ACTIVE CONTOUR ATTRACTED BY A REFERENCE CONTOUR: A REGION-BASED APPROACH

Cédric De Roover, Jacek Czyz and Benoit Macq

Laboratoire de Télécommunications et Télédétection, Université catholique de Louvain
Bâtiment Stévin, place du Levant - 2
B-1348 Louvain-la-Neuve, Belgium
email: deroover,czyz,macq@tele.ucl.ac.be

ABSTRACT

Lack of temporal coherence in video segmentation algorithms often leads to flickering or to discrepancies on the segmentation mask boundaries. Active contour video segmentation algorithms can lead to very smooth segmentation masks when they are defined in three dimensions (two spatial and one temporal).

In this paper, we will use an active contour method to attract an active surface toward a non-smoothed segmentation mask boundary. This active surface will produce a new segmentation mask which will be smoother than the first one. To achieve this task, we introduce a new region-based term for active contour segmentation in the variational framework. This term attracts the evolving curve to a reference contour. The energy criterion and the evolution equation are defined in n-dimension and we investigate the particular case of two regions.

The use of active surfaces to smooth video segmentation masks appears to be a very powerful tool for video postprocessing. The resulting masks are smoother than the original ones, the discrepancies of the segmentation masks are removed and the flickering on the boundary of the segmentation masks are considerably reduced.

1. INTRODUCTION

Active contours are powerful tools for image and video segmentation or tracking. They can be formulated in the framework of variational methods. The basic principle is to derive a Partial Differential Equation (PDE) from the minimization of an energy criterion. By solving the PDE, the contour is iteratively deformed until it converges towards a (local) minimum of the criterion hopefully corresponding to boundaries of the objects to be segmented. The curve, noted Γ, evolves in its normal direction, under a velocity field deduced from the minimization of an energy criterion. Originally, snakes [1] or geodesic active contours [2] are driven towards the boundaries of the objects through the minimization of a boundary integral:

$$J(\Gamma) = \int_{\Gamma} k^{\Gamma}(s) ds, \tag{1}$$

where $k^{\Gamma}(s)$ is generally a function of the gradient of the image. Active contours driven by the minimization of region functionals in addition to boundary functionals have been developed in [3, 4, 5]. In region-based active contours, the domain is divided in different regions. Let us examine the case of two regions Ω^{in} and Ω^{out}. The energy J of such an active contour can be written as:

$$J(\Omega^{in}, \Omega^{out}) = \int_{\Omega^{in}} k^{in}(\mathbf{x})d\mathbf{x} + \int_{\Omega^{out}} k^{out}(\mathbf{x})d\mathbf{x} + \int_{\Gamma} k^{\Gamma}(\mathbf{x})d\mathbf{x}. \tag{2}$$

The functionals k^{in} and k^{out}, called descriptors, characterize the regions. The descriptors may depend on the region features, $k^{i}(\Omega^{i}, \mathbf{x})$, e.g. by considering different region statistics like the mean intensity or the variance [6].

In this paper, we study the attraction of an active contour to a reference contour. This is equivalent to consider the reference contour as being a shape prior for the active contour. The use of shape prior in active contour has been studied by several researchers. First approaches were statistical models of shape variation [7, 8, 9, 10]. More recently, variational model have been proposed [11, 12, 13, 14].
Gastaud et al. [13] proposed an interesting variational approach based on an active contour technique including a shape prior. The criterion is based on the distance between the active contour and the reference contour. The derivation of the contour evolution criterion introduced in [13] assumes that the evolving curve does not cross the skeleton of the reference contour. This assumption is not often satisfied when dealing with multiple regions to segment or when the reference curve is not smooth enough. With a boundary-based approach, when a point of the evolution curve is on the skeleton of the reference contour, it has theoretically two or more than two possibilities of evolution directions because there exists more than one closest reference point. Nevertheless, in practice, the implementation and the neighboring points drive the evolution in the good direction [15]. With a region-based criterion, the assumption of not crossing the skeleton could be discarded.

For these reasons, we propose a *region-based* criterion for active contour segmentation. Our region-based term is defined in n-dimension for a domain separated in K different regions. Foulonneau et al. [14] also presented a region-based criterion to attract an active contour to a shape prior. The difference with our method is that their criterion is defined in 2D as a distance between shape descriptors based on the Legendre moments of the characteristic function. They can deal with affine deformation of the shape prior but extension of their approach to 3D and K regions would be difficult.

We apply the proposed method to video segmentation mask smoothing. A video segmentation mask, obtained by a fast watershed algorithm [16], is used as a tri-dimensional shape prior (or reference surface). An active surface is then attracted to this reference using

Thanks to Belgian Walloon Region for funding.

the proposed active contour criterion, resulting in a both temporally and spatially smoothed segmentation mask.
The general contour evolution criterion to be minimized in n-D with K regions is introduced in Section 2. This criterion is then minimized and an evolution equation is introduced for the particular case of two regions in n-D in Section 3. Finally, we show in Section 4 that our new criterion is well suited for video segmentation mask smoothing applications and quantitative evaluation is performed in Section 5.

2. A GENERAL FRAMEWORK

Consider a reference partition $\Omega_1^{ref}, \ldots, \Omega_K^{ref}$ of a domain $\mathcal{X} \subset \mathbb{R}^n$. Let Γ^{ref} be the reference boundary delimiting the K different regions $\Omega_1^{ref}, \ldots, \Omega_K^{ref}$. Note that :

$$\bigcup_{i=1}^{K} \partial\Omega_i^{ref} = \Gamma^{ref}, \tag{3}$$

where $\partial\Omega_i^{ref}$ is the boundary of the region Ω_i^{ref}. For each region Ω_i^{ref}, we define a distance map which indicates the distance of each point of coordinates $\mathbf{x}$ of the domain $\mathcal{X} \subset \mathbb{R}^n$ to the boundary region $\partial\Omega_i^{ref}$. Let $\varphi(d_i(\mathbf{x}, \partial\Omega_i^{ref}))$ be the distance map of region Ω_i^{ref}. $\varphi : \mathbb{R} \rightarrow \mathbb{R}$, is a differentiable, odd, and increasing function of the geometric signed distance d_i. In our applications we will use $\varphi(d) = d$. The function $d_i(\mathbf{x}, \partial\Omega_i^{ref})$ is the distance between $\mathbf{x}$ and the reference boundary $\partial\Omega_i^{ref}$. We use the following convention:

$$d_i(\mathbf{x}, \partial\Omega_i^{ref}) = \begin{cases} min_{\mathbf{y}\in\partial\Omega_i^{ref}} |\mathbf{x} - \mathbf{y}| & \text{if } \mathbf{x} \text{ is outside } \Omega_i^{ref} \\ -min_{\mathbf{y}\in\partial\Omega_i^{ref}} |\mathbf{x} - \mathbf{y}| & \text{if } \mathbf{x} \text{ is inside } \Omega_i^{ref} \\ 0 & \text{if } \mathbf{x} \in \partial\Omega_i^{ref} \end{cases}$$

This definition implies:

$$\forall \mathbf{x} \in \Omega_i^{ref} : \varphi(d_i(\mathbf{x}, \partial\Omega_i^{ref})) = \min_{1\leq j\leq K} \varphi(d_j(\mathbf{x}, \partial\Omega_j^{ref})). \tag{4}$$

From those definitions, it is possible to prove that:

Theorem 1 *A partition $\Omega_1, \ldots, \Omega_K$ of the domain $\mathcal{X} \subset \mathbb{R}^n$ minimizes the following energy criterion:*

$$J_K(\Omega_1, \ldots, \Omega_K) = \sum_{i=1}^{K} \int_{\Omega_i} \varphi(d_i(\mathbf{x}, \partial\Omega_i^{ref}))d\mathbf{x}, \tag{5}$$

if and only if $\Omega_i = \Omega_i^{ref}$, $\forall i \in [1, \ldots, K]$.

The proof of the theorem is given in Appendix A. The theorem implies that in order to find $\partial\Omega_i$, one has to minimize J_K. In the next section, we introduce how this criterion J_K can be minimized in the special case of two regions.

3. A TWO REGION CRITERION

Assume that we have two regions, Ω_{in} and Ω_{out}, in the domain $\mathcal{X} \subset \mathbb{R}^n$. Those regions have a common boundary, Γ (see Fig. 1) . Note that a region can be a set of different unconnected objects. Let also assume that we have a reference boundary, Γ^{ref}, making the separation between two reference regions Ω_{in}^{ref} and Ω_{out}^{ref}. We compute the two distance maps $d_{in}(\mathbf{x}, \Gamma^{ref})$ and $d_{out}(\mathbf{x}, \Gamma^{ref})$ according to Equation (4) (see Fig. 2 in the one-dimensional case).

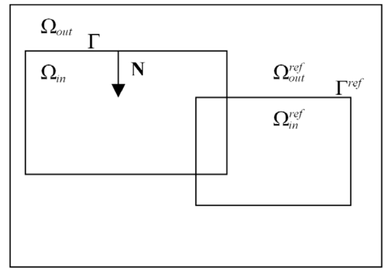


Fig. 1. The initial partition of the image in two regions, Ω_{in}^{ref} and Ω_{out}^{ref}, and the evolving curve Γ delimiting the current regions Ω_{in} and Ω_{out}.

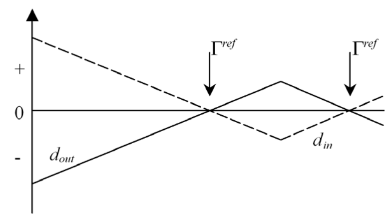


Fig. 2. Distance functions d_{in} and d_{out} associated to the reference regions Ω_{in}^{ref} and Ω_{out}^{ref} in the one dimension case. The distances are negative within their corresponding region and positive outside.

In the particular case of two regions, $\Omega_{in}^{ref} = \mathcal{X}\backslash\Omega_{out}^{ref}$ and, consequently, $d_{in}(\mathbf{x}, \Gamma^{ref}) = -d_{out}(\mathbf{x}, \Gamma^{ref})$.
Using (5), the proposed region-based criterion, J_s, that attracts the boundary Γ to the shape prior Γ^{ref}, becomes:

$$\begin{aligned} J_s(\Omega_{in}, \Omega_{out}, \Gamma) &= \int_{\Omega_{in}} \varphi(d_{in}(\mathbf{x}, \Gamma^{ref}))d\mathbf{x} \\ &+ \int_{\Omega_{out}} \varphi(d_{out}(\mathbf{x}, \Gamma^{ref}))d\mathbf{x} \\ &+ \int_{\Gamma} \lambda d\vec{s}, \end{aligned} \tag{6}$$

where $\mathbf{x} \in \mathbb{R}^n$ and λ is a smoothing constant parameter. The addition of the smoothing term will allow to smooth to obtain a smooth final contour. We perform the derivation of this criterion following an Eulerian framework as in [6].
Equation (6) requires the computation of the two distance maps d_{in} and d_{out}.
However, since $d_{in}(\mathbf{x}, \Gamma^{ref}) = -d_{out}(\mathbf{x}, \Gamma^{ref})$, and using the definition of φ, the resulting evolution equation can be writen as:

$$\frac{\partial\Gamma_s(\tau)}{\partial\tau} = \left[2\varphi(d_{in}(\mathbf{x}, \Gamma^{ref})) + \lambda\kappa_m\right]\mathbf{N}, \tag{7}$$

where κ_m expresses the mean curvature of the surface.

4. APPLICATION: SMOOTHING

4.1. Implementation

Our implementation of Equation (7) uses a level set framework [17]. The reference contour is obtained by performing a watershed segmentation using the algorithm introduced in [18]. These segmentation masks returned by the watershed algorithm appear to be not

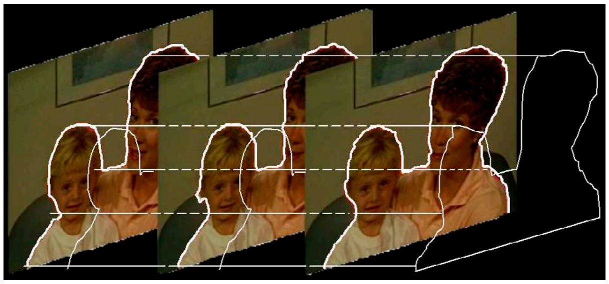


Fig. 3. 3D-smoothing: an active surface is attracted by the segmentation mask obtained by a first segmentation algorithm.

(a) Reference contour on frame 1 (b) Final contour on frame 1

(c) Reference contour on frame 2 (d) Final contour on frame 2

Fig. 4. (Left) Reference contour, (Right) 3D-smoothed contour

very smooth, both spatially and temporally (see Fig. 3). From those binary masks, we built a 3D distance map using the fast algorithm introduced in [19]. This map defines the distance of each point of the video sequence to the closer point of the reference contour which is not necessary in the same frame. This 3D-distance map is given as input to the level set algorithm. To speed up the convergence of the algorithm, we initialize the level set function on the reference surface ($\Gamma(0) = \Gamma^{ref}$). Then, the evolving surface converges in less than 100 iterations to a new surface which is smoother than the original but keeping the same shape.

4.2. Results

Fig. 4 shows the result on the first frames of the sequence *Erik*. We observe that the final contour is smoother than the reference contour obtained by the first segmentation algorithm. Thanks to the use of a 3D active surface, the segmentation mask is smoothed in the spatial axis but also in the temporal axis.

Fig. 5 also shows that the final contour is smoother than the reference contour. By looking at the reference contours, one can see that the contour are not smooth and not regular along the temporal axis. For instance, we notice two large discrepancies. The first one is on frame 2, close to the shoulder and the head of the daughter.

(a) frame 2 (b) frame 3 (c) frame 2 (d) frame 3

Fig. 5. Zoom on 3D-smoothing. (a) (b) are the initial contour, (c) (d) are the final contour. The temporal discrepancies are removed by the rigidity of the 3D-active surface.

The second one is on frame 3, close to the ear of the daughter. As those discrepancies appear only on one frame, the 3D-level set removes them thanks to the temporal rigidity. The final contour is then smoother in a spatial point of view but also in a temporal point of view.

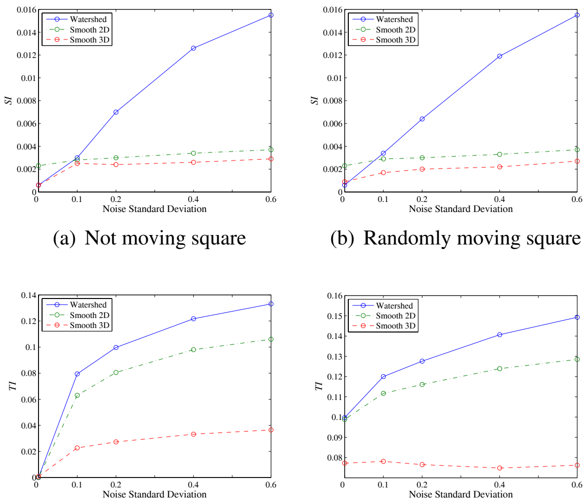
5. VALIDATION

In this section we quantitatively evaluate the improvement of 3D smoothing of video segmentation masks. For doing this we use the spatial perceptual information (SI) to compare a smooth segmentation mask to a reference segmentation mask on a spatial point of view, and the temporal perceptual information (TI) to have a comparison on a temporal point of view. Both metrics were recommended in a ITU-T Recommendation and may be found in [20]. For our experiments, we created two synthetic video sequences of 100 frames. The first one is a non moving black square over a white background and the other one is a square randomly moving over the sequence. The sequences are corrupted with zero-mean gaussian noises, non correlated over time, with standard deviation of 0.1, 0.2, 0.4 and 0.6. We then apply a watershed segmentation on each frame of the sequences. Then, we smoothed the watershed masks with 2D level-sets and with 3D level-sets. Fig. 6 plots average SI and TI values versus noise standard deviation. We can see that when noise increases, segmentation performance decreases. From a spatial point of view (SI), the 3D smoothing does not bring much more performances than 2D smoothing. But on a temporal point of view (TI) the improvement brought by the 3D smoothing is obvious. The flickering is reduced substantially. Note that in case of no noise, the 2D smoothing smooth the angles of the square. The segmentation results are then worse with a smoothing than with a simple watershed segmentation (Fig. 6 (a),(b)). However, we do not observe this effect with 3D smoothing. This is because, with 3D smoothing, a point of the evolving surface is influenced by the points in the other frames. Therefore, the curvature should be higher than with a 2D smoothing in order to smooth the angles of the square.

6. CONCLUSION

This paper has proposed a new region-based term for active contour segmentation in the variational framework. This term attracts the evolving curve to a reference contour. An application of such a criterion illustrated here is the smoothing of segmentation masks.

On the basis of a reference contour, we first compute a signed distance map per region. In the case of two regions, one signed distance map is sufficient. The active contour evolves through this distance



(a) Not moving square (b) Randomly moving square

(c) Not moving square (d) Randomly moving square

Fig. 6. Spatial perceptual information (SI) and temporal perceptual information (TI).

map, being attracted by the reference contour and constrained by the smoothness term. The main advantage of this smoothing is the spatial and the temporal coherence. This prevents time artifacts and considerably reduces the flickering. Moreover, thanks to the 3D, a point of the boundary is influenced by the previous and the following frames.

A. PROOF OF THEOREM 1

For any partition $\Omega_1, \ldots, \Omega_K$ of the domain $\mathcal{X} \subset \mathbb{R}^n$ we have :

$$\sum_{i=1}^{K} \int_{\Omega_i} \varphi(d_i(\mathbf{x}, \partial\Omega_i^{ref}))d\mathbf{x} \geq \sum_{i=1}^{K} \int_{\Omega_i} \min_{1 \leq j \leq K} \left(\varphi(d_j(\mathbf{x}, \partial\Omega_j^{ref}))\right) d\mathbf{x} \tag{8}$$

On the one hand, if $\Omega_i = \Omega_i^{ref}, \forall i \in [1, \ldots, K]$, then Equation (8) becomes an equality according to Equation (4).
On the other hand, Equation (8) is minimized when we have the equality. In that case,

$$\varphi(d_i(\mathbf{x}, \partial\Omega_i^{ref})) = \min_{1 \leq j \leq K} \left(\varphi(d_j(\mathbf{x}, \partial\Omega_j^{ref}))\right), \tag{9}$$

This occurs when $\mathbf{x} \in \Omega_i^{ref}$. Ω_i is therefore at least a subset of Ω_i^{ref}. As the definition of a partition implies $\bigcup_{i=1}^{K} \Omega_i^{ref} = \mathcal{X}$ and $\bigcup_{i=1}^{K} \Omega_i = \mathcal{X}$, and as $\Omega_i \subseteq \Omega_i^{ref}$, $\forall i \in [1, \ldots, K]$, then, the partitions must satisfy $\Omega_i = \Omega_i^{ref}$.

B. REFERENCES

[1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *IJCV*, vol. 1, pp. 321–332, 1988.

[2] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *IJCV*, vol. 22, no. 1, pp. 61–79, 1997.

[3] N. Paragios and R. Deriche, "Geodesic active regions for motion estimation and tracking," in *IEEE ICCV*, 1999, vol. II, pp. 224–240.

[4] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. on Im. Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[5] J. Kim, J.W. Fisher, A. Yezzi, M. Cetin, and A.S. Willsky, "Nonparametric methods for image segmentation using information theory and curve evolution," in *ICIP*, 2002, vol. 3, pp. 797–800.

[6] S. Jehan-Besson, M. Barlaud, and G. Aubert, "Dream2s: Deformable regions driven by an eulerian accurate minimization method for image and video segmentation," *IJCV*, vol. 53, no. 1, pp. 45–70, 2003.

[7] Y. Wang and L. Staib, "Boundary finding with correspondence using statistical shape models," in *IEEE Int. Conf. in Computer Vision and Pattern Recognition*, 1998, pp. 338–345.

[8] M. Leventon, E. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contour," in *CVPR*, Hilton Head Island, South Carolina, 2000, pp. 1316–1323.

[9] D. Cremers, F. Tischhäuser, J. Weickert, and C. Schnorr, "Diffusion-snakes: Introducing statistical shape knowledge into the mumford-shah functional," vol. 50, no. 3, pp. 295–313, 2002.

[10] M. Rousson and N. Paragios, "Shape priors for level set representations," in *ECCV*, Copenhagen, Denmark, 2002, pp. 78–92.

[11] Y. Chen, H.D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K.S. Gopinath, R.W. Briggs, and E.A. Geiser, "Using prior shapes in geometric active contours in a variational framework," *IJCV*, vol. 50, no. 3, pp. 315–328, 2002.

[12] X. Bresson, P. Vandergheynst, and J.P. Thiran, "A priori information in image segmentation: Energy functional based on shape statistical model and image information," in *ICIP*, September 2003, vol. 3, pp. 428–428.

[13] M. Gastaud, M. Barlaud, and G. Aubert, "Combining shape prior and statistical features for active contour segmentation," *IEEE TCSVT*, vol. 14, no. 5, pp. 726–734, May 2004.

[14] A. Foulonneau, P. Charbonnier, and T. Heitz, "Geometric shape priors for region-based active contours," in *ICIP*, 2003, vol. III, pp. 413–416.

[15] A. Gouze, C. De Roover, A. Herbulot, E. Debreuve, M. Barlaud, and B. Macq, "Watershed-driven active contours for moving object segmentation," in *ICIP*, Genova, Italy, 2005.

[16] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE PAMI*, vol. 13, no. 6, 1991.

[17] S. Osher and J.A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on hamilton-jacobi formulation," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.

[18] C. De Roover, M. Gabbouj, and B. Macq, "An accurate semiautomatic segmentation scheme based on watershed and change detection mask," *Proc. of SPIE, Im. and Video Comm. and Process.*, vol. 5685, pp. 479–488, January 2005.

[19] O. Cuisenaire and B. Macq, "Fast euclidean distance transformations by propagation using multiple neighbourhoods," *Comput. vis. and Im. understanding*, vol. 76, no. 2, pp. 163–172, 1999.

[20] P.L. Correia and F. Pereira, "Objective evaluation of video segmentation quality," *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 186–200, February 2003.