Information-theory-based snake adapted to inhomogeneous intensity variations

Frédéric Galland* and Philippe Réfrégier

Physics and Image Processing Group–Fresnel Institute UMR CNRS 6133–École Centrale de Marseille, Université Paul Cézanne–Domaine Universitaire de Saint-Jérôme–13397 Marseille Cedex 20, France

**Corresponding author: frederic.galland@fresnel.fr*

Received June 1, 2007; accepted July 15, 2007;
posted July 27, 2007 (Doc. ID 83422); published August 17, 2007

A new snake-based segmentation technique of a single object (simply connected) in the presence of inhomogeneous Gaussian noise is proposed, in which the mean in each region is modeled as a polynomial function of the coordinates and which is thus adapted to inhomogeneous illumination. It is shown that the minimization of the stochastic complexity of the image, which can be implemented efficiently, allows one to automatically estimate not only the number and the position of the nodes of the polygonal contour used to describe the object but also the degree of the polynomials that model the variations of the mean. © 2007 Optical Society of America

OCIS codes: 100.2000, 100.5010, 110.4280.

Noisy image segmentation is a key problem in image processing. This paper addresses the particular case where the image is composed of a single target (simply connected) lying on a background. The goal is then to retrieve as precisely as possible the contour of this target, i.e., the segmentation task is reduced to a more restricted contour estimation task. Instead of detecting the edges in the image and then linking them to obtain the shape of the object, snakelike approaches [1] generally consist of deforming a curve. One of the main issues of such snake methods is related to the choice of the criterion used to drive the curve to the desirable position [2–4].

In [4], a statistical approach that presents clear optimal properties in the context of statistical estimation theory for a given image model was developed with a polygonal description for the snake. This method has been generalized in [5] and has been implemented using a fast computation scheme. Furthermore, it has been demonstrated that the number of control points for a B-spline representation of the contour [6] or the number of nodes for a polygonal contour [7] can be estimated by minimizing the stochastic complexity of the image [8].

All these statistical techniques rely on the hypothesis of statistically homogeneous regions. The pixel gray levels in each region are thus assumed to be distributed with a probability density function (pdf) whose parameters remain constant all over the region.

In the following, a more general model is considered, involving the possibility of inhomogeneous illuminations. More precisely, the pixel gray levels are assumed to be distributed inside each region as a Gaussian pdf with a constant variance inside each region but whose mean parameter can vary as a polynomial function of the image coordinates. To reduce the computation complexity, at most second-degree polynomials will be considered for the experimental validations (i.e., constant, affine, or quadratic polynomials), even if the theoretical developments are valid whatever the polynomial degree is. Such variations of the mean values, which can notably be caused by nonhomogeneous image enlightenment or when objects with different surface orientations are observed, have already been taken into account, notably in [9] in a minimum description length [8] region merging scheme.

In this paper, a polygonal snake whose convergence relies on the minimization of the stochastic complexity of the image is proposed. This approach allows one to automatically estimate the number of nodes of the polygon and thus to regularize the contour without needing to tune any parameter in the optimized criterion. Moreover, it is shown that the maximum likelihood (ML) estimate of the polynomial coefficients and of the variance can be obtained with a fast algorithm.

Let us consider an image $\mathbf{s}=\{s(x,y)\}$ composed of two regions Ω_r, $r\in\{a,b\}$, where index a denotes the target region and b the background region. In each region Ω_r, the pixel gray levels are assumed to be independent realizations of Gaussian random variables, with a variance σ_r^2 constant over the whole region (but different between the regions). On the contrary, the mean parameter of this Gaussian pdf is assumed not to be constant inside each region but to vary as a degree $d\leqslant 2$ polynomial function of the (x,y) coordinates. The mean parameter in Ω_r is thus approximated with

$$m_r^d(x,y)=\sum_{0\leqslant i+j\leqslant d} a_r^d[i,j]x^iy^j \text{ with } (i,j)\in\mathbb{N}^2, \qquad (1)$$

where $a_r^d[i,j]$ are the polynomial coefficients.

Let $\mathbf{w}=\{w(x,y)\}$ be a two-valued function that denotes a partition of the image so that $w(x,y)=r\Leftrightarrow(x,y)\in\Omega_r$. The stochastic complexity $\Delta(\mathbf{w})$ of the image $\mathbf{s}$ associated with the partition $\mathbf{w}$ is the sum of three terms: the code length $\Delta_L(\mathbf{w})$ of the gray-level description when the partition $\mathbf{w}$ and the parameters of the pdf inside each region are known, the code length $\Delta_P(\mathbf{w})$ of the description of the pdf parameters

0146-9592/07/172514-3/$15.00

inside each region, and the code length $\Delta_C(\mathbf{w})$ of the description of the snake contour, i.e., the partition $\mathbf{w}$.

It has been shown in [6,7] that in the case of a single simply connected object, the approximation for the description of the contour $\Delta_C(\mathbf{w})=k\log N$ can be used, where k is the number of nodes of the snake polygon and N is the pixel number in the image.

According to [8], the code length needed to encode the pdf parameters in region Ω_r is equal to $\alpha(d_r)/2\log N_r$, where N_r is the number of pixels in Ω_r and $\alpha(d_r)$ is the number of parameters of the pdf in Ω_r, when using a polynomial of degree d_r. The pdf in Ω_r depends on the mean parameters $\{a_r^{d_r}[i,j],i+j\leqslant d_r\}$ and on the parameter σ_r, leading to a total number of parameters equal to $\alpha(d_r)=1+[(d_r+1)\times(d_r+2)]/2$ [9]. The total code length to encode the pdf parameters in both Ω_a and Ω_b is thus equal to

$$\Delta_P(\mathbf{w})=\sum_{r\in\{a,b\}}\frac{\alpha(d_r)}{2}\log N_r,\quad \alpha(d_r)=\frac{{d_r}^2+3d_r+4}{2}. \tag{2}$$

The last term $\Delta_L(\mathbf{w})$ is equal to $\Delta_L(\mathbf{w})=-\Sigma_{r\in\{a,b\}}\mathcal{L}(\Omega_r|\theta_r^{d_r})$, where $\mathcal{L}(\Omega_r|\theta_r^{d_r})$ is the log-likelihood of the pixel gray levels inside region Ω_r and where $\theta_r^{d_r}$ is the parameter vector that embedded the parameters σ_r and $\{a_r^{d_r}[i,j]\}$ of the Gaussian pdf in Ω_r, when using a polynomial of degree d_r for the mean. Since $\theta_r^{d_r}$ is *a priori* unknown, it needs to be estimated. Because of good behaviors of ML estimates in the exponential family [10,11], this parameter will be estimated in the ML sense, which is equivalent to minimizing the stochastic complexity. The log-likelihood inside Ω_r is equal to

$$\mathcal{L}(\Omega_r|\theta_r^{d_r})=-\frac{N_r}{2}\log 2\pi\sigma_r^2-\sum_{(x,y)\in\Omega_r}\frac{[s(x,y)-m_r^{d_r}(x,y)]^2}{2\sigma_r^2}, \tag{3}$$

where $m_r^{d_r}(x,y)$ depends on $\{a_r^{d_r}[i,j]\}$. The ML estimates $\hat{a}_r^{d_r}[i_0,j_0]$ of $a_r^{d_r}[i_0,j_0]$ for each couple (i_0,j_0) is obtained from $\partial\mathcal{L}(\Omega_r|\theta_r^{d_r})/\partial a_r^{d_r}[i_0,j_0]=0$, leading to

$$\sum_{i+j\leqslant d_r}S_r[i_0+i,j_0+j,0]\hat{a}_r^{d_r}[i,j]=S_r[i_0,j_0,1], \tag{4}$$

where $S_r[i,j,k]=\Sigma_{(x,y)\in\Omega_r}x^iy^j[s(x,y)]^k$. Since the different terms $S_r[i,j,k]$ are simple summations over Ω_r, each of the polynomial coefficients $\hat{a}_r^{d_r}[i_0,j_0]$ can be obtained by inverting the linear system of Eq. (4).

Moreover, the ML estimate of the variance σ_r^2 in each region is equal to

$$\hat{\sigma}_r^2=\frac{1}{N_r}\Bigg[S_r[0,0,2]-2\sum_{i+j\leqslant d_r}S_r[i,j,1]\hat{a}_r^{d_r}[i,j] + \sum_{i+j\leqslant d_r}\sum_{k+l\leqslant d_r}S_r[i+k,j+l,0]\hat{a}_r^{d_r}[i,j]\hat{a}_r^{d_r}[k,l]\Bigg]. \tag{5}$$

Finally, when substituting these expressions into Eq. (3), the following profile log-likelihood is obtained [4], in which each unknown parameter has been substituted with its ML estimate:

$$\mathcal{L}(\Omega_r|\hat{\theta}_r^{d_r})=-\frac{N_r}{2}\log(2\pi\hat{\sigma}_r^2)-\frac{N_r}{2}. \tag{6}$$

Using the expression of $\Delta_C(w)$ and Eqs. (2) and (6), the stochastic complexity of an image can thus be calculated. It can be shown that using polynomials of degree d, $(5d^2+9d+6)/2$ summations $S_r[i,j,k]$ over the regions have to be performed (including $N_r=S_r[0,0,0]$), i.e., 3 summations when $d=0$, 10 when $d=1$, 22 when $d=2$, and 39 when $d=3$. This increase notably explains why experimental results have been restricted to $d\leqslant 2$. These summations over the target and the background can be very time consuming since they must be re-evaluated after each modification of the partition $\mathbf{w}$. Fortunately, some methods have been proposed to provide fast computation of such summations [5,12]. In our case, the methodology introduced in [5] has been implemented, which consists of replacing these 2D summations over the regions with 1D summations over their contour. In the following, the computation time t (in seconds) obtained on a standard PC with a 3.2 GHz processor will be shown under each segmentation result.

The optimal partition $\mathbf{w}^{opt}$ minimizes the stochastic complexity $\Delta(\mathbf{w})$. In the following, the two-step minimization algorithm proposed in [7] has been implemented. In the first step, the number of nodes is progressively increased while the contour converges to capture all the details of the object. The second step consists of reducing the complexity of the contour to estimate the number of nodes. This two-step technique allows one to estimate the number of nodes and their position for highly nonconvex objects (see [7] for details).

In Fig. 1 (top row), different segmentation results are presented on a gray-scale video image, when imposing a polynomial degree d for the mean for both the target and the background, equal to $d=0$, 1, and 2. The case $d=0$ illustrates the result obtained using a classical Gaussian model [7]. This approach can easily be generalized to multicomponent images, e.g.,

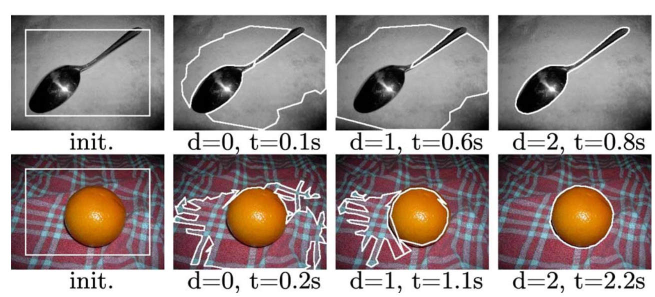

Fig. 1. (Color online) Segmentation of two video images (320×240 pixels): a gray-scale image of a spoon (top row) and a color image of an orange lying on a red fabric (bottom row). Column 1, initial contour; columns 2–4, segmentation results and computation time t using a polynomial for the mean with a degree $d=0$ (column 2), $d=1$ (column 3), and $d=2$ (column 4).

for the color image of Fig. 1 (bottom row). In the case where the three RGB components are supposed to be independent (see, notably, [9] for the correlated case), the pdf of a color pixel is the product of the pixel gray-level pdf of each component. The term Δ_L (Δ_P) for the color image is thus simply the sum of the terms Δ_L (Δ_P) obtained on each component (see [13] for details), whereas the term Δ_C remains unchanged. The segmentation results obtained on this color image are shown in the bottom row of Fig. 1. These two experiments show that in case of nonhomogeneous enlightenment of the scene, constant and affine models for the mean variations can be insufficient: using a quadratic model can thus greatly improve the segmentation results.

In the examples presented in Fig. 1, the degree of the polynomial function for the mean was assumed known and identical in every region. We propose in the following to automatically estimate the polynomial degree $\hat{d}_r$ inside each region Ω_r by minimizing the stochastic complexity of the image, which leads to $\hat{d}_r = \arg\min_{d \in \{0,1,2\}} [\alpha(d)/2 \log N_r - \mathcal{L}(\Omega_r | \hat{\theta}_r^d)]$ and thus

$$\Delta(\mathbf{w}) = \Delta_C(\mathbf{w}) + \sum_r \min_d \left[\frac{\alpha(d)}{2} \log N_r - \mathcal{L}(\Omega_r | \hat{\theta}_r^d) \right]. \tag{7}$$

This new expression of the stochastic complexity no longer requires knowledge of the polynomial degrees.

This approach is illustrated in Fig. 2 with two synthetic images corrupted with Gaussian noise [Figs. 2(a) and 2(c)]. These results confirm the good behavior of the algorithm, notably concerning the estimation of the number and position of the nodes of the polygonal contour and the degrees of the mean polynomials. Indeed, the estimated polynomial degrees correspond to their true value, i.e., $d_a=0$ and $d_b=2$ in Figs. 2(b) and 2(d). The polynomial degrees can also be estimated on the two real images of Fig. 1, as shown in Figs. 3(a) and 3(b), where both the target and the background are estimated to be quadratic polynomials, which is coherent with the results previously obtained in Fig. 1.

In this Letter, a new polygonal snake algorithm has been proposed to deal with noisy images corrupted with Gaussian noise whose mean is inhomogeneous inside the inner or the outer regions of the snake and can vary as a polynomial function of the image coordinates. This segmentation algorithm is based on the minimization of the stochastic complexity, leading to a criterion without tuning parameter, which allows one to automatically estimate the position and the number of nodes of the polygonal contour and the degree of the polynomial function of the mean gray-level values inside each region. The computation of this stochastic complexity involves only simple summations over the regions, which can be implemented with an efficient technique.

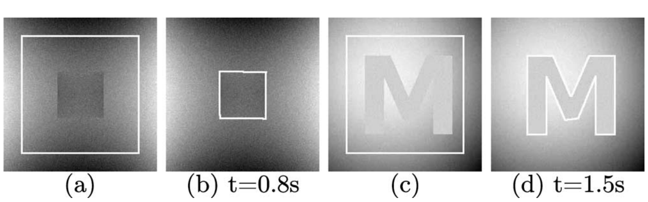

Fig. 2. Segmentation with estimation of the polynomial degree of the mean inside each region. (a), (c) synthetic images (256×256 pixels) corrupted with Gaussian noise ($d_a=0$, $d_b=2$); (b), (d) segmentation results ($\hat{d}_a=0$, $\hat{d}_b=2$). The initial contours are shown in (a), (c).

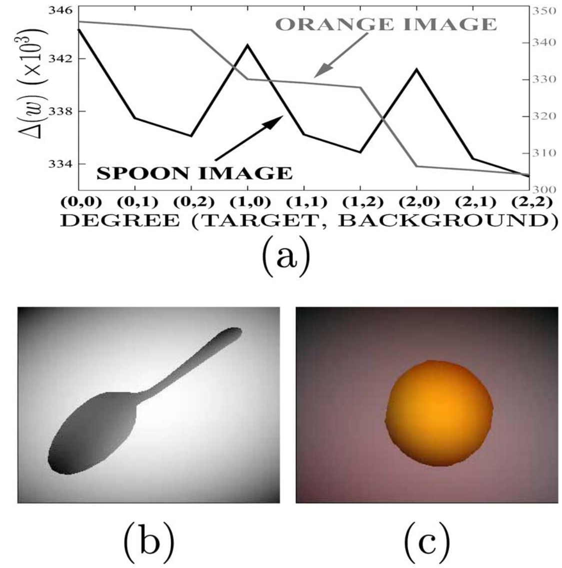

Fig. 3. (Color online) Variation of the stochastic complexity (a) as a function of $(\hat{d}_a, \hat{d}_b)$ on the two real images of Fig. 1, with a minimum for $(\hat{d}_a, \hat{d}_b)=(2,2)$. (b), (c) reconstructions of the corresponding estimated means.

References

1. M. Kass, A. Witkin, and D. Terzopoulos, Int. J. Comput. Vis. **1**, 321 (1988).
2. C. Xu and J. L. Prince, IEEE Trans. Image Process. **7**, 359 (1998).
3. R. Ronfard, Int. J. Comput. Vis. **2**, 229 (1994).
4. O. Germain and P. Réfrégier, Opt. Lett. **21**, 1845 (1996).
5. C. Chesnaud, P. Réfrégier, and V. Boulet, IEEE Trans. Pattern Anal. Mach. Intell. **21**, 1145 (1999).
6. M. Figueiredo, J. Leitão, and A. K. Jain, IEEE Trans. Image Process. **9**, 1075 (2000).
7. O. Ruch and P. Réfrégier, Opt. Lett. **26**, 977 (2001).
8. J. Rissanen, *Stochastic Complexity in Statistical Inquiry*, Vol. 15 of Series in Computer Science (World Scientific, 1989).
9. T. Kanungo, B. Dom, W. Niblack, and D. Steele, in *Proceedings of Computer Vision and Pattern Recognition 1994* (IEEE, 1994), p. 609.
10. T. S. Ferguson, *Mathematical Statistics, a Decision Theoretic Approach* (Academic, 1967).
11. P. Garthwaite, I. Jolliffe, and B. Jones, *Statistical Inference* (Prentice-Hall Europe, 1995).
12. P. Viola and M. Jones, *Proceedings of 13th Conference on Computer Vision and Pattern Recognition* (IEEE, 2001), p. 511.
13. F. Galland, N. Bertaux, and P. Réfrégier, Pattern Recogn. **38**, 1926 (2005).