

Contours actifs basés régions pour la segmentation des objets en mouvement dans les séquences à caméra fixe ou mobile

Stéphanie JEHAN-BESSON¹, Michel BARLAUD¹, Gilles AUBERT²

¹Laboratoire I3S, CNRS-UNSA
2000 route des Lucioles, 06903 Sophia Antipolis, France

²Laboratoire J.A. Dieudonné, CNRS-UNSA
Parc Valrose, 06108 Nice Cedex 2, France
jehan@i3s.unice.fr, barlaud@i3s.unice.fr
gaubert@math.unice.fr

Résumé –

Nous adressons ici le problème de la segmentation des objets en mouvement dans les séquences d'images en utilisant des contours actifs basés régions. Nous proposons tout d'abord un cadre de travail général pour les contours actifs basés régions. A partir de la minimisation d'un critère général, dans lequel les régions sont caractérisées chacune par un descripteur, nous mettons en oeuvre une méthode Eulérienne qui nous permet de calculer l'équation d'évolution d'un contour actif. La vitesse calculée est celle qui permet de faire évoluer le contour actif le plus rapidement possible vers un minimum du critère. Cette nouvelle méthode Eulérienne nous permet de prendre facilement en compte les cas des descripteurs dépendant de caractéristiques attachées à la région et qui évoluent donc au cours de la propagation du contour.

Ce cadre de travail général est ensuite appliqué à la segmentation des objets en mouvement dans les séquences à caméra fixe ou mobile. Nous proposons d'intégrer le modèle de mouvement de la caméra dans le critère à minimiser et ainsi d'effectuer conjointement la compensation du mouvement de la caméra et la segmentation.

Abstract –

We address here the problem of moving objects segmentation in video sequences using region-based active contours. From the minimization of a general criterion, where the regions are each characterized by a descriptor, we propose a new Eulerian method to compute the evolution equation of the active contour. The velocity vector computed makes the active contour evolve as fast as possible towards a minimum of the criterion. This new Eulerian method allows us to take easily into account the case of descriptors depending upon features globally attached to the regions and so varying with the evolution of the curve.

This general framework is then applied to moving objects segmentation in video sequences acquired by a static or a mobile camera. We propose to jointly perform camera compensation and segmentation by taking directly into account the camera model in the criterion to minimize.

1 Introduction

La nouvelle norme de compression MPEG-4 [1] nécessite la segmentation préalable d'une séquence en objets vidéo. Dans ce but, nous proposons ici une méthode pour segmenter les objets en mouvement dans des séquences vidéo en utilisant des contours actifs basés régions.

De manière générale, on peut distinguer deux approches pour la segmentation : les approches basées régions comme les champs de Markov [2], et les approches basées contours. A l'origine, les contours actifs étaient des méthodes basées contours : les snakes [3], ou les contours actifs géodésiques [4] évoluent vers les gradients de l'image par la minimisation d'un critère. Ce critère utilise uniquement un terme sur le contour. De manière à utiliser les contours actifs pour segmenter les objets en mouvement, des informations basées régions, comme par exemple le champ des vecteurs mouvement, doivent être incorporées dans l'équation d'évolution du contour actif.

Notre première contribution est de définir un cadre de

travail général très complet pour les contours actifs basés régions. Nous introduisons tout d'abord un critère incluant à la fois des termes basés contours et des termes basés régions. A partir de ce critère initial, nous mettons en place un cadre théorique nous permettant de calculer l'équation d'évolution du contour actif. Nous utilisons une nouvelle démonstration Eulérienne qui assure la décroissance la plus rapide du contour actif vers un minimum du critère. Dans le critère, chaque région est décrite par une fonction que nous appelons génériquement "descripteur" de la région. Pour une application particulière, l'utilisateur doit seulement choisir des descripteurs bien adaptés.

La deuxième contribution de cet article est de proposer un ensemble de "descripteurs" pour la détection des objets en mouvement avec compensation du mouvement de la caméra. Le modèle de mouvement de la caméra est directement intégré dans le critère à minimiser. Il est estimé d'une image à l'autre par estimation robuste, en utilisant le théorème semi-quadratique [5].

La cadre de travail sur les contours actifs basés régions est développé section 2, tandis que la segmentation des objets vidéo est exposée dans la section 3. En section 4, des résultats expérimentaux sur des séquences réelles montrent le potentiel de cette approche.

2 Un cadre de travail général pour les contours actifs basés régions

Notre but est de trouver la partition de l'image qui minimise un critère incluant des termes basés régions et des termes basés contours. Chaque région est décrite par une fonction nommée "descripteur" de la région.

Dans la littérature, des auteurs [6, 7, 8, 9] ont proposé d'ajouter des termes basés régions pour faire évoluer un contour actif. Ces travaux montrent le potentiel des approches basées régions, voir [10] pour une revue détaillée de ces méthodes. Notre contribution est ici de mettre en oeuvre un cadre de travail général pour les contours actifs basés régions et de proposer une nouvelle méthode pour calculer l'équation d'évolution du contour actif sans utiliser les équations d'Euler Lagrange. Nous proposons d'introduire le schéma dynamique directement dans le critère à minimiser et d'utiliser un théorème issu de la mécanique des fluides pour calculer la dérivée du critère. Avec cette méthode, le cas des descripteurs dépendant de l'évolution de la courbe (qui dépendent de caractéristiques attachées aux régions en évolution) peut être aisément pris en compte. Nous montrons que cette dépendance induit des termes additionnels dans l'équation d'évolution du contour actif [10].

2.1 Du critère vers l'équation d'évolution d'un contour actif

Une image $I(x, y)$ est une fonction définie pour $(x, y) \in \Omega$, où Ω est le domaine de l'image. Ce domaine est constitué de deux parties : la partie contenant les objets à segmenter notée Ω_{in} , et le fond de l'image noté Ω_{out} . La frontière commune entre ces deux domaines est notée Γ . Nous cherchons alors les domaines qui minimisent le critère suivant :

$$J(\Omega_{out}, \Omega_{in}, \Gamma) = \iint_{\Omega_{out}} k^{(out)} + \iint_{\Omega_{in}} k^{(in)} + \int_{\Gamma} k^{(b)} \quad (1)$$

Comme on peut le voir dans l'équation (1), les deux premiers termes sont des intégrales sur les régions alors que le troisième terme est une intégrale de contour. Les fonctions $k^{(out)}(\cdot)$, $k^{(in)}(\cdot)$ et $k^{(b)}(\cdot)$ sont appelées respectivement descripteur de Ω_{out} , Ω_{in} et Γ .

Afin de calculer une solution optimale, le critère est plongé dans un schéma dynamique dans lequel les régions deviennent des fonctions d'un paramètre d'évolution τ .

$$J(\tau) = \iint_{\Omega_{out}(\tau)} k^{(out)} + \iint_{\Omega_{in}(\tau)} k^{(in)} + \int_{\Gamma(\tau)} k^{(b)} \quad (2)$$

Le contour $\Gamma(\tau)$ est alors modélisé comme un contour actif qui converge vers la segmentation voulue. Soit Γ_0

le contour initial défini par l'utilisateur, nous cherchons l'équation d'évolution de $\Gamma(\tau)$ sous la forme suivante :

$$\frac{\partial \Gamma(\tau)}{\partial \tau} = \mathbf{v} \quad (3)$$

L'inconnu du problème est \mathbf{v} , le vecteur vitesse du contour, qui doit faire évoluer le plus rapidement possible $\Gamma(\tau)$ vers un minimum du critère.

Le détail de la dérivation du critère $J(\tau)$ est donné dans [10]. On trouve :

$$J'(\tau) = \iint_{\Omega_{in}(\tau)} \frac{\partial k^{(in)}}{\partial \tau} dx dy + \iint_{\Omega_{out}(\tau)} \frac{\partial k^{(out)}}{\partial \tau} dx dy \quad (4)$$

$$+ \int_{\Gamma(\tau)} (k^{(out)} - k^{(in)} - k^{(b)} \cdot \kappa + \nabla k^{(b)} \cdot \mathbf{N})(\mathbf{v} \cdot \mathbf{N}) ds$$

avec $\kappa(x, y, \tau)$ la courbure de $\Gamma(x, y, \tau)$. Dans le cas des descripteurs ne dépendant pas de τ , les deux dernières intégrales sont nulles. Nous exprimons ici la dérivée dans le cas général lorsque les descripteurs peuvent dépendre de τ .

Nous prenons $k^{(in)}$ et $k^{(out)}$ dans le cas général comme des fonctions dépendant de caractéristiques attachées à la région ou au contour. On a :

$$k^{(in)}(x, y, \tau) = g^{(in)}(x, y, G_1^{(in)}(\tau), \dots, G_n^{(in)}(\tau)) \quad (5)$$

Avec :

$$G_j^{(in)}(\tau) = \begin{cases} \iint_{\Omega_{in}(\tau)} \psi_j^{(in)}(x, y) dx dy & \text{if } j \in I_{in} \\ \int_{\Omega_{in}(\tau)} \psi_j^{(in)}(x, y) ds & \text{if } j \in J_{in} \end{cases} \quad (6)$$

avec I_{in} les indices j où $G_j^{(in)}$ est une intégrale de surface et J_{in} les indices j où $G_j^{(in)}$ est une intégrale de contour. On obtient alors :

$$\iint_{\Omega_{in}} \frac{\partial k^{(in)}}{\partial \tau} = - \sum_{j \in I_{in}} A_j^{(in)} \int_{\Gamma(\tau)} \psi_j^{(in)}(\mathbf{v} \cdot \mathbf{N}) ds \quad (7)$$

$$- \sum_{j \in J_{in}} A_j^{(in)} \int_{\Gamma(\tau)} (\psi_j^{(in)} \kappa - \nabla \psi_j^{(in)} \cdot \mathbf{N})(\mathbf{v} \cdot \mathbf{N}) ds$$

avec :

$$A_j^{(in)} = \iint_{\Omega_{in}(\tau)} \frac{\partial g^{(in)}}{\partial G_j^{(in)}}(x, y, G_1^{(in)}(\tau), \dots, G_n^{(in)}(\tau)) dx dy \quad (8)$$

Nous exprimons de la même manière l'intégrale sur $k^{(out)}$ et nous obtenons ainsi l'expression générale de la dérivée de $J(\tau)$. De cette expression, nous déduisons l'expression du vecteur vitesse \mathbf{v} en utilisant le théorème de Cauchy-Schwartz. Le contour va évoluer le plus rapidement possible vers un minimum du critère avec l'équation d'évolution suivante :

$$\frac{\partial \Gamma(\tau)}{\partial \tau} = [k^{(in)} - k^{(out)} + k^{(b)} \cdot \kappa - \nabla k^{(b)} \cdot \mathbf{N} \quad (9)$$

$$+ \sum_{j \in I_{in}} A_j^{(in)} \psi_j^{(in)} - \sum_{j \in I_{out}} A_j^{(out)} \psi_j^{(out)}$$

$$+ \sum_{j \in J_{in}} A_j^{(in)} (\psi_j^{(in)} \kappa - \nabla \psi_j^{(in)} \cdot \mathbf{N})$$

$$- \sum_{j \in J_{out}} A_j^{(out)} (\psi_j^{(out)} \kappa - \nabla \psi_j^{(out)} \cdot \mathbf{N})] \mathbf{N}$$

Nous montrons différents exemples de descripteurs dépendant de τ et nous calculons la valeur des termes additionnels issus de cette variation dans [10].

De manière à implémenter l'EDP (9), nous utilisons la méthode des ensembles de niveaux introduite par Osher et Sethian [11], voir [10] pour plus de détails.

3 Segmentation des objets vidéo

Le cadre de travail général introduit en section 2 est maintenant utilisé pour la segmentation des objets en mouvement. Notons I_n l'intensité de l'image numéro n de la séquence d'images. Les notations de la partie 2 sont maintenant indicées par n , où n représente le numéro de l'image dans la séquence. Les fonctions $k^{(n,out)}$, $k^{(n,in)}$ et $k^{(n,b)}$ représentent respectivement les descripteurs de la région du fond notée $\Omega_{n,out}$, de la région contenant les objets en mouvement notée $\Omega_{n,in}$ et du contour noté Γ_n pour l'image numéro n . Nous proposons deux types de descripteurs.

3.1 Premier ensemble de descripteurs

Nous proposons un premier ensemble de descripteurs avec estimation du fond de la séquence.

$$\begin{cases} k^{(n,out)}(x, y) &= (B(x, y) - I_n(x, y))^2 \\ k^{(n,in)}(x, y) &= \alpha_c \\ k^{(n,b)}(x, y) &= \lambda \end{cases} \quad (10)$$

Avec B l'image du fond de la séquence calculée par estimation robuste sur une fenêtre de m images contenant l'image à segmenter. Les paramètres λ et α_c sont deux constantes positives. Si la séquence est à caméra mobile, le mouvement de la caméra doit être compensé au préalable.

En minimisant $k^{(n,out)}$ sur la région du fond, nous voulons que $I_n(x, y)$ soit proche de $B(x, y)$ lorsque le pixel (x, y) appartient à la région du fond. Nous minimisons aussi la surface de l'objet plus ou moins suivant la valeur de la constante de pondération α_c . Ainsi le paramètre α_c va jouer le rôle d'un seuil de détection. Le paramètre λ est quant à lui un paramètre de lissage.

3.2 Second ensemble de descripteurs

Nous proposons également les descripteurs suivants avec compensation du mouvement de la caméra:

$$\begin{cases} k^{(n,out)}(p) &= \sum_{p_i \in \mathcal{V}} (I_n(p_i) - Proj(I_{n-1})(p_i))^2 \\ k^{(n,in)}(p) &= \alpha_c \\ k^{(n,b)}(p) &= \lambda \end{cases} \quad (11)$$

Avec $p = (x, y)$ et $Proj(I_{n-1})$ une image qui est la projection de l'image I_{n-1} dans le référentiel de l'image I_n . Le terme \mathcal{V} désigne un voisinage de p .

3.2.1 Le modèle de mouvement de la caméra

Le mouvement de la caméra peut être modélisé par un modèle de mouvement affine à six paramètres qui est un bon compromis entre complexité et représentativité [2]. Ainsi le mouvement apparent d'un point (x, y) du fond de

l'image, entre les images I_{n-1} et I_n , peut être modélisé par:

$$\mathbf{w}_n(p) = \mathbf{A}_n p + \mathbf{t}_n = \begin{bmatrix} a_{11}^n x + a_{12}^n y + t_1^n \\ a_{21}^n x + a_{22}^n y + t_2^n \end{bmatrix} \quad (12)$$

avec $p = (x, y)$. L'image projection $Proj(I_{n-1})$ est donnée par:

$$Proj(I_{n-1})(p + \mathbf{w}_n(p)) = I_{n-1}(p) \quad (13)$$

Des méthodes d'interpolation sont utilisées pour remplir l'image projection.

3.2.2 Estimation du modèle de la caméra

Nous cherchons les six paramètres du modèle de mouvement de la caméra qui minimisent le critère suivant:

$$\sum_{p \in \Omega} \varphi(|\mathbf{u}_n - \mathbf{A}_n p - \mathbf{t}_n|) \quad (14)$$

Dans le cadre Bayésien, φ est connue comme une fonction potentielle. Cette fonction est introduite pour éliminer les points aberrants dus au mouvement des objets vidéo. Nous choisissons l'estimateur de Geman et Mc Lure [12]:

$$\varphi(t) = \frac{t^2}{\sigma + t^2} \quad (15)$$

avec σ une constante positive.

Le champ des vecteurs mouvement \mathbf{u}_n est classiquement calculé en utilisant un algorithme de Block Matching entre les images I_{n-1} et I_n .

De manière à minimiser le critère (14), nous utilisons les propriétés de la régularisation semi-quadratique avec la stratégie des minimisations alternées [5]:

Initialisation $(\mathbf{A}_n^0, \mathbf{t}_n^0)$

Répéter

$$\begin{cases} b^k = \frac{\varphi'(r^k)}{2r^k} \\ (\mathbf{A}_n^{k+1}, \mathbf{t}_n^{k+1}) = \operatorname{argmin}_{(\mathbf{A}, \mathbf{t})} \sum_{\Omega} b^k(r^{k+1})^2 \end{cases}$$

Jusqu'à la convergence.

Nous notons $r^k = |\mathbf{u}_n - \mathbf{A}_n^k p - \mathbf{t}_n^k|$, et k le numéro de l'itération. La minimisation de $\sum_{\Omega} b^k(r^{k+1})^2$ est effectuée en utilisant une méthode de descente de gradient.

Nous obtenons ainsi $(\mathbf{A}_n, \mathbf{t}_n)$, et nous pouvons calculer le descripteur $k^{(n,out)}$.

3.3 Équation d'évolution du contour actif

Pour le premier et le second ensemble de descripteurs, l'équation d'évolution du contour actif se réduit à l'expression suivante (ces descripteurs ne dépendant pas de τ):

$$\frac{\partial \Gamma_n(\tau)}{\partial \tau} = [k^{(n,in)} - k^{(n,out)} + k^{(n,b)} \cdot \kappa - \nabla k^{(n,b)} \cdot \mathbf{N}] \mathbf{N} \quad (16)$$

Le contour initial Γ_n^0 est choisi comme un rectangle près des bords de l'image. Ensuite, nous utilisons l'EDP (16), avec le premier ou le second ensemble de descripteurs, pour faire évoluer le contour actif vers les objets en mouvement à segmenter. La segmentation finale de l'image numéro n peut être utilisée pour calculer une courbe initiale pour l'image numéro $n + 1$.

4 Résultats expérimentaux

L'algorithme a été testé sur des séquences réelles, "Akiyo" une séquence à caméra fixe, et deux séquences à caméra mobile, "Coastguard" du groupe de recherche COST 211 et "Edberg" du groupe MPEG. Nous montrons l'évolution du contour figure 1 pour la séquence "Akiyo". Nous prenons pour cette séquence le premier ensemble de descripteurs (10) avec $\alpha_c = 20$ et $\lambda = 10$. Pour les deux séquences à caméra mobile, nous choisissons le second ensemble de descripteurs (11) avec $\sigma = 0.01$ et $\lambda = 10$. Pour "Coastguard", nous prenons $\alpha_c = 50$ et pour "Edberg", $\alpha_c = 25$. Dans les figures 2 et 3, le contour final est en blanc. Les objets en mouvement sont bien détectés ce qui montre le potentiel des contours actifs pour la segmentation des objets vidéo.

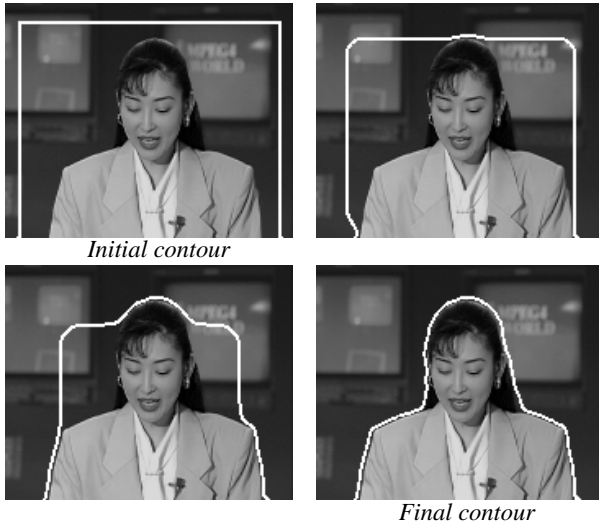


FIG. 1: "Akiyo": propagation du contour actif sur une image

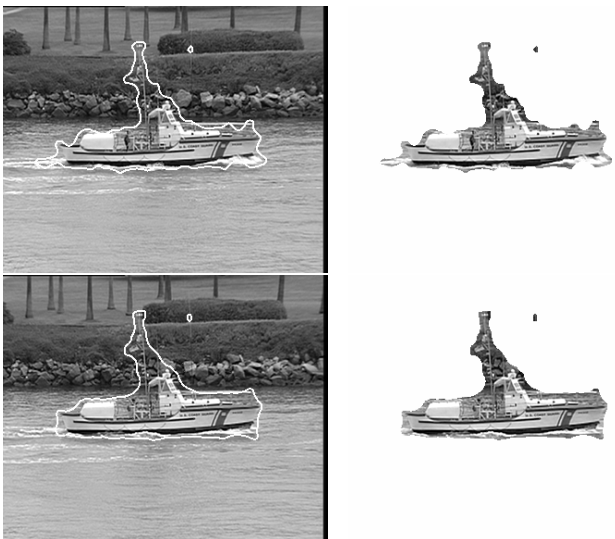


FIG. 2: "Coastguard", **Gauche**: Contours finaux **Droite**: objets vidéo.



FIG. 3: "Edberg": Contours finaux en blanc

5 Références

- [1] International Organization for Standardization, "Overview of the MPEG-4 standard," March 1999, ISO/IEC JTC1/SC29/WG11 N2725.
- [2] J. Odobez and P. Bouthemy, "Detection of multiple moving objects using multiscale MRF with camera motion compensation," in *ICIP*, Texas, 1994.
- [3] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *IJCV*, vol. 1, pp. 321–332, 1988.
- [4] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *IJCV*, vol. 22, no. 1, pp. 61–79, 1997.
- [5] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Processing*, vol. 6, no. 2, pp. 298–311, february 1997.
- [6] S. Zhu and A. Yuille, "Region competition: unifying snakes, region growing, and bayes/MDL for multi-band image segmentation," *IEEE PAMI*, vol. 18, pp. 884–900, september 1996.
- [7] A. Chakraborty, L. Staib, and J. Duncan, "Deformable boundary finding in medical images by integrating gradient and region information," *IEEE Trans. Medical Imag.*, vol. 15, pp. 859–870, 1996.
- [8] A.R. Mansouri and J. Konrad, "Motion segmentation with level sets," in *ICIP*, Kobe, Japan, 1999.
- [9] N. Paragios and R. Deriche, "Geodesic active regions for motion estimation and tracking," in *ICCV*, Corfu Greece, 1999.
- [10] S. Jehan-Besson, M. Barlaud, and G. Aubert, "Video object segmentation using eulerian region-based active contours," in *ICCV*, Vancouver, Canada, 2001.
- [11] J.A. Sethian, *Level Set Methods*, Cambridge Univ. Press, 1996.
- [12] S. Geman and D.E. McClure, "Bayesian image analysis: an application to single photon emission tomography," in *Proc. Statist. Comput. Sect.*, 1985.