SHAPE GRADIENT FOR IMAGE SEGMENTATION USING INFORMATION THEORY

A.Herbulot †, *S.Jehan-Besson* ‡, *M.Barlaud* †, *G.Aubert* §

†Laboratoire I3S - CNRS UNSA, 2000 route des Lucioles, 06903 Sophia-Antipolis FRANCE
‡Laboratoire GREYC-images, 6 Bd Marechal Juin, 14050 Caen Cedex FRANCE
§Laboratoire J.A.Dieudonne - CNRS, Parc Valrose, 06108 Nice Cedex 2 FRANCE
{herbulot,barlaud}@i3s.unice.fr, stephanie.jehan@greyc.ismra.fr, gaubert@math.unice.fr

ABSTRACT

This paper deals with video and image segmentation using region based active contours. We consider the problem of segmentation through the minimization of a new criterion based on information theory. We first propose to derive a general criterion based on the probability density function using the notion of shape gradient. This general derivation is then applied to criterions based on information theory, such as the entropy or the conditional entropy for the segmentation of sequences of images. We present experimental results on grayscale images and color videos showing the accuracy of the proposed method.

1. INTRODUCTION

In many image processing problems such as segmentation, tracking or classification, the purpose is to extract image regions that minimize an energy. In this paper, we propose to minimize a criterion based on the entropy for color images and videos segmentation.

The issue is to find a region Ω with homogeneous features, such as mean, variance, texture...This region is characterized by a minimum of an energy criterion including region and boundary features. The region features are modeled as a combination of region integrals of a descriptor $k(x, \Omega)$ that depends on this region Ω and on its features. We note $J(\Omega)$ the criterion:

$$J(\Omega) = \int_{\Omega} k(x, \Omega) dx \tag{1}$$

We use the shape gradient method presented in [1] to derive this criterion and obtain a force F that we apply to an active contour. Given an initial contour Γ_0, the active contours method consists in applying a force to this initial contour such that it evolves towards the object of interest. The active contour is modeled by a parametric curve $\Gamma(s, \tau)$ where s is its arc-length and τ an evolution parameter. It evolves using the following Partial Differential Equation:

$$\frac{\partial \Gamma(s, \tau)}{\partial \tau} = \mathbf{v} = F\mathbf{N} \quad \text{with} \quad \Gamma(\tau = 0) = \Gamma_0 \tag{2}$$

where $\mathbf{v}$ is the velocity vector of $\Gamma(s, \tau)$, F its amplitude along the unit inward normal $\mathbf{N}$ of the curve.

Active contours were originally boundary methods and have been introduced in [2], and geodesic active contours in [3]. The energy to minimize includes region and boundary functionals, like in [4, 5, 6].

First, we present the problem of optimization of region and boundary functionals with active contours, then the general framework of shape derivation. In section 3, we define a criterion based on the entropy and use derivation tools to obtain the equation of evolution of the active contour. We do the same with two other criterions: an approximation of the entropy, and the conditional entropy. Then we show some experimental results in section 4.

2. PROBLEM STATEMENT AND GENERAL FRAMEWORK

2.1. Problem statement

Let us define a general criterion.

$$J(\Omega) = \int_{\Omega} \varphi(q(I(x), \Omega)) dx \tag{3}$$

where

- $q(I(x), \Omega)$ is the probability to have the intensity $I(x)$ with x in the region Ω
- φ is a function: $\mathcal{R}^+ \to \mathcal{R}^+$ of this probability which may be relative to the entropy, or the conditional entropy

We compute the derivative of this criterion by using the notion of shape gradient proposed in [7]. From this derivative, we obtain the velocity of the curve evolution. So that the active contour converges to a local minimum, and hopefully towards the boundary of the object Ω_{in}. Let us see how to derive the functional 3.

2.2. Derivation tool

In this section we use the general criterion $J(\Omega)$ defined in (3). The probability density function of the intensity of the image in the region Ω is estimated using the Parzen window method:

$$q(I(x),\Omega) = \frac{1}{|\,\Omega\,|}\int_{\Omega} K(I(x) - I(\hat{x}))d\hat{x} \tag{4}$$

where K is the gaussian kernel of this estimation with 0-mean and σ-variance

We can not compute a direct derivation of this criterion with respect to Ω. A solution is to use the shape gradient method using a dynamic scheme where the region Ω becomes continuously dependent on an evolution parameter τ. The criterion is then defined as follows:

$$J(\Omega(\tau)) = \int\limits_{\Omega(\tau)} \varphi\Big(q(I(x),\Omega(\tau))\Big)dx \tag{5}$$

To simplify the notations, we note $\Omega = \Omega(\tau)$.
The contour evolution equation is obtained by derivating this criterion with respect to τ in an Eulerian framework. The Eulerian derivative dJr of this criterion in the direction **V** represents the variation of $J(\Omega(\tau))$ due to both the deformation of integration domain $\Omega(\tau)$ in the direction of **V** and the variation of φ (see [7, 8] for details). This derivative is:

$$\begin{aligned} dJ_r(\Omega,\mathbf{V}) &= \int_{\Omega} \varphi_r'(q(I(x),\Omega),\mathbf{V})dx \\ &\quad - \int_{\partial\Omega} \varphi(q(I(s),\Omega))(\mathbf{V}\cdot\mathbf{N})ds \end{aligned} \tag{6}$$

where $\varphi_r'(q(I(x),\Omega),\mathbf{V})$ is the domain derivative of φ in the direction **V** and **N** is the unit inward normal of the curve.

The first term is the integral of the domain derivative of φ. It comes from the dependence of the descriptor φ upon the region Ω, whereas the second term comes from the evolution of the region Ω itself.

The domain derivative of φ is the following:

$$\begin{aligned} \varphi_r'(q(I(x),\Omega),\mathbf{V}) &= \frac{1}{|\,\Omega\,|}\int\limits_{\partial\Omega} \varphi'(q(I(x),\Omega)[q(I(x),\Omega) \\ &- K(I(x) - I(s))](\mathbf{V}\cdot\mathbf{N})ds \end{aligned} \tag{7}$$

where $\varphi'(q)$ is the derivative of φ with respect to q.

From this Eulerian derivative, we deduce the velocity vector of the active contour that will make it evolve as fast as possible towards a minimum of the functional. According to the Cauchy-Schwartz inequality, the fastest decrease of $dJr(\Omega)$ is obtained with the following equation:

$$\frac{\partial\Gamma}{\partial\tau} = \mathbf{v} = \Big(\varphi(q(I(x),\Omega)) + A(x,\Omega)\Big)\mathbf{N} \tag{8}$$

where $A(x,\Omega)$ is a term coming from the dependence of the descriptors with the region and will be detailed in the following examples.

3. THE ENTROPY

In this section we present a functional based on information theory: the entropy. We present the criterion and the equations obtained by using derivation tools described in the previous section.

3.1. Minimization of entropy

Let us consider the general functional introduced in section 2. For the entropy we use the following function φ:

$$\varphi(q(I(x),\Omega)) = -q(I(x),\Omega)\ln q(I(x),\Omega) \tag{9}$$

The functional we want to minimize is then given by the following expression:

$$E(\Omega) = \int_{\Omega} -q(I(x),\Omega)\ln q(I(x),\Omega)dx \tag{10}$$

We derive this criterion by using the method proposed in [7] and we obtain the Eulerian derivative in the direction **V**.

Let us first compute the domain derivative φ_r' whose expression is given by equation (7).
We have:

$$\varphi'(q(I(x),\Omega)) = -\ln q(I(x),\Omega) - 1$$

Hence, we obtain:

$$\begin{aligned} \varphi_r'(q(I(x),\Omega),\mathbf{V}) &= \frac{1}{|\,\Omega\,|}\int\limits_{\partial\Omega}[-\ln q(I(x),\Omega) - 1]. \\ &\quad [q(I(x),\Omega) - K(I(x) - I(s))](\mathbf{V}\cdot\mathbf{N})ds \end{aligned} \tag{11}$$

With this domain derivative, we can write the first term of the Eulerian derivative:

$$\begin{aligned} \int_{\Omega}\varphi_r'(q(I(x),\Omega),\mathbf{V})dx &= \int_{\Omega}\frac{1}{|\,\Omega\,|}\int\limits_{\partial\Omega}[-\ln q(I(x),\Omega) - 1]. \\ &\quad [q(I(x),\Omega) - K(I(x) - I(s))](\mathbf{V}\cdot\mathbf{N})ds\,dx \end{aligned}$$

We switch the order of integration and we obtain the following formulation:

$$\begin{aligned} \int_{\Omega}\varphi_r'(q(I(x),\Omega),\mathbf{V})dx &= \int\limits_{\partial\Omega}\Big(\frac{1}{|\,\Omega\,|}[E(\Omega) - 1 \\ &+ \int_{\Omega}K(I(x) - I(s))\ln q(I(x),\Omega)dx] \\ &+ q(I(s),\Omega)\Big)(\mathbf{V}\cdot\mathbf{N})ds \end{aligned}$$

Thus, the Eulerian derivative of the criterion is:

$$\begin{aligned} dE_r(\Omega,\mathbf{V}) &= \int\limits_{\partial\Omega}\Big[\frac{1}{|\,\Omega\,|}\Big(E(\Omega) - 1 \\ &+ \int_{\Omega}K(I(x) - I(s))\ln q(I(x),\Omega)dx\Big) \\ &+ q(I(s),\Omega) \\ &+ q(I(x),\Omega)\ln q(I(x),\Omega)\Big](\mathbf{V}\cdot\mathbf{N})ds \end{aligned}$$

From which we deduce the following evolution equation:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \tau} &= \Big[-q(I(\hat{x},\Omega))(\ln q(I(\hat{x},\Omega))+1) - \frac{1}{\mid \Omega \mid}\Big(E(\Omega) \\ &- 1 + \int_{\Omega} K(I(x)-I(\hat{x}))\ln q(I(x),\Omega)dx\Big)\Big]\mathbf{N} \quad (12)\end{aligned}$$

In the experimentations, we use a competition between the background region and the object region and the criterion to minimize is:

$$J(\Omega_{in},\Omega_{out},\Gamma) = E(\Omega_{in}) + E(\Omega_{out}) + \int_{\Gamma} \lambda ds \quad (13)$$

where λ is a regularization parameter.

3.2. Minimization of an approximation of entropy

Another criterion based on the entropy has been studied in the same framework. As shown in [9], we obtain an approximation of the entropy using the weak law of large numbers. The function φ defined in section 2 is denoted as follows:

$$\varphi(q(I(x),\Omega)) = -\ln q(I(x),\Omega)$$

and the criterion to minimize is:

$$E_A(\Omega) = \int_{\Omega} -\ln q(I(x),\Omega)dx \quad (14)$$

Then, the eulerian derivative of this criterion has the following formulation:

$$\begin{aligned}d_r E_A(\Omega,\mathbf{V}) &= \int_{\partial\Omega} \Big(-1 + \frac{1}{\mid \Omega \mid}\int_{\Omega} \frac{K(I(x)-I(s))}{q(I(x),\Omega)}dx \\ &+ \ln q(I(s),\Omega)(\mathbf{V}\cdot\mathbf{N})ds \quad (15)\end{aligned}$$

Hence, the expression of the velocity of the curve is:

$$\frac{\partial \Gamma}{\partial \tau} = \Big[1 - \frac{1}{\mid \Omega \mid}\int_{\Omega} \frac{K(I(x)-I(\hat{x}))}{q(I(x),\Omega)}dx - \ln q(I(\hat{x}),\Omega)\Big]\mathbf{N} \quad (16)$$

As the previous criterion we use a competition between the background region and the object region:

$$J(\Omega_{in},\Omega_{out},\Gamma) = E_A(\Omega_{in}) + E_A(\Omega_{out}) + \int_{\Gamma} \lambda ds \quad (17)$$

3.3. Minimization of the conditional entropy

We consider that we want to segment an object region Ω_{in}. We note Ω_{out} the background region, and Γ is the common boundary. Let us define the following functional:

$$\begin{aligned}E_C(\Omega_{in},\Omega_{out}) &= E(\Omega_{in}) \mid \Omega_{in} \mid + E(\Omega_{out}) \mid \Omega_{out} \mid \\ &= \mid \Omega_{in}\mid \int_{\Omega_{in}} -q(I(x),\Omega_{in})\ln q(I(x),\Omega_{in})dx \quad (18) \\ &+ \mid \Omega_{out}\mid \int_{\Omega_{out}} -q(I(x),\Omega_{out})\ln q(I(x),\Omega_{out})dx\end{aligned}$$

This criterion represents the conditional entropy.
We use the derivation tools to obtain the Eulerian derivative of the functional $J(\Omega) = E(\Omega) \mid \Omega \mid$:

$$\begin{aligned}dJ_r(\Omega,\mathbf{V}) &= \partial E_r \mid \Omega \mid - E(\Omega)\int_{\partial\Omega}(\mathbf{V}\cdot\mathbf{N})d\hat{x} \\ &= \int_{\partial\Omega} \Big[K(I(x)-I(\hat{x}))(\ln q(I(x),\Omega)+1) - 1 \\ &+ \mid \Omega \mid q(I(\hat{x}),\Omega)\ln q(I(\hat{x}),\Omega)\Big](\mathbf{V}\cdot\mathbf{N})d\hat{x}\end{aligned}$$

Then we deduce the velocity of the curve:

$$\begin{aligned}\frac{\partial \Gamma}{\partial \tau} &= \Big[- \mid \Omega \mid q(I(\hat{x}),\Omega)\ln q(I(\hat{x}),\Omega) + 1 \quad (19) \\ &- \int_{\Omega} K(I(x)-I(\hat{x}))(\ln q(I(x),\Omega)+1)dx\Big]\mathbf{N}\end{aligned}$$

4. EXPERIMENTAL RESULTS

4.1. Implementation

In these experiments, we use a parametric method to implement the evolution equation: smoothing B-splines. We use this method instead of usual level-sets methods because it is less time consuming (see [10]). Furthermore smoothing B-spline approach combines a very low computational cost and a global robustness to noisy data.

4.2. Results on grayscale medical images

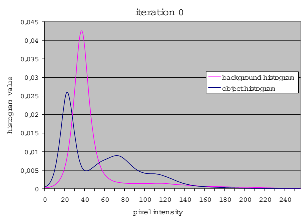
We consider an osteoporosis grayscale image, for a simple representation of the histogram. We use the criterion of the entropy (equation (10) and evolution equation (12)). Fig. 1 shows how histograms are evolving during the segmentation.We notice that during the segmentation processing, both histograms and segmentations are evolving and histograms are getting more distinct due to the statistic probabilities of the regions.

4.3. Results on color video

For color images and videos segmentation we use 3D-histograms of the regions. Thus, we define the probability $q(I(x),\Omega) = q(I_1(x),I_2(x),I_3(x),\Omega)$ with $(I_1,I_2,I_3) = (H,S,V)$ where H,S and V represent the three components of the Hue-Saturation-Value color space. This probability is the joint probability. We quantify this histogram with an uniform step quantization, identical for the three components and we estimate it with the Parzen method with a parameter σ between 2 and 5.
Regions of interest are homogeneous regions, like the face on the sequence *Erik*. Fig. 2 shows that the criterion using the minimization of an approximation of the entropy (equation (14)) fails to segment correctly the region near the ear

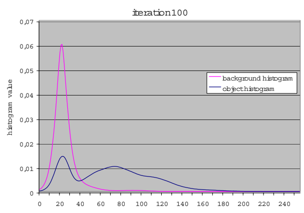
because the background color is close to the skin one. The segmentation obtained with the minimization of the entropy (evolution equation (12)) is more accurate.

(a) Initial curve



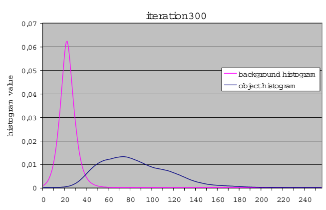
(b) Initial histograms

(c) Iteration 100



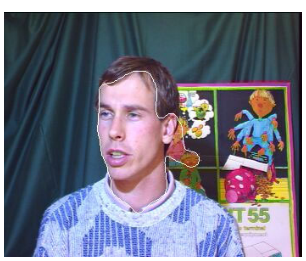
(d) Iteration 100

(e) Final segmentation

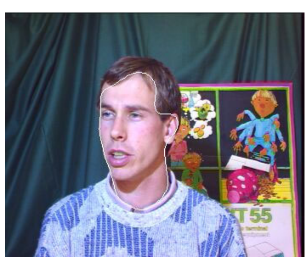


(f) Final histograms

Fig. 1. Evolution of segmentation and histograms with the minimization of the entropy



(a) Segmentation with an approximation of the entropy



(b) Segmentation with the entropy

Fig. 2. Results with the criterion minimizing an approximation of the entropy and the criterion minimizing the entropy

5. CONCLUSION

In this paper we have presented a general framework based on information theory for image segmentation using active contours. We use a non-parametric and statistic method to define the functionals we want to minimize. By derivating these functionals using a gradient shape method, we obtain the curve evolution. This general derivation is applied to descriptors like the entropy and the conditional entropy and we show some experimental results on grayscale and color images. Further research may be concerned with the evaluation of other criterions like for example the mutual entropy.

6. REFERENCES

[1] M.C. Delfour and J.P. Zolésio, *Shape and geometries*, Advances in Design and Control, SIAM, 2001.

[2] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *IJCV*, vol. 1, 1988.

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

[4] S. Zhu and A. Yuille, “Region competition: unifying snakes, region growing, and bayes/MDL for multiband image segmentation,” *PAMI*, vol. 18, pp. 884–900, september 1996.

[5] N. Paragios and R. Deriche, “Geodesic active regions and level set methods for supervised texture segmentation,” *International Journal of Computer Vision*, vol. 46, no. 3, pp. 223, 2002.

[6] T. Chan and L. Vese, “Active contours without edges,” *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[7] S. Jehan-Besson, M. Barlaud, and G. Aubert, “DREAM^{2}S: Deformable regions driven by an eulerian accurate minimization method for image and video segmentation,” *IJCV*, vol. 53, no. 1, 2003.

[8] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, “Image segmentation using active contours: Calculus of variations or shape gradients ?,” *SIAM Applied Mathematics*, vol. 1, no. 2, 2003.

[9] J. Kim, J.W. Fisher, A. Yezzi, M. Cetin, and A.S. Wilsky, “Nonparametric methods for image segmentation using information theory and curve evolution,” in *ICIP*, 2002, pp. III: 797–800.

[10] F. Precioso, M. Barlaud, T. Blu, and M. Unser, “Smoothing B-spline active contour for fast and robust image and video segmentation,” in *ICIP*, Barcelona, Spain, September 2003, IEEE, vol. 1, pp. 137–140.