SHAPE GRADIENT FOR MULTI-MODAL IMAGE SEGMENTATION USING JOINT INTENSITY DISTRIBUTIONS

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

[†]Laboratoire I3S - CNRS UNSA, 2000 route des Lucioles, 06903 Sophia-Antipolis FRANCE
{herbulot,barlaud}@i3s.unice.fr
[‡]Laboratoire GREYC-images, 6 Bd Marechal Juin, 14050 Caen Cedex FRANCE
stephanie.jehan@greyc.ismra.fr
[§]Laboratoire J.A.Dieudonne - CNRS, Parc Valrose, 06108 Nice Cedex 2 FRANCE
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 joint entropy for the segmentation of sequences of images. We present experimental results on multimodal images 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 multimodal 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 includes region and boundary functionals, like in [4, 5].

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 extend the framework to joint entropy and mutual information. Then we show some experimental results in sec.4.

2. PROBLEM STATEMENT AND GENERAL FRAMEWORK

2.1. Problem statement

We note:

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

Let us define a general criterion.

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

We compute the derivative of this criterion by using the notion of shape gradient proposed in [6]. From this derivative, we obtain the velocity of the curve evolution.

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:

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

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, let us denote $\Omega = \Omega(\tau)$.
The contour evolution equation is obtained by deriving 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 [6, 7] for details). This derivative is:

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

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:

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

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.

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:

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

We derive this criterion by using the method proposed in [6] 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:

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

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

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

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

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

Thus, the Eulerian derivative of the criterion is:

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

From which we deduce the following evolution equation:

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

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

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

where λ is a regularization parameter.

3.2. Minimization of joint entropy

Let X and Y denote two random variables with marginal probability distributions $q_X(x)$ and $q_Y(y)$. $q_{XY}(x,y)$ is the joint probability distribution. The joint entropy has the following definition:

$$H(X,Y)=-\int_{\Omega_X}\int_{\Omega_Y} q_{XY}(x,y)\ln q_{XY}(x,y)\,dx\,dy \quad (14)$$

In the case of multimodal images, we compute the joint entropy between two channels of the image. If we note $I(x)=(I_X(x),I_Y(x))$, the probability distributions are:

$$\begin{aligned} q_X(x)=q(I_X(x),\Omega) \quad &; \quad q_Y(x)=q(I_Y(x),\Omega) \\ q_{XY}(x,y) &= q(I_X(x),I_Y(y),\Omega) \end{aligned} \quad (15)$$

and the joint entropy is:

$$H_{XY}(\Omega)=-\int_\Omega q(I_X(x),I_Y(y),\Omega)\ln q(I_X(x),I_Y(y),\Omega)dx\,dy$$

Using the joint probability distributions instead of the probability distributions in equation(12), we obtain:

$$\begin{aligned} \frac{\partial\Gamma}{\partial\tau} &= \Big[-q(I_X(\hat{x}),I_Y(\hat{x}),\Omega)(\ln q(I_X(\hat{x}),I_Y(\hat{x}),\Omega)+1) \\ &- \frac{1}{|\,\Omega\,|}\Big(H_{XY}(\Omega)-1+\int_\Omega K_{XY}(I_X(x)-I_X(\hat{x}), \\ &\quad I_Y(x)-I_Y(\hat{x})).\ln q(I_X(x),I_Y(x),\Omega)dx\Big)\Big]\mathbf{N} \end{aligned} \quad (16)$$

where

$$K_{XY}(x,y)=\frac{1}{2\pi\sigma^2}\exp-\frac{1}{2\sigma^2}(x\ y)(x\ y)^T \quad (17)$$

In the experiments we minimize the following criterion:

$$J(\Omega_{in},\Omega_{out},\Gamma)=H_{XY}(\Omega_{in})+H_{XY}(\Omega_{out})+\int_\Gamma \lambda ds \quad (18)$$

3.3. Mutual information

Let $H(X)$ and $H(Y)$ denote the entropy of X and Y respectively, and $H(X,Y)$ their joint entropy.
The *Mutual Information* (MI) noted $MI(X,Y)$ or relative entropy mesures the degree of dependence of X and Y by measuring the Kullback-Leibler distance between the joint distribution and the product of the distributions. MI can be written like follows:

$$MI(X,Y) = H(X)-H(X/Y) \quad (19)$$

Like in [8], we define the binary label L determined by the curve Γ as a mapping from the image domain to $\{R_{in},R_{out}\}$:

$$L(x)=\begin{cases} R_{in} & \text{if } x\in\Omega_{in} \\ R_{out} & \text{if } x\in\Omega_{out} \end{cases}$$

We consider the mutual information between the label and the image intensity:

$$MI(I(X),L(X))=H(I(X))-H(I(X)|L(X)) \quad (20)$$

with X a random variable uniformly distributed over the image domain.

The mutual information is maximized if $R_{in}=\Omega_{in}$ and $R_{out}=\Omega_{out}$, ie if the segmentation is correct. The functional to minimize is then given by:

$$J(\Omega_{in},\Omega_{out},\Gamma)=-MI(I(X),L(X))+\int_\Gamma \lambda ds \quad (21)$$

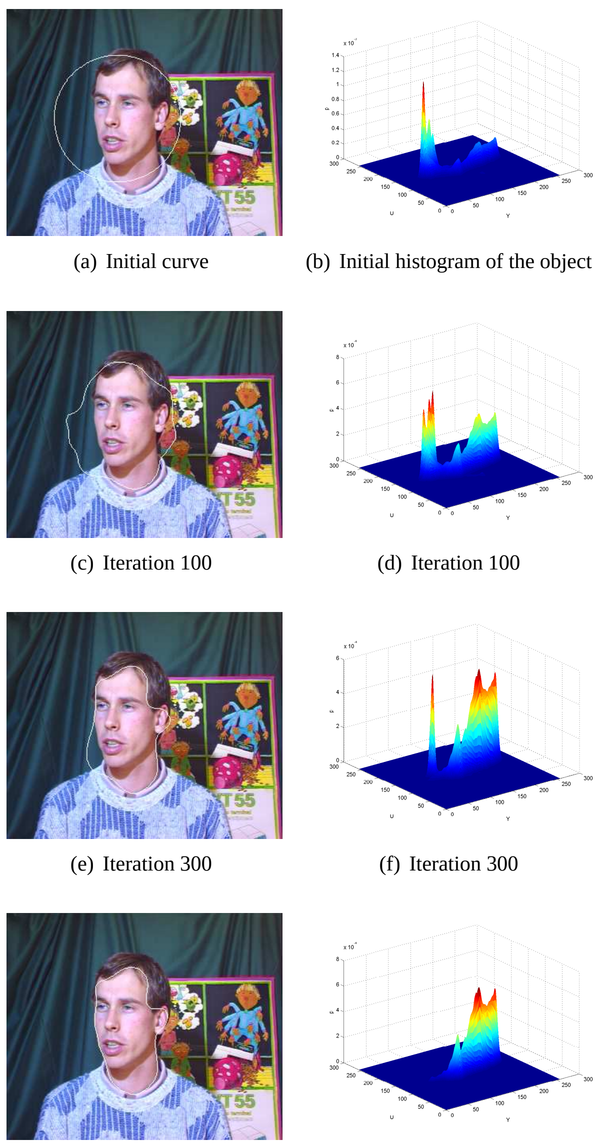
where λ is a regularization parameter. See [9] for more details.

4. EXPERIMENTAL RESULTS

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.
Regions of interest are homogeneous regions, like the face on the sequence *Erik*. The color images are in the YUV color space. For these experiments we use 2D-histograms for a easier visualization. Thus we use the channels Y and U and we ignore the V one. Indeed, for face segmentation, the channel V (the blue chrominance) doesn't bring any meaningful information. We use the criterion of joint entropy with the channels Y and U. We quantify this histogram with an uniform step quantization, identical for the two components and we estimate it with the Parzen method with a parameter σ between 2 and 5.
We use the criterion(18) and the evolution equation(16). Fig.(4) shows the evolution of the curve and the evolution of the histogram of the object (region inside the curve).

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 deriving these functionals using a gradient shape method, we obtain the curve evolution. This general derivation is applied to descriptors like the entropy and the joint entropy and we show some experimental results on color images. Some other criterions are studied like mutual information, an information measure which is very used for medical image registration (see [11, 12]).



(a) Initial curve (b) Initial histogram of the object

(c) Iteration 100 (d) Iteration 100

(e) Iteration 300 (f) Iteration 300

(g) Final segmentation (h) Final histogram

Fig. 1. Evolution of segmentation and histograms with the minimization of the joint 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, pp. 321–332, 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] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[6] 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.

[7] 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.

[8] 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 *ICIP02*, 2002, pp. III: 797–800.

[9] A. Herbulot, S. Jehan-Besson, M. Barlaud, and G. Aubert, "Shape gradient for image segmentation using mutual information," in *ICIP*, (submitted) 2004.

[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.

[11] W. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, no. 1, pp. 35–51, 1996.

[12] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, no. 2, 1997.