An Approach for Dynamic Combination of Region and Boundary Information in Segmentation

Mohand Saïd Allili and Djemel Ziou
Département d'Informatique, Université de Sherbrooke, J1K 2R1, QC, Canada
Emails: {ms.allili,d.ziou}@usherbrooke.ca

Abstract

Image segmentation combining boundary and region information has been the subject of numerous research works in the past. This combination is usually subject to arbitrary weighting parameters (hyper-parameters) that control the contribution of boundary and region features during segmentation. In this work, we investigate a new approach for estimating the hyper-parameters adaptively to segmentation. The approach takes its roots from the physical properties of the energy functional controlling segmentation and a Bayesian formulation of segmentation and hyper-parameters estimation.

1. Introduction

Image segmentation is one of the most studied topics in computer vision and still the focus of many recent works. To enforce the precision and robustness of segmentation, several works in the past investigated the combination of boundary and region information in the segmentation [5, 7]. These approaches take advantage of the complementary information provided by boundary and region features to build homogeneous regions whose boundaries lie on strong discontinuities of the image. Quite often, the combination of boundary and region information in those approaches is a weighted sum of two objective functions, each of which achieving a segmentation by optimizing a criterion based on either region or boundary information [3, 5]. In this field, variational models have been used for their consistent and natural way of formulating segmentation by energy minimization. The final solution for segmentation is achieved by minimizing energy functionals with respect to the region contours, via the resolution of partial differential equations (PDEs) [3, 8]. It remains that the combination of boundary and region information in those functionals is formulated as an arbitrary weighting sum, where the weights are fixed in advance by the user. The influence of this on segmentation quality and convergence has not been shown so far.

In this work, we propose an automatic estimation of the *hyper-parameters* for a segmentation combining boundary and region information. We suppose the image is composed of a fixed, but arbitrary number of regions. The region information is formulated using a mixture of Gaussians, whose parameters are calculated adaptively to segmentation. Boundary information is formulated using a multi-band edge detector. We show on various examples that the approach speeds up convergence of segmentation while enhancing its robustness to shadowing and illumination changes. This paper is organized as follows: Section 2 presents the problem of fixed hyper-parameters in segmentation. Section 3 gives a formulation of the proposed model. Section 4 presents some experiments for our model validation.

2. Problem statement

Let $U = (u_1, \ldots, u_n)$ be a multi-valued image defined on the domain $\Omega = \mathbb{Z}^+ \times \mathbb{Z}^+$; where $\mathbb{Z}^+$ is the set of positive integers and $n \geq 1$. We use the notation Ω_k and $\partial\Omega_k$ to designate, respectively, a region and its boundaries. The segmentation aims to build a partition of the image composed of M regions $P = \{\Omega_1, ..., \Omega_M\}$, where $\bigcup_{i=1}^{M} \Omega_k = \Omega$. Variational segmentation combining boundary and region information has been formulated in several research works [3, 8] by minimizing the following functional:

$$E(\bigcup_{k=1}^{M} \partial\Omega_k(s), \mathbf{\Theta}) = \alpha E_b + \beta E_r, \qquad (1)$$

$$\text{where } E_b = \sum_{k=1}^{M} \oint_{\partial\Omega_k} g(\nabla I(\partial\Omega_k(s)))ds \qquad (2)$$

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

$$\text{and } E_r = \sum_{k=1}^{M} \iint_{\Omega_k} -\log\left[t_k(\mathbf{x})\right] dxdy, \quad (3)$$

where $\mathbf{\Theta}$ designates the mixture parameters, composed of the parameters θ_k of each pdf and the mixing probabilities π_k, $k \in \{1, \ldots, M\}$. $t_k(\mathbf{x}) = p(\theta_k|U(\mathbf{x})) = \frac{\pi_k p(U(\mathbf{X})|\theta_k)}{\sum_{j=1}^{M} \pi_j p(U(\mathbf{X})|\theta_j)}$. The boundary information is added in the first term of functional (1), using the formalism of the GAC model [2], where the color boundary plausibility $\nabla I(\partial\Omega_k(s))$ is calculated by the method proposed in [4], and g is a strictly decreasing function. The second term of the functional assigns pixels to regions by minimizing the Bayes error classification. To segment an image, functional (1) is minimized using Euler-Lagrange equations. After introducing the level set formalism [2], the following motion equation is given for each region contour:

$$\frac{d\phi_k}{dt} = (\alpha V_b - \beta V_r)\,|\nabla\phi_k| \quad (4)$$

where

$$\left\{ \begin{array}{ll} V_b & = g(\nabla I(\phi_k))\kappa + \nabla g(\nabla I(\phi_k))\frac{\nabla\phi_k}{|\nabla\phi_k|} \\ V_r & = \log\left(\pi_k p(U|\theta_k)\right) - \log\left(\pi_h p(U|\theta_h)\right) \end{array} \right. \quad (5)$$

where $\phi_k : \mathbb{R}^2 \rightarrow \mathbb{R}$ is a level set function associated with the contour $\partial\Omega_k$ ($\phi_k(\mathbf{x}) < 0$ if $\mathbf{x}$ is inside the contour $\partial\Omega_k$ and $\phi_k(\mathbf{x}) > 0$ otherwise). κ is the curvature of the zero level set. The term V_b represents the boundary velocity that regularizes the curve and aligns it with the region boundaries. The term V_r represents the region velocity and has the role of moving the contours inside the regions by minimizing the Bayes error classification of the pixels. Here, the region term is made as a competition for a given pixel between the current region Ω_k and the region $\Omega_h \neq \Omega_k$ that has the maximum posterior probability for the pixel value.

In Eq. (4), α and β control the contribution of the boundary and region terms. By setting a small value for β and a large one for α, the contour was stopped at pixels with small gradient magnitude, as shown in Fig. 1.b. By setting a small value for α and a large one for β, the contour flicked to capture parts of the background, as shown in Fig. 1.c. Finally, setting similar values for ($\alpha = \beta = 0.5$) leads to the result shown in Fig. 1.d, where the contour captures the region of interest, but it is stopped by the eyes region, even though it does not constitute a real boundary to the face.

If we want to include the eyes into the face region, one can tune the hyper-parameters in such way that boundary information is activated only when the contour is in the vicinity of a true boundary. Our approach in this work tries to achieve this objective by giving priority to boundary information only if the neighborhood of a pixel is characterized by a high gradient response. Otherwise, the region module is favored.

(a) (b) (c) (d)

Figure 1. Segmentations using fixed hyper-parameters α and β: (a) represents the contour initialization; (b), (c) and (d) represent segmentation results obtained using ($\alpha = 0.9, \beta = 0.1$), ($\alpha = 0.1, \beta = 0.9$) and ($\alpha = 0.5, \beta = 0.5$), respectively.

3. Formulation of the model

To establish adaptive values for the hyper-parameters α and β, we first rewrite the segmentation as probability maximization. The optimal segmentation of each region is given by a contour which minimizes functional (1). Let E_b and E_r be the values of the boundary and region energies. We may claim that the values of theses energies are dissipative (since the PDE minimizing them is in a parabolic form) [6]. Thus, these energies are expected to diminish over time, yielding exponentially decaying solution for the PDE. We, then, obtain the optimal location for any boundary $\partial\Omega_k$ by maximizing the following probability over $\partial\Omega_k$ and $\mathbf{\Theta}$:

$$\begin{aligned} p(E_b, E_r|\Omega_k, \mathbf{\Theta}, \alpha, \beta) = \alpha\text{exp}\big\{ - \alpha E_b(\partial\Omega_k, \mathbf{\Theta})\big\} \\ \times\beta\text{exp}\big\{ - \beta E_r(\partial\Omega_k, \mathbf{\Theta})\big\} \quad (6) \end{aligned}$$

The multiplication by α and β is for normalizing the probabilities. The above formulation of the energy distribution relies on the following assumptions. First, we supposed the boundary and region segmentation probabilities are independent, which is reasonable in nature. Second, boundary and region energy probabilities are modeled using exponential distributions. The exponential distribution comes out as a natural choice since it gives the maximum of segmentation probability when the energy is minimized.

Eq. (6) defines a probability of a global state of segmentation for which we assume the hyper-parameters α and β are the same for all the pixels on which the energies E_b and E_r are calculated. If we want to make the segmentation independent for each pixel (up to its neighborhood), we can build a local segmentation map for the pixel where we suppose the hyper-parameters α and β are constant in its neighborhood, as illustrated in

Fig. 2. Let us denote by $W(\mathbf{x})$ a square window which surrounds the pixel $\mathbf{x} = (x, y)$. The size of $W(\mathbf{x})$ is $N = m \times m$. We denote by $\tilde{E}_b(\mathbf{x})$ and $\tilde{E}_r(\mathbf{x})$ the neighborhood boundary and region energies, respectively. These energies are given by:

$$\tilde{E}_b(\mathbf{x}) = \frac{1}{N} \sum_{(u) \in W(\mathbf{x})} E_b(u) \quad (7)$$

$$\tilde{E}_r(\mathbf{x}) = \frac{1}{N} \sum_{(u) \in W(\mathbf{x})} E_r(u) \quad (8)$$

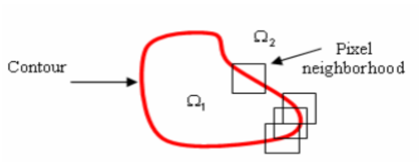

Figure 2. Neighborhood system used for hyper-parameters calculation.

To make the segmentation formulated by (6) independent of α and β, a correct Bayesian treatment is to integrate them out over any prediction [1]. The evidence of each region contour localization is then given by integrating the energy (segmentation) probability over α and β. I each pixel neighborhood, this is given by:

$$p(\tilde{E}_b(\mathbf{x})) = \int_{-\infty}^{\infty} p(\alpha, \tilde{E}_b(\mathbf{x})) d\alpha \quad (9)$$

$$\propto \int_{0}^{\infty} p(\tilde{E}_b(\mathbf{x})|\alpha) p(\alpha) d\alpha \quad (10)$$

$$p(\tilde{E}_r(\mathbf{x})) \propto \int_{0}^{\infty} p(\tilde{E}_r(\mathbf{x})|\beta) p(\beta) d\beta \quad (11)$$

The integrals (10) and (11) represent the expectations of the boundary and region energies with respect to the hyper-parameters α and β. The calculation of these integrals requires choosing specific priors for α and β. Since we assume that no information is available in advance about them, one proper choice is to use non-informative priors, as proposed in [1]:

$$p(\alpha) \propto \alpha^{-1}; p(\beta) \propto \beta^{-1} \quad (12)$$

In Eqs. (10) and (11), we assume the hyper-parameters are constant in the neighborhood, which simplifies the integrals to the following ones:

$$p(\tilde{E}_b(\mathbf{x})) \propto \int_0^{\infty} \exp\{-\alpha \tilde{E}_b(\mathbf{x})\} d\alpha = \left[\tilde{E}_b(\mathbf{x})\right]^{-1} \quad (13)$$

$$p(\tilde{E}_r(\mathbf{x})) \propto \int_0^{\infty} \exp\{-\beta \tilde{E}_r(\mathbf{x})\} d\beta = \left[\tilde{E}_r(\mathbf{x})\right]^{-1} \quad (14)$$

Putting to logarithm, then differentiating the above equations with respect to each contour $\partial\Omega_k$, we obtain:

$$-\nabla \log \left[p(\tilde{E}_b(\mathbf{x})) p(\tilde{E}_r(\mathbf{x}))\right]$$

$$\propto \nabla \log \left[\tilde{E}_b(\mathbf{x})\right] + \nabla \log \left[\tilde{E}_r(\mathbf{x})\right] \quad (15)$$

If we perform the same differentiation but without integrating the hyper-parameters (i.e., by defining $p(\tilde{E}_b(\mathbf{x}))$ and $p(\tilde{E}_r(\mathbf{x}))$ as formulated in Eq. (6)), we obtain:

$$-\nabla log \left[p(\tilde{E}_b(\mathbf{x})) p(\tilde{E}_r(\mathbf{x}))\right]$$

$$\propto \alpha \cdot \nabla \tilde{E}_b(\mathbf{x}) + \beta \cdot \nabla \tilde{E}_r(\mathbf{x}) \quad (16)$$

By comparing Eqs. (15) and (16) term to term, we obtain the values of α and β as follows:

$$\alpha \propto [\tilde{E}_b(\mathbf{x})]^{-1}; \beta \propto [\tilde{E}_r(\mathbf{x})]^{-1} \quad (17)$$

Eq. (17) gives the values of α and β in Eq. (6) which minimize the energy locally with respect to each region contour $\partial\Omega_k$. Thus, replacing the values of α and β in Eq. (4) will yield better minimization of the energy and, therefore, better localization of the contour. Indeed, Eq. (17) means that whenever the parts inside and outside a contour are in the same region, the energy $\tilde{E}_r$ will be small and β should be increased. If the contour is in the vicinity of a region boundary, then $\tilde{E}_b$ will be small and α should be increased. The final algorithm for segmentation is given as follows:

1- Initialize the region contours.

2- Until the convergence of the contours do:

a- Evolve the contours using Eq. (5).

b- Re-initialize the level set functions.

c- Re-estimate the region parameters $\mathbf{\Theta}$ by:

$$\hat{\pi}_k = \frac{\iint_{\Omega_k} t_k(\mathbf{x}) d\mathbf{x}}{\iint_{\Omega_k} d\mathbf{x}} \quad (18)$$

$$\hat{\mu}_k = \frac{\iint_{\Omega_k} t_k(\mathbf{x}) U(\mathbf{x}) d\mathbf{x}}{\iint_{\Omega_k} t_k(\mathbf{x}) d\mathbf{x}} \quad (19)$$

$$\hat{\Sigma}_k = \frac{\iint_{\Omega_k} t_k(\mathbf{x}) [U(\mathbf{x}) - \mu_k][U(\mathbf{x}) - \mu_k]^T d\mathbf{x}}{\iint_{\Omega_k} t_k(\mathbf{x}) d\mathbf{x}} \quad (20)$$

d- Re-estimate the hyper-parameters using Eqs. (17).

In the previous equations, π_k, μ_k and Σ_k designate, respectively, the mixing parameter, the mean vector and the covariance matrix of the region Ω_k.

4. Experiments

To show the effectiveness of the proposed model, we conducted experiments related to the segmentation of images (from the Berkeley dataset[1]) composed of multiple regions. Fig. 3 shows six examples of segmentations obtained using fixed and dynamic values of α and β, respectively. For the examples with fixed hyper-parameters, we set $\alpha = \beta = 0.5$. For dynamic hyper-parameters, we use a neighborhood $W(\mathbf{x})$ of size 9×9 pixels. For each tested image, we set the number of regions and initialized manually the contours. For each region, we deform its contour from the initial state until convergence, using the previous algorithm. The final partition of the image is defined by the union of all the contours, plus the region formed by pixels that do not belong to any inside of contour. Finally, since the segmentation is completely data-driven, we avoid overlapping between contours by stopping the evolution of the latter at contacting contour pixels.

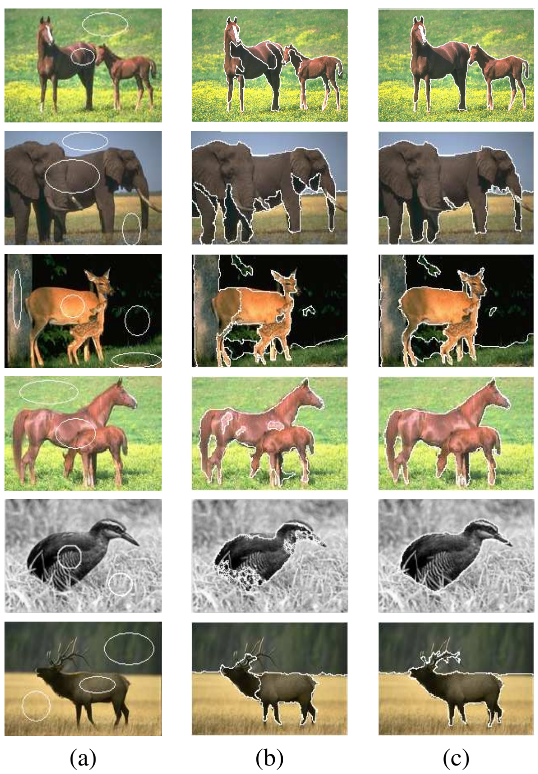

Figure 3. Examples of image segmentation: (a) represents the original image with the initialization, (b) and (c) represent segmentations obtained using fixed and dynamic hyper-parameters, respectively.

We can see in Fig. 3 that dynamic hyper-parameters diminished considerably poor boundary localization caused by self-shadowing or objets specularity. By having fixed hyper-parameters, the contour was stopped by weak edges due to the these phenomena. Furthermore, as shown in Tab. 1, convergence of the contours has been enhanced by having dynamic hyper-parameters, since giving priority to one of the information permits to speed up energy minimization and, therefore, contour motion.

image	fixed α, β	dynamic α, β	gain (%)
1	5.3	4.1	22
2	4.6	3.2	30
3	6.2	4.6	25
4	4.1	3.1	24
5	3.5	2.4	31
6	5.2	3.3	36

Table 1. Values of time (in seconds) taken to perform the segmentations in Fig. (3).

5. Conclusions

In this work, we proposed a principled framework for automatic estimation of hyper-parameters weighting boundary and region information in segmentation. We demonstrated that this approach gives better convergence for segmentation, as well as robustness to over-segmentation due to specularities and shadowing.

References

[1] C. Bishops. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, 1995.

[2] V. Caselles, R. Kimmel, and G. Shapiro. A fast level set method for propagating surfaces. *IJCV*, 22:61–79, 1997.

[3] A. Chakraborty and J. Duncan. Game-theoretic integration for image segmentation. *IEEE PAMI*, 21(1):12–30, 1999.

[4] C. Drewniok. Multispectral edge detection: Some experiments on data from landsat-tm. *IJRS*, 15(18):3743–3765, 1994.

[5] J. Freixenet, X. Munoz, D. Raba, J. Marti, and X. Cuffi. Yet another survey on image segmentation: Region and boundary information integration. *ECCV*, pages 408–422, 2002.

[6] M. Heath. *Scientific Computing*. McGraw-Hil, 2002.

[7] X. Munoz, J. Freixenet, X. Cufi, and J. Marti. Strategies for image segmentation combining region and boundary information. *Patt. Recogn. Lett.*, 24(1-3):408–422, 2003.

[8] N. Paragios and R. Deriche. Unifying boundary and region-based information for geodesic active tracking. *IEEE CVPR*, 2:300–305, 1999.

[1] http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/