Minimal-complexity segmentation with a polygonal snake adapted to different optical noise models

Olivier Ruch* and Philippe Réfrégier

Physics and Image Processing Group, Fresnel Institute, Ecole Nationale Supérieure de Physique de Marseille, Domaine Universitaire de Saint-Jérôme, 13 397 Marseille Cedex 20, France

Received March 19, 2001

Polygonal active contours (snakes) have been used with success for target segmentation and tracking. We propose to adapt a technique based on the minimum description length principle to estimate the complexity (proportional to the number of nodes) of the polygon used for the segmentation. We demonstrate that, provided that an up-and-down multiresolution strategy is implemented, it is possible to estimate efficiently this number of nodes without *a priori* knowledge and with a fast algorithm, leading to a segmentation criterion without free parameters. We also show that, for polygonal-shaped objects, this new technique leads to better results than using a simple regularization strategy based on the smoothness of the contour.

OCIS codes: 100.0100, 100.2960, 100.5010.

Image segmentation and object-shape estimation are among the most challenging problems for pattern recognition, in particular in new optical and microwave imagery systems for which the level of noise can be important. A new approach to segmenting objects in images corrupted by different types of noise present in optical and radar imagery systems, called a statistically independent region snake, was recently proposed.[1–3] In particular, the efficiency of this approach has been demonstrated with images corrupted with speckle noise, Poisson noise, or Gaussian noise (see Ref. 3 for a review). This region-based approach is analogous to other recently reported techniques involving active contours[4,5] (snakes) but presents clear optimal statistical properties. To determine the shape of the object that is to be segmented, one uses a polygonal description of the contour. This contour model allows one not only to estimate efficiently the shapes of manufactured objects but also to obtain a very fast technique, since segmentation can be performed in few hundred milliseconds. However, as with all the techniques based on a parametric contour model, it is necessary to specify the complexity of the contour model, which is equivalent, in our particular case, to the number of nodes of the polygon.

In this Letter we propose to adapt a technique based on the minimum description length principle to estimate this polygon complexity. A classical difficulty with snake-based approaches is their sensitivity to the initial shape used for the segmentation. In Ref. 3, a simple multiresolution method in which the number of nodes is progressively increased was thus introduced. This approach improves both the robustness of the snake against initialization and its convergence properties. We demonstrate here that, provided that an original approach is implemented, it is possible to estimate efficiently the number of nodes needed to describe the shape of the segmented objects with a fast technique. This approach thus leads to a segmentation criterion without free parameters. We also show that, for polygonal objects, this new technique leads to better results than using a regularization strategy based on the smoothness of the contour and that the technique is still robust if the object's shape is not a simple polygon.

In the following mathematical developments, one-dimensional notation is used for simplicity, and bold symbols denote N-dimensional vectors. Let us consider a scene $\mathbf{s} = \{s_i \,|\, i \in [1, N]\}$ composed of N pixels. The gray levels of the target, $\mathbf{a}$, and of the background, $\mathbf{b}$, are considered independent random vectors with statistically independent components distributed with probability density functions (PDF) $P_{\theta_a}(i)$ and $P_{\theta_b}(i)$, respectively. θ_a and θ_b are the statistical parameters that must differ between the two regions. Let $\mathbf{w} = \{w_i \,|\, i \in [1, N]\}$ denote a binary window function that defines a certain shape for the target, so that $w_i = 1$ within this shape and $w_i = 0$ elsewhere. The image is thus divided into two regions: $\Omega_a = \{i \in [1, N] \,|\, w_i = 1\}$ and $\Omega_b = \{i \in [1, N] \,|\, w_i = 0\}$.

The purpose of segmentation is therefore to estimate the shape, $\mathbf{w}$, of the target in the scene. The k-node polygonal snake defines the boundary of the shape, and $\mathbf{w}$ is then a polygon-bounded support function, with one value on and within the snake and zero value elsewhere. In the approaches reported in Refs. 1–3, the number of nodes, k, was arbitrarily chosen, and the estimation of target shape $\mathbf{w}$ was performed by maximization of the likelihood function determined from image $\mathbf{s}$ and a hypothesis for the shape, $\mathbf{w}$. To estimate the number of nodes, k, we propose the use of an approach analogous to the one developed in Refs. 6 and 7 and that consists of minimizing the length Δ of the description of the image [this approach is well known as the minimum description length (MDL) principle introduced by Rissanen; see Ref. 8, for example]. Since the image is divided into three parts (the target, the background, and the contour), Δ is the sum of three terms: the length Δ_a of the description of the target gray levels, the length Δ_b of the description of

0146-9592/01/130977-03$15.00/0

the background gray levels, and the length Δ_w of the description of polygon $\mathbf{w}$. Let us first provide an approximation of Δ_w. The number of possible locations for one node is N. Thus, for k nodes the number of different sets of locations is N^k, and we consider N^k an approximation of the number of different polygons. The number of bits necessary to describe the polygon (if all polygons are assumed to be equally likely) is thus approximately $\log_2(N^k)$, where $\log_2$ is the base 2 logarithm, which can be considered the complexity of the polygon. It has been well known since the work of Shannon[9] that the average number of bits needed to describe N_l random variables distributed with PDF $P_{\theta_l}(x)$ is $\Delta_l \simeq N_l S_l$, where S_l is the entropy of the PDF and is given by $S_l = -\int P_{\theta_l}(x) \log_2[P_{\theta_l}(x)]\mathrm{d}x - \log_2(q)$ and q is the quantization precision. Since the contribution of $\log_2(q)$ will consist only of adding a constant term to the description length, it will not be taken into account in the following. The average number of bits needed to describe the object region (the background region) is thus $\Delta_a \simeq N_a S_a$ ($\Delta_b \simeq N_b S_b$). So the total description length is then

$$\Delta \simeq N_a S_a + N_b S_b + k \log_2(N). \qquad (1)$$

It is well known that if entropy is approximated by use of the empirical mean instead of the statistical average, one obtains $N_l S_l \simeq -\sum_{i\in\Omega_l} \log_2[P_{\theta_l}(i)]$, where $l = a$ or b. It is thus easy to see that $-N_a S_a - N_b S_b$ is the base 2 log likelihood $l_2[\mathbf{s}\,|\,\mathbf{w}, \theta_a, \theta_b]$ of the hypothesis that, in image $\mathbf{s}$, the shape of the target is $\mathbf{w}$. This is nothing but the statistical criterion that was optimized in Ref. 3. One can thus see that the MDL principle leads to the minimization

$$\Delta' \simeq -l_e[\mathbf{s}\,|\,\mathbf{w}, \theta_a, \theta_b] + k \ln(N), \qquad (2)$$

where ln is the natural logarithm and $l_e[.]$ denotes the natural logarithm likelihood.

When the statistical parameters θ_a, θ_b are unknown, we adopt the same approach as in Ref. 3, which consists of considering their maximum likelihood estimates $\hat{\theta}_a(\mathbf{s}, \mathbf{w})$, $\hat{\theta}_b(\mathbf{s}, \mathbf{w})$. The pseudolikelihood is thus $l[\mathbf{s}\,|\,\mathbf{w}, \hat{\theta}_a(\mathbf{s}, \mathbf{w}), \hat{\theta}_b(\mathbf{s}, \mathbf{w})]$, and in Table 1 we list its expressions for different PDFs (see also Ref. 3).

The shape estimate $\hat{\mathbf{w}}$ is thus obtained by simultaneous determination of the values of k and $\mathbf{w}$ that minimize Δ'. This double-optimization problem is not trivial, and the adopted strategy may have a strong influence on the quality and relevance of the application of the MDL principle. The simplest strategy consists of determining the shapes $\hat{\mathbf{w}}^{(k)}$ that optimize the criterion Δ' for different fixed values of k and then selecting the value of k that leads to the minimal value of Δ'.

Since, in our approach, a polygonal description is used and the range of possible values of k is large (typically 4–200), this method appears to be inefficient. Let us illustrate this point with a synthetic image of a boat (whose shape is polygonal with 10 nodes) corrupted with speckle noise of order 1 [see Figs. 1(a) and 1(b)]. Figure 1(c) presents the shape obtained after the convergence of the snake with the true number of nodes ($k = 10$), when the multiresolution strategy described in Ref. 3 is used. One can see that the segmentation is not efficient because the contour does not describe all the details of the shape. The solid curve in Fig. 2 shows the evolution of the value of Δ' as k increases when the multiresolution strategy is used. This result shows that one cannot directly apply this simple strategy.

To overcome this problem we propose a new approach: an up-and-down multiresolution strategy with two basic steps. Since efficient convergence is obtained when the number of nodes is large, we first perform a segmentation with a multiresolution strategy, by increasing the number of nodes to a point at which the distance between two consecutive nodes does not exceed a small value (typically one or two pixels).[2] Thus we typically end up with an overestimated number of nodes, k_0. The second step is a complexity reduction technique and consists of pruning the contour. For this purpose we sequentially consider each node of the contour, and we determine the value of the criterion Δ' obtained if this node is removed. Then the pruned contour is defined by suppression of the node that leads to the minimal value of Δ', and a new convergence of the snake is

Table 1. Expressions of the Pseudolikelihood $l[\mathbf{s}\,|\,\mathbf{w}, \hat{\theta}_a(\mathbf{s}, \mathbf{w}), \hat{\theta}_b(\mathbf{s}, \mathbf{w})]$ for Different Gray-Level PDFs[a]

Gamma	$N_a \log \hat{m}_a + N_b \log \hat{m}_b + A$
Gaussian	$N_a \log \hat{\sigma}_a + N_b \log \hat{\sigma}_b + B$
Poisson	$-N_a\hat{m}_a \log \hat{m}_a - N_b\hat{m}_b \log \hat{m}_b + C$

[a] N_l is the number of pixels in region Ω_l, $\hat{m}_l = (1/N_l \sum_{i\in\Omega_l} s_i)$, and $\hat{\sigma}_l{}^2 = (1/N_l \sum_{i\in\Omega_l} s_i{}^2 - \{1/N_l \sum_{i\in\Omega_l} s_i\}^2)$ are the maximum-likelihood estimates of m_l and $\sigma_l{}^2$, where $l = a$ or b. A, B, and C are independent of N_l, $\hat{m}_l$, and $\hat{\sigma}_l{}^2$ ($l = a$ or b).

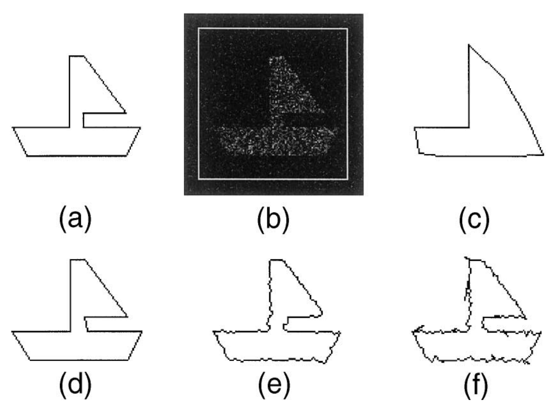


Fig. 1. Scene with speckle noise. (a) Boat whose shape is a polygon with 10 nodes in a 128 by 128 pixel image. (b) Speckled image of (a) with a contrast equal to 4 (initialization of the contour in white). (c) Final state of the snake after optimization of the log likelihood with the multiresolution strategy when the number of nodes on the contour is equal to the true one: $k = 10$. (d) Final state of the snake after optimization of the MDL criterion with the up-and-down multiresolution strategy. (e) Final state of the snake after optimization of the log likelihood with a regularizing term and $\alpha = 0.2$ (a value that leads to the minimal number of misclassified pixels). (f) Final state of the snake after optimization of the log likelihood without a regularizing term.

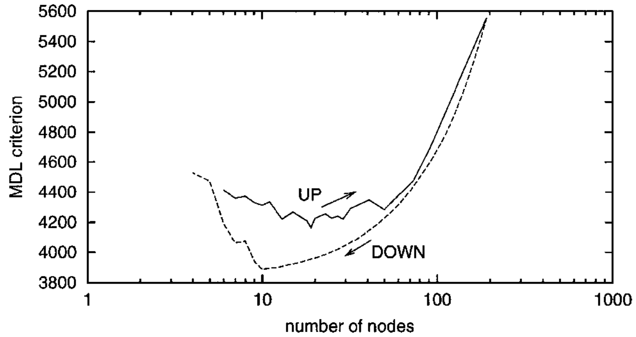
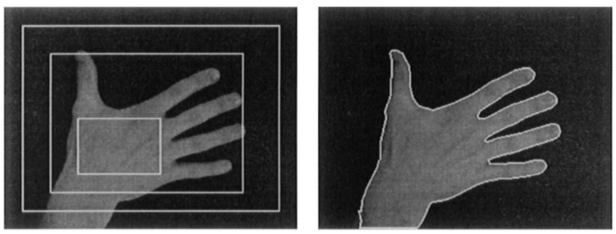


Fig. 2. Solid curve, values of the MDL criterion Δ' obtained with a direct approach. Dashed curve, values of the MDL criterion Δ' obtained with the up-and-down multiresolution strategy.



(a) (b)

Fig. 3. (a) Real image of a hand (240×320 pixel image) and initialization of the contour. (b) Result of the segmentation with the up-and-down multiresolution strategy.

performed on this pruned contour. The process is continued until the number of nodes is smaller than a limit value (typically four or five).

In Fig. 2 we show the value of Δ' obtained with this procedure for decreasing values of k starting from k_0. One can observe that this technique is very efficient for the selected test image, since one obtains the true number of nodes ($k = 10$), although k varies in a large domain (from 192 to 5). We show in Fig. 1(d) the polygonal snake obtained with the MDL segmentation of an image of a boat. In Fig. 1(e) we compare this result with the solution obtained when a regularizing term $\alpha U_{\mathrm{reg}}(\mathbf{w})$ as described in Ref. 3 is added to the log likelihood and when the distance between two consecutive nodes is two pixels. The value of the weighted coefficient α was chosen so that the number of pixels misclassified between the segmented contour and the true contour is minimal. The solution when no regularizing term is added but the conditions are the same is shown in Fig. 1(f). These results illustrate the efficiency of the up-and-down multiresolution strategy. It is worth noting that this strategy has also the advantage that it leads to a fast algorithm. For example, the total computing time (with a typical PC-700 MHz) is less than 3 s in the example shown in Fig. 1, for which the distance between two consecutive nodes at the end of the first step is two pixels. This computing time can be reduced to less than 500 ms, with no loss in the quality of the contour, if this distance is set equal to four pixels. Analogous results were obtained from images corrupted with Poisson or Gaussian noise. In Fig. 3 we show the results obtained with an image of a real hand when a Gaussian noise model is used.

The authors acknowledge the support of Thales Optronique for this work and thank Henri Maitre for fruitful discussions and for his encouragement to analyze MDL techniques. P. Réfrégier's e-mail address is philippe.refregier@fresnel.fr.

*Also with Thales Optronique SA, Rue Guynemer, BP 55, 78283 GuyanCourt Cedex, France.

References

1. O. Germain and Ph. Réfrégier, "Optimal snake-based segmentation of a random luminance target on a spatially disjoint background," Opt. Lett. **21,** 1845–1847 (1996).
2. C. Chesnaud, V. Pagé, and Ph. Réfrégier, "Robustness improvement of the statistically independent region snake-based segmentation method," Opt. Lett. **23,** 488–490 (1998).
3. C. Chesnaud, Ph. Réfrégier, and V. Boulet, "Statistical region snake-based segmentation adapted to different physical noise models," IEEE Trans. Pattern Anal. Mach. Intell. **21,** 1145–1157 (1999).
4. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," Int. J. Comput. Vision **1,** 321–331 (1988).
5. R. Ronfard, "Region-based strategies for active contour models," Int. J. Comput. Vision **2,** 229–251 (1994).
6. S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," IEEE Trans. Pattern Anal. Mach. Intell. 874–900 (1996).
7. M. A. T. Figueiredo, J. M. N. Leitão, and A. K. Jain, "Unsupervised contour representation and estimation using B-splines and a minimum description length criterion," IEEE Trans. Image Process. **9,** 1075–1087 (2000).
8. J. Rissanen, *Stochastic Complexity in Statistical Inquiry* (World Scientific, Singapore, 1989).
9. C. E. Shannon, "A mathematical theory of communication," Bell Syst. Tech. J. **27,** 379–423 (1948).