent être séparées en deux familles : les modèles basés contour [1, 2, 3] et basés région [4, 5, 6]. Les premiers n'utilisent que l'information présente sur la courbe, f étant intégrée uniquement sur C . Ces modèles restent donc limités à la segmentation d'objets possédant des frontières pouvant être caractérisées par de simples gradients, et sont incapables de traiter des images texturées. Les contours actifs basés région utilisent une information plus globale de l'image sous la forme de descripteurs de régions. Ceux-ci sont généralement intégrés à l'intérieur de la région formée par la courbe ou parfois sur l'image entière.

Il existe deux représentations d'un contour actif : paramétrique et implicite. Le principe de la représentation paramétrique est de rééchantillonner la courbe et de guider

son évolution à l'aide de points de contrôle. Ce mode de représentation, bien que rapide, ne permet pas au modèle de gérer naturellement les changements de topologie. La représentation implicite, introduite par Osher et Sethian dans [7] comme l'implémentation en ensemble de niveaux (*Level Sets*), permet de faire varier la topologie du modèle et fait évoluer la courbe en déterminant son équation d'évolution de la forme :

$$\frac{\partial C}{\partial t} = FN \quad (2)$$

N représente le vecteur normal à la courbe et F est sa fonction vitesse déterminée à l'aide de critères déduits de l'équation (1).

Dans le cadre de la segmentation d'images texturées, les modèles développés sont généralement des contours actifs basés région utilisant des descripteurs de texture comme descripteurs de région. Malheureusement, le pouvoir discriminant des descripteurs de texture peut varier considérablement d'un type d'image à un autre. Une manière de pallier cet inconvénient est d'intégrer le contour actif dans un processus de segmentation supervisé [5].

Nous proposons dans cet article un modèle de contour actif pour la segmentation d'images texturées basé sur les coefficients d'Haralick [8]. Afin de permettre à notre modèle de s'adapter automatiquement au type d'image traité, un classificateur binaire supervisé appris grâce à une segmentation manuelle est utilisé pour guider la courbe.

La section suivante présente le processus complet de seg-

mentation et détail de quelle manière le classificateur a été introduit dans le contour actif.

2 Processus de segmentation

Le processus de segmentation proposé est composé de deux étapes : une étape supervisée et une étape de segmentation. Dans un premier temps, l'utilisation de classificateurs supervisés rend nécessaire l'extraction d'un ensemble d'apprentissage. Une étape supervisée est donc utilisée afin de déterminer dans une image d'apprentissage deux régions C_{in} et C_{out} créées à partir d'une segmentation experte. C_{in} correspond à l'objet recherché, alors que C_{out} décrit la région située sur les bordures de l'objet.

Intuitivement, C_{in} devrait être définie comme la région à l'intérieur de C^* , et C_{out} comme le reste de l'image, tel que $\Omega^* = C_{in} \cup C_{out}$. Mais comme il a été énoncé dans [9], la région à l'extérieur de l'objet segmenté n'est que très rarement homogène et contient même parfois des textures très similaires à celle de l'objet recherché. De plus, même si ce phénomène n'apparaît pas dans l'image d'apprentissage Ω^* , il peut encore être présent dans les autres images de la série. Cela entrainera alors une grande variabilité de la texture de C_{out} sur toute la série d'images, rendant difficile le travail du classificateur.

Aussi, afin de maximiser la distance entre les textures de C_{in} et de C_{out} , garantissant un plus grand pouvoir discriminant aux caractéristiques d'Haralick, et de conserver des propriétés de textures constantes pour chacune des deux régions, C_{in} est définie comme la région intérieure à C^* et C_{out} comme une bande étroite fixée par l'expert autour de C_{in} (ainsi $C^* \in C_{out}$).

Les caractéristiques de textures d'Haralick de chaque pixel appartenant à ces deux régions sont alors déterminées à l'aide d'une fenêtre de voisinage local. Un ensemble d'apprentissage \mathbf{X} composé, pour chaque individu, de ses coefficients d'Haralick et de sa classe (objet ou fond de l'image) est ainsi extrait. La création de C_{in} , C_{out} et de l'ensemble d'apprentissage pour la série d'images échographiques est illustrée figure (1).

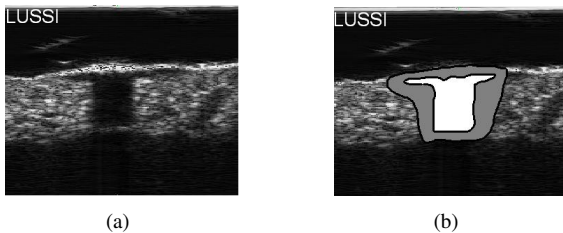


FIG. 1 – Image d'apprentissage pour une série d'images échographiques. a) Image d'origine b) Définition de C_{in} et C_{out} . Les coefficients d'Haralick des pixels définissant C_{in} (en blanc) et C_{out} (en gris) sont déterminés afin de composer l'ensemble d'apprentissage \mathbf{X} .

Un classificateur nécessite généralement le réglage de plusieurs paramètres, ces derniers influençant fortement le résultat

de la classification. Ainsi, il a été décidé de séparer \mathbf{X} en deux sous-ensembles \mathbf{X}_l et \mathbf{X}_e tel que $\mathbf{X} = \mathbf{X}_l \cup \mathbf{X}_e$. Chaque classificateur est appris selon différentes configurations de paramètres variant de manière régulière sur l'ensemble d'apprentissage \mathbf{X}_l , puis testé sur l'ensemble d'évaluation \mathbf{X}_e . La configuration de paramètres ayant obtenu le meilleur taux de classification sur \mathbf{X}_e permet alors de déterminer le classificateur introduit dans le modèle supervisé de contour actif. Afin de conserver un grand nombre d'individus pour l'apprentissage, la proportion d'individus dans chaque sous-ensemble est telle que $Card(\mathbf{X}_l) = \frac{2}{3}Card(\mathbf{X})$ et $Card(\mathbf{X}_e) = \frac{1}{3}Card(\mathbf{X})$ (l'ensemble \mathbf{X} possédant environ 10000 individus, une proportion de 1/3 est suffisante pour évaluer les classificateurs). En théorie, une fois les paramètres déterminés, il est recommandé de relancer un apprentissage sur l'ensemble complet \mathbf{X} , afin d'apprendre le maximum de données. En pratique, l'ensemble \mathbf{X} est suffisamment grand pour conserver le classificateur appris uniquement avec \mathbf{X}_l (correspondant à plus de 6600 individus).

Une fois l'apprentissage du classificateur réalisé, celui-ci est utilisé afin de guider le contour dans plusieurs images tests. Afin de réaliser l'intégration du classificateur directement dans l'équation d'évolution d'un contour actif, l'implémentation en ensemble de niveaux de Shi et Karl [10] est utilisée. Celle-ci part du principe qu'il est possible de n'utiliser que la valeur du signe de la fonction vitesse F de l'équation (2) pour guider un contour actif, chaque point de la courbe évoluant en fonction du signe de sa fonction F et de celle de ses voisins directs.

Ainsi, le signe de F détermine pour chaque point de la courbe son appartenance à l'objet recherché ou au fond de l'image. La fonction vitesse agit donc comme un classificateur binaire déterminant pour chaque individu (les points de la courbe), la classe à laquelle il appartient (objet ou fond). Dans notre approche, nous proposons de remplacer complètement la fonction vitesse par un classificateur binaire supervisé. Le processus complet de segmentation est décrit figure (2).

Trois types de classificateurs ont été testés : un algorithme des k-plus-proches-voisins (K-PPV), un réseau de neurones artificiels (RNA) et un modèle de machines à vecteurs supports (SVM). La section suivante compare leurs performances à celles du modèle de contour actif région de Chan et Vese [6]. Dans les sections suivantes nous noterons CA-CV le contour actif de Chan et Vese, CA-PPV le contour actif guidé par K-PPV, CA-SVM le contour actif guidé par machines à vecteurs supports et CA-RNA le contour actif guidé par réseau de neurones artificiels.

3 Résultats

Afin de mettre en évidence l'indépendance de nos modèles par rapport aux types d'images traités, nous présentons les résultats de deux problèmes de segmentation distincts : la segmentation de *naevus* dans des échographies et la segmentation de zones urbaines dans des images satellites (issues de la base

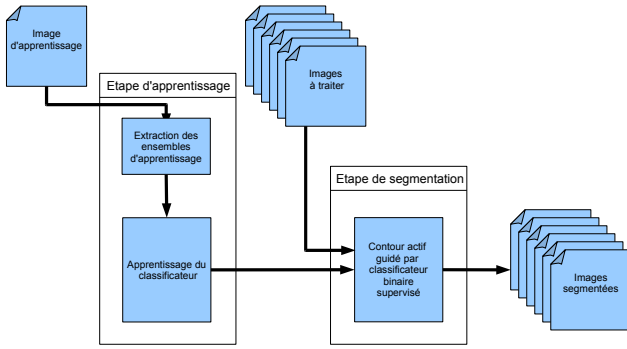


FIG. 2 – Processus complet de segmentation : dans un premier temps les ensembles X_l et X_e extraits de l’image d’apprentissage sont utilisés afin de réaliser l’algorithme d’apprentissage du classificateur. Dans un deuxième temps, celui-ci est utilisé pour guider le contour actif sur les autres images de la série.

ASTER). Pour chaque type d’image, une image d’apprentissage comprenant une segmentation experte est utilisée afin de réaliser l’étape d’apprentissage de chaque classificateur.

Pour évaluer les performances de nos modèles par rapport à une vérité terrain, deux critères sont utilisés : la mesure générique d’anomalie (*Generic Discrepancy Measure*) [11] notée d_{mga} , et l’opposé du critère de Pratt [12] noté PRA_m . Pour les deux mesures, une valeur de 0 correspondra à une segmentation parfaite. La plus mauvaise segmentation vaudra 1 pour l’opposé du critère de Pratt et environ 0,5 pour la mesure générique d’anomalie. Les tableaux (1) et (2) comparent les résultats de segmentation de nos trois modèles au contour actif de Chan et Vese sur les deux séries d’images. Les figures (3.a) à (3.d) présentent la segmentation obtenue par le contour actif guidé par un réseau de neurones sur deux images choisies dans chacune des deux problématiques de segmentation.

Ces résultats mettent en avant la capacité de segmentation des modèles proposés, en particulier pour le CA-SVM et le CA-RNA. Toutefois, la plus grande rapidité du CA-RNA tant pour l’apprentissage que pour la segmentation lui permettent d’être considéré comme le meilleur modèle. En conséquence, celui-ci a été choisi pour être adapté à la segmentation d’images échographiques 3D.

4 Segmentation d’échographies 3D

La reconstruction d’objet dans les échographies 3D représente un outil d’aide au diagnostic précieux. En effet, les dermatologues doivent régulièrement enlever de manière chirurgicale des *naevus* (“grains de beauté”). Afin d’assurer leur geste, une grande marge de sécurité est prise lors de l’ablation, entraînant parfois des conséquences disgracieuses (cicatrices). Afin de permettre au médecin de réaliser un diagnostic plus précis sur les dimensions du *naevus*, tout en évitant des techniques d’imagerie trop invasives, notre modèle de contour actif guidé par classificateur a été étendu à la reconstruction dans

des échographies 3D de la peau.

Le processus de segmentation reste inchangé. Une échographie 3D étant une succession de coupes 2D, l’image d’apprentissage est une unique coupe 2D mais les coefficients d’Haralick extraits sont calculés en trois dimensions. Une fois le classificateur appris, le contour actif (devenu une surface active) est alors lancé dans l’image complète afin de réaliser la reconstruction. Les reconstructions obtenues ont été validées par des médecins dermatologues et des experts en imagerie échographique. Les figures (4.a) à (4.d) présentent un exemple des reconstructions obtenues.

5 Conclusion et perspectives

Dans cet article, trois modèles de contours actifs guidés par un classificateur binaire supervisé ont été proposés. Les tests sur plusieurs types d’images ont montré que le modèle guidé par réseau de neurones est le plus efficace. Cela peut en partie s’expliquer par la bonne capacité de ce type de classificateur à sélectionner les caractéristiques. Le contour actif guidé par réseau de neurones a été adapté avec succès à la reconstruction dans des images échographiques 3D. Afin d’améliorer la précision de nos modèles, un contour actif multi-classificateur est en développement. De plus, la précision du classificateur étant dépendante des caractéristiques de textures utilisées pour créer l’ensemble d’apprentissage, une de nos perspectives est de développer un modèle prenant en compte plusieurs types de caractéristiques de textures (filtres de Gabor, ondelettes...).

TAB. 1 – Comparaison de la précision de segmentation des méthodes testées sur la série d’images satellites. Les mesures de précision sont multipliées par un facteur 100.

| Images satellites ASTER | | | |
|-------------------------|---------|--------------|--------------|
| Image | Méthode | PRA_m | d_{mga} |
| SAT 1 | CA-CV | 64,12 | 12,04 |
| | CA-PPV | 59,63 | 9,36 |
| | CA-SVM | 52,11 | 12,27 |
| | CA-RNA | 49,23 | 9,10 |
| SAT 2 Figure (3.a) | CA-CV | 92,02 | 47,57 |
| | CA-PPV | 76,67 | 13,13 |
| | CA-SVM | 76,90 | 12,79 |
| | CA-RNA | 77,06 | 12,70 |
| SAT 3 Figure (3.b) | CA-CV | 77,59 | 14,51 |
| | CA-PPV | 63,37 | 10,68 |
| | CA-SVM | 58,74 | 9,69 |
| | CA-RNA | 59,64 | 9,57 |
| SAT 4 | CA-CV | 82,73 | 15,59 |
| | CA-PPV | 74,40 | 12,25 |
| | CA-SVM | 70,15 | 11,93 |
| | CA-RNA | 69,98 | 10,21 |

TAB. 2 – Comparaison de la précision de segmentation des méthodes testées sur la série d’images échographiques.

| Images échographiques | | | |
|------------------------|---------|--------------|-------------|
| Image | Méthode | PRA_m | d_{mga} |
| Echo 1 | CA-CV | 66,91 | 1,17 |
| | CA-PPV | 69,85 | 1,20 |
| | CA-SVM | 48,51 | 1,05 |
| | CA-RNA | 50,91 | 0,53 |
| Echo 2 Figure (3.c) | CA-CV | 82,07 | 2,85 |
| | CA-PPV | 73,58 | 2,26 |
| | CA-SVM | 60,06 | 1,42 |
| | CA-RNA | 66,51 | 1,64 |
| Echo 3 Figure (3.d) | CA-CV | 66,95 | 2,31 |
| | CA-PPV | 58,05 | 1,85 |
| | CA-SVM | 44,20 | 1,10 |
| | CA-RNA | 43,86 | 1,10 |
| Echo 4 | CA-CV | 92,55 | 2,61 |
| | CA-PPV | 74,73 | 1,24 |
| | CA-SVM | 51,36 | 0,53 |
| | CA-RNA | 48,32 | 0,47 |

Références

[1] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : active contour models. *Int. J. of Comp. Vis.*, 1(4) :321–331, 1988.

[2] L. D. Cohen. On active contour models and balloons. *Computer Vision, Graphics, and Image Processing : Image Understanding*, 53(2) :211–218, 1991.

[3] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *Int. J. of Comp. Vis.*, 22(1) :61–79, 1997.

[4] S. Zhu and A. Yuille. Region competition : unifying snake/balloon, region growing, and Bayes/MDL/energy for multi-band image segmentation. *IEEE Trans. PAMI*, 18(9) :884–900, 1996.

[5] N. Paragios and R. Deriche. Geodesic active regions for supervised texture segmentation. In *proc. of the Int. Conf. on Computer Vision*, volume 2, pages 926–932, Washington, DC, USA, 1999. IEEE Computer Society.

[6] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. on Im. Proc.*, 10(2) :266–277, 2001.

[7] S. Osher and J. A. Sethian. Fronts propagation with curvature-dependent speed : algorithms based on Hamilton-Jacobi formulations. *J. of Comp. Ph.*, 79 :12–49, 1988.

[8] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. on Systems, Man, and Cybernetics*, 3 :610–621, 1973.

[9] J. Mille, R. Boné, P. Makris, and H. Cardot. 2D and 3D deformable models with narrow band region energy. In *proc. of the IEEE Int. Conf. on Image Processing*, pages 57–60, San Antonio, USA, 2007.

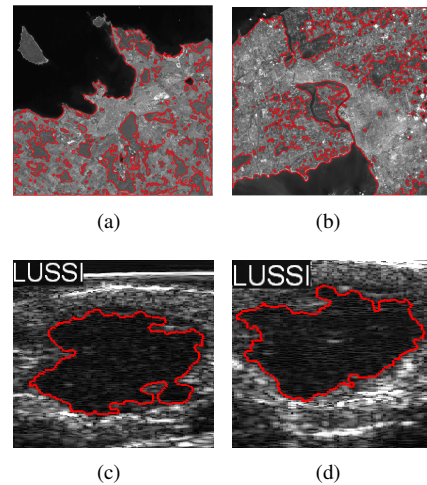


FIG. 3 – Exemples de segmentations avec le réseau de neurones : a-b) images satellites c-d) échographies

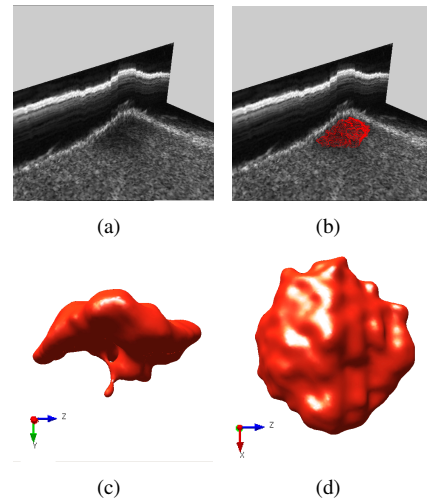


FIG. 4 – Exemple de reconstruction 3D. a) image échographique 3D b) reconstruction du naevus c-d) maillage obtenu

[10] Y. Shi and W. Karl. A fast level set method without solving PDEs. In *proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, volume 2, pages 97–100, Philadelphia, USA, 2005.

[11] J. S. Cardoso and L. Corte-Real. Toward a generic evaluation of image segmentation. *IEEE Trans. on Image Processing*, 14(11) :1773–1782, 2005.

[12] W. K. Pratt, O. D. Faugeras, and A. Gagalowicz. Visual discrimination of stochastic texture fields. *IEEE Trans. on Systems, Man and Cybernetics*, 8(11) :796–804, November 1978.