UNSUPERVISED CONTOUR ESTIMATION

Mário A. T. Figueiredo, and José M. N. Leitão

Instituto de Telecomunicações, and
Departamento de Engenharia Electrotécnica e de Computadores.
Instituto Superior Técnico,
1096 Lisboa Codex, PORTUGAL

ABSTRACT

We introduce a fully adaptive active contour model in which no parameters have to be set *a priori* or tuned by the user. It is based on elliptic Fourier contour description and on the *minimum description length* (MDL) principle. The proposed technique estimates all the observation model parameters (e.g., noise variances), the order of the contour description (number of Fourier coefficients), and the contour itself.

1. INTRODUCTION AND PREVIOUS WORK

Contour estimation is one of the most important, interesting, and challenging problems in image processing and computer vision. Originating from the seminal work of Kass, Witkin, and Terzopoulos [11], *snake*-type approaches (in which we may include deformable contours, active contours, dynamic contours, and deformable templates) constitute one of the most successful approaches.

In the original version [11], snakes work by minimizing an *energy* function composed of an (internal) elastic-type term which increases with the contour deformations, and an (external) attraction potential linking the contour with the image. The goal is a compromise between contour *smoothness* and adequacy to the observed data. In recent years, several improvements, modifications, and reformulations have overcome limitations of the traditional model such as sensitivity to initialization, *myopia* (i.e., insensitivity to distant features), and inability to reparametrize itself during the deformation process (see [1], [2], [8], [12], [13], [15], [18], and references therein).

Deformable templates, though related to *snake*-type approaches, use shape descriptions with small numbers of parameters; the deformation energy term is usually unnecessary (see [3], [10], [19] and references therein).

From a Bayesian estimation angle, deformable models are interpretable as *maximum a posteriori* (MAP) estimators; the internal energy and the external potential terms are associated with the *a priori* probability function and the likelihood function, respectively; for details, see [6], [10], [19], [20]. The Bayesian estimation perspective has the advantage of giving meaning to all the involved entities; e.g., the form of the energy term that links the contour with the image contents, i.e. the likelihood function (in Bayesian terms) can be derived from knowledge about the observation model rather than simply from common sense arguments [5], [6]. The main difficulty in this approach is the choice of the parameters involved in the definition of the *a priori* probability function and of the observation model (e.g., noise variances). In [6], we have proposed an adaptive Bayesian approach for a ventricular contour estimation problem. In [5], a technique which adaptively estimates the observation model parameters was proposed.

In this paper, we introduce a new fully adaptive contour estimation technique in which no parameters have to be set *a priori* or tuned by the user. It is based on Fourier contour description and on the *minimum description length* (MDL) principle [16], [17]. In the proposed criterion, the contour shapes are described by as few parameters as possible. Basically, it is a template matching approach with adaptive parametrization.

2. PROPOSED TECHNIQUE

2.1. Fourier Representation

Let a closed curve (i.e., a closed contour) on the image plane be represented by a periodic vector function $\mathbf{v}(t) = [x(t)\ y(t)]^T$, of period 2π, i.e., of unit fundamental angular frequency. The complex Fourier series description of the closed curve is defined as

$$\mathbf{v}(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \sum_{k=-\infty}^{\infty} \begin{bmatrix} c_k \\ d_k \end{bmatrix} e^{j\,k\,t}, \quad t \in [0,\ 2\pi[, \tag{1}$$

0-7803-3258-X/96/$5.00 © 1996 IEEE

where

$$\begin{bmatrix} c_k \\ d_k \end{bmatrix} = \frac{1}{2\pi} \int_0^{2\pi} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} e^{-jkt} dt \qquad (2)$$

are the complex Fourier coefficients [9], [19]. The discrete version of this representation is obtained by considering a discretization of a period of the curve $\mathbf{v}(t)$ into N points $\{\mathbf{v}_i,\ i = 0, 1, ..., N-1\}$. The discrete complex Fourier series representation is

$$\mathbf{v}_i \equiv \begin{bmatrix} x(i2\pi/(N-1)) \\ y(i2\pi/(N-1)) \end{bmatrix} = \sum_{k=0}^{N-1} \begin{bmatrix} e_k \\ f_k \end{bmatrix} e^{jki2\pi/N} \qquad (3)$$

with

$$\begin{bmatrix} e_k \\ f_k \end{bmatrix} \equiv \mathbf{g}_k = \frac{1}{N} \sum_{i=0}^{N-1} \begin{bmatrix} x(i2\pi/(N-1)) \\ y(i2\pi/(N-1)) \end{bmatrix} e^{-jki2\pi/N}. \qquad (4)$$

By truncating series (3) to K terms (with $K < N$), a smoothed version of the curve is obtained. We denote the vector of $2K$ complex coefficients by

$$\boldsymbol{\theta}_{(K)} = [e_0, f_0, e_1, f_1, ..., e_{K-1}, f_{K-1}]. \qquad (5)$$

The true unobserved contour is assumed to be *smooth*, i.e., it can be exactly described by some $\boldsymbol{\theta}_{(K)}$.

2.2. The Likelihood Function

Given an N-point contour defined by a K-order discrete complex Fourier series, the observed image $\mathbf{I}$ is modeled by the likelihood function $p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi})$, where $\boldsymbol{\phi}$ is the vector of parameters of the observation mechanism. Although this is an often overlooked aspect, great care has to be put on the derivation of the likelihood function. For specific applications (e.g., finding organ boundaries in medical images), all the available knowledge about the image acquisition process should be included [5], [6]. Not doing so may result in disastrous results, specially on very low quality images (see [5]). This approach naturally leads to region based strategies which, unlike gradient based ones, are robust in the presence of noisy or low contrast images [5], [6], [8], [18].

We now make the following assumptions:

Conditional independence: given the true (unobserved) contour, the image pixels are independently distributed.

Inside and outside regions homogeneity: The conditional probability function of each pixel depends only on whether it belongs to the inside or outside region of the contour; i.e., all the pixels inside (resp. outside) have a common distribution characterized by a parameter vector $\boldsymbol{\phi}_{\text{in}}$ (resp. by $\boldsymbol{\phi}_{\text{out}}$), with $\boldsymbol{\phi} = [\boldsymbol{\phi}_{\text{in}}, \boldsymbol{\phi}_{\text{out}}]$.

Accordingly,

$$p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) = \left(\prod_{(i,j) \in \mathcal{I}(\boldsymbol{\theta}_{(K)})} p(I_{(i,j)}|\boldsymbol{\phi}_{\text{in}}) \right) \left(\prod_{(i,j) \in \mathcal{O}(\boldsymbol{\theta}_{(K)})} p(I_{(i,j)}|\boldsymbol{\phi}_{\text{out}}) \right)$$

where $\mathcal{I}(\boldsymbol{\theta}_{(K)})$ and $\mathcal{O}(\boldsymbol{\theta}_{(K)})$ are the inside and outside regions of the contour defined by $\boldsymbol{\theta}_{(K)}$, respectively; likewise, $p(I_{(i,j)}|\boldsymbol{\phi}_{\text{in}})$ and $p(I_{(i,j)}|\boldsymbol{\phi}_{\text{out}})$ are the pixel-wise conditional probabilities, of the inner and outer regions, respectively.

2.3. The Estimation Criterion

If K was known, maximum likelihood (ML) estimates could in principle be obtained by maximizing $p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi})$ with respect to $\boldsymbol{\theta}_{(K)}$ and $\boldsymbol{\phi}$. However, since K is unknown, there is a model order problem which can be stated as (assuming, for simplicity, known $\boldsymbol{\phi}$):

- for each K, there is a parameter space $\Theta_{(K)} \ni \boldsymbol{\theta}_{(K)}$, of dimension K;
- the spaces are *nested*, that is, for each $\boldsymbol{\theta}_{(K)} \in \Theta_{(K)}$, there is some $\boldsymbol{\theta}'_{(K+1)} \in \Theta_{(K+1)}$ such that[1]

$$p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) = p(\mathbf{I}|\boldsymbol{\theta}'_{(K+1)}, \boldsymbol{\phi}); \qquad (6)$$

- consequently, K can not be estimated directly by maximizing the likelihood function since $p(\mathbf{I}|\widehat{\boldsymbol{\theta}}_{(K)}, \boldsymbol{\phi})$, where $\widehat{\boldsymbol{\theta}}_{(K)}$ is the ML estimate of $\boldsymbol{\theta}_{(K)}$ given K, is a non-decreasing function of K [14].

To overcome this difficulty, we adopt Rissanen's MDL principle (see [16] or [17] for details) where

$$\left(\widehat{\boldsymbol{\theta}_{(K)}}, \widehat{\boldsymbol{\phi}}\right) = \arg\min \left\{ -\log p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) + 2K \log N \right\}, \qquad (7)$$

with $\widehat{\boldsymbol{\theta}_{(K)}}$ standing for the joint estimates[2] of K and $\boldsymbol{\theta}_{(K)}$. The $2K \log N$ term results from the fact that a K-order parametrization involves $2K$ complex coefficients, i.e. $4K$ real ones, and the MDL principle penalizes each real coefficient with a $(\log N)/2$ cost [17].

[1] In our case, $\boldsymbol{\theta}_{(K)} = [e_0, f_0, e_1, f_1, .., e_{K-1}, f_{K-1}]$ and $\boldsymbol{\theta}'_{(K+1)} = [e_0, f_0, e_1, f_1, .., e_{K-1}, f_{K-1}, 0, 0]$ describe the same contour.

[2] Notice the different notation to be used throughout the paper: $\widehat{\boldsymbol{\theta}_{(K)}}$ denotes joint estimates of both K and $\boldsymbol{\theta}_{(K)}$, while $\widehat{\boldsymbol{\theta}}_{(K)}$ denotes an estimate of $\boldsymbol{\theta}_{(K)}$ for a given K.

From a Bayesian point of view, (7) can be interpreted as a MAP estimator,

$$\begin{aligned}\left(\widehat{\theta_{(K)}}, \widehat{\phi}\right) &= \arg\max\left\{p\left(\theta_{(K)}, \phi\,|I\right)\right\} \\ &= \arg\max\left\{\log p(\mathbf{I}|\theta_{(K)}, \phi) + \right. \\ &\qquad \left. \log p\left(\theta_{(K)}, \phi\right)\right\},\end{aligned}$$

with the prior $p(\theta_{(K)}, \phi) \propto \exp\{-2K \log N\}$. Since K is the number of terms in the Fourier description of the contour, this is basically a smoothing prior (as in regularization) expressed in the Fourier domain; this has the advantage of avoiding the shrinkage associated with smoothing priors directly expressed on the contour coordinates [19]. Finally, we stress that the estimator (7) does not require the previous specification of parameters, thus being fully unsupervised.

2.4. Implementing the Estimation Criterion

2.4.1. Introduction

To deal with the difficult task of solving (7) we developed an iterative scheme which can be seen as related to the *expectation-maximization* (EM) algorithm of Dempster *et al* [4] and to the auxiliary variables methods described by Cohen [3].

Given some $\theta_{(K)}$, let $\mathbf{v}(\theta_{(K)}) = [\mathbf{v}_0, \mathbf{v}_1, ...\mathbf{v}_{N-1}]$ be an explicit contour description obtained from the Fourier series (3); of course, $\mathbf{v}(\theta_{(K)})$ contains the same information as $\theta_{(K)}$ itself. Then, writing $p(\mathbf{I}|\theta_{(K)}, \phi)$ is equivalent to writing $p(\mathbf{I}|\mathbf{v}(\theta_{(K)}), \phi)$.

Now, let $\mathbf{r}$ be a noisy version of $\mathbf{v}(\theta_{(K)})$; specifically, $\mathbf{r} = \mathbf{v}(\theta_{(K)}) + \mathbf{n}$, where the elements of $\mathbf{n}$ are all independent and identically distributed zero mean Gaussian variables of variance σ^2. Assume that the observed image is not a function of the true contour $\mathbf{v}(\theta_{(K)})$, but rather of the noisy contour $\mathbf{r}$. As in *incomplete data* problems, we consider $\mathbf{r} = [\mathbf{r}_0, \mathbf{r}_1, ..., \mathbf{r}_{N-1}]$ as *missing data* and write

$$\left(\widehat{\theta_{(K)}}, \widehat{\phi}\right) = \arg\min\{-\log p\left(\mathbf{I}, \mathbf{r}|\theta_{(K)}, \phi\right) + 2K \log N\} \tag{8}$$

where the *complete data* $(\mathbf{I}, \mathbf{r})$ has an *observed* part $\mathbf{I}$ and a *missing* part $\mathbf{r}$ [4]. Here, $\mathbf{r}$ plays the role of the auxiliary variables used in Cohen's methods [3]. Invoking Bayes law we can write

$$\begin{aligned}p\left(\mathbf{I}, \mathbf{r}|\theta_{(K)}, \phi\right) &= p\left(\mathbf{I}|\mathbf{r}, \theta_{(K)}, \phi\right) p\left(\mathbf{r}|\theta_{(K)}, \phi\right) \\ &= p\left(\mathbf{I}|\mathbf{r}, \phi\right) p\left(\mathbf{r}|\theta_{(K)}\right)\end{aligned} \tag{9}$$

since, given the contour $\mathbf{r}$, the observed image $\mathbf{I}$ does not depend on $\theta_{(K)}$, and, given $\theta_{(K)}$, the contour $\mathbf{r}$ does not depend of ϕ.

Notice that, if K was known, we would be facing a typical *missing data* problem [4]; i.e., in the presence of both $\mathbf{I}$ and $\mathbf{r}$, ML estimates of $\theta_{(K)}$ and ϕ are easy to obtain:

- given $\mathbf{r}$, and since $\mathbf{r} = \mathbf{v}(\theta_{(K)}) + \mathbf{n}$, the ML estimate $\widehat{\theta}_{(K)}$ is simply obtained by computing the K first Fourier series coefficients of $\mathbf{r}$; this is true independently of the value of σ^2 [14];
- given $\mathbf{I}$ and $\mathbf{r}$, ML estimates of ϕ_{in} and ϕ_{out} are

$$\widehat{\phi_{\text{in}}} = \arg\max_{\phi_{\text{in}}} \prod_{(i,j)\in\mathcal{I}(\mathbf{r})} p(I_{(i,j)}|\phi_{\text{in}}) \tag{10}$$

$$\widehat{\phi_{\text{out}}} = \arg\max_{\phi_{\text{in}}} \prod_{(i,j)\in\mathcal{O}(\mathbf{r})} p(I_{(i,j)}|\phi_{\text{out}}) \tag{11}$$

  where $\mathcal{I}(\mathbf{r})$ and $\mathcal{O}(\mathbf{r})$ denote the inside and outside regions of the contour $\mathbf{r}$, respectively.

However, $\mathbf{r}$ is not observed, it is *missing*; moreover, K is also unknown. Criterion (8) can then be classified as an MDL principle with incomplete data. In [7], we have proposed a similar criterion for adaptive discontinuity-preserving image restoration.

2.4.2. Algorithms

Assuming temporarily that K is known, the problem reduces to

$$\left(\widehat{\theta}_{(K)}, \widehat{\phi}\right) = \arg\min\{-\log p\left(\mathbf{I}|\mathbf{v}, \phi\right) - \log p\left(\mathbf{v}|\theta_{(K)}\right)\} \tag{12}$$

Let $\widehat{\mathbf{r}}^{(t)}$, $\widehat{\theta}_{(K)}^{(t)}$, and $\widehat{\phi}^{(t)}$ denote estimates in iteration t; our iterative scheme proceeds as follows:

Fixed-K Algorithm

Step 0 (Initialization): Let $t = 0$ and assume some initial contour estimate $\widehat{\mathbf{r}}^{(0)}$.

Step 1: From $\widehat{\mathbf{r}}^{(t)}$, the ML estimate $\widehat{\phi}^{(t)}$, which is given by (10) and (11), is computed.

Step 2: From the current contour estimate $\widehat{\mathbf{r}}^{(t)}$, the ML estimate $\widehat{\theta}_{(K)}^{(t)}$ is obtained by computing the K first complex Fourier series coefficients of $\widehat{\mathbf{r}}^{(t)}$.

Step 3: Given $\widehat{\theta}_{(K)}^{(t)}$, an intermediate contour is obtained according to:

$$\mathbf{v} = \mathbf{v}(\widehat{\theta}_{(K)}^{(t)}). \tag{13}$$

Step 4: Each point of $\mathbf{v}$ is moved in the direction of ascent of $\log p(\mathbf{I}|\mathbf{v}, \widehat{\phi}^{(t)})$. To avoid instability, this ascent step should be small [3]; typically,

$$\mathbf{r}^{(t+1)} = \mathbf{v} + \alpha \nabla_{\mathbf{v}} \log p(\mathbf{I}|\mathbf{v}, \widehat{\phi}^{(t)}). \qquad (14)$$

where α is a small constant and $\nabla_{\mathbf{v}}$ denotes an approximate gradient obtained by local differences.

Step 5: If $|\widehat{\theta}_{(K)}^{(t)} - \widehat{\theta}_{(K)}^{(t)}| < \varepsilon_K$ the algorithm stops; otherwise, t is incremented and we go back to **Step 1**. Here, ε_K is a suitable stopping threshold.

Notice that steps 1 and 2 perform a projection of the current contour estimate onto the subspace of contours with only K Fourier coefficients different from zero. Since a subspace is a convex set, this scheme is also related to the *projection onto convex sets* (POCS) [21].

When K is unknown, which is the general case, the previous algorithm is inserted into an outer loop which sweeps a range of values $\{1, 2, ..., K_{\max}\}$.

Unknown-K Algorithm

Step 1: For each $K \in \{1, 2, ..., K_{\max}\}$, run **Fixed-K Algorithm** and store the final estimates $\widehat{\theta}_{(K)}$ and $\widehat{\phi}$

Step 2: For each pair $\widehat{\theta}_{(K)}$, $\widehat{\phi}$ compute

$$-\log p(\mathbf{I}|\widehat{\theta}_{(K)}, \widehat{\phi}) + 2K \log N \qquad (15)$$

and choose the one which yields the lowest value.

3. FINAL NOTE

For lack of space, experimental results can not be shown here. These will be included in an expanded forthcoming paper.

4. REFERENCES

[1] V. Caselles, R. Kimmel, and G. Sapiro. "Geodesic active contours". *in Proc. of the Intern. Conf. on Comp. Vis. – ICCV'95*, pp. 694–699, Cambridge, 1995.

[2] L. Cohen. "On active contour models and baloons". *Comp. Vision, Graphics, and Image Proc. (CVGIP): Image Understanding*, vol. 53, pp. 211–218, 1991.

[3] L. Cohen. "Auxiliary variables and two-step iterative algorithms in computer vision problems". *Intern. Jour. of Comp. Vis.*, vol. 6, pp. 59–83, 1996.

[4] A. Dempster, N. Laird, and D. Rubin. "Maximum likelihood estimation from incomplete data via the EM algorithm". *Jour. of the Royal Statist. Soc. B*, vol. 39, pp. 1–38, 1977.

[5] J. Dias and J. Leitão. "Wall position and thickness estimation from sequences of echocardiograms images". *IEEE Trans. on Medical Imag.*, vol. 15, pp. 25-38, 1996.

[6] M. Figueiredo and J. Leitão. "Bayesian estimation of ventricular contours in angiographic images". *IEEE Trans. on Medical Imag.*, vol. 11, pp. 416–429, 1992.

[7] M. Figueiredo and J. Leitão. "Adaptive discontinuity location in image restoration". *in Proc. of the First IEEE Intern. Conf. on Image Proc. – ICIP'94*, vol. II, pp. 665–669, Austin, 1994.

[8] J. Ivins and J. Porrill. "Active region models for segmenting medical images". *in Proc. of the First IEEE Intern. Conf. on Image Proc. – ICIP'94*, pp. 227–231, Austin, 1995.

[9] A. Jain. *Fundamentals of Digital Image Proc.*. Prentice Hall, Englewood Cliffs, 1989.

[10] A. Jain, Y. Zhong, and S. Lakshmanan. "Object matching using deformable templates". Tech. Report TR-CPS-94-66, Pattern Recognition and Image Proc. Lab., Michigan State University, East Lansing, 1995.

[11] M. Kass, A. Witkin, and D. Terzopoulos. "Snakes: Active contour models". *Intern. Jour. of Comp. Vis.*, vol. 1, pp. 259–268, 1987.

[12] R. Malladi, J. Sethian, and B. Vemuri. "Shape modeling with front propagation". *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. 17, pp. 158–175, 1995.

[13] T. McInerney and D. Terzopoulos. "Topologically adaptable snakes". *in Proc. of the Intern. Conf. on Comp. Vis. – ICCV'95*, pp. 840–845, Cambridge, 1995.

[14] B. Porat. *Digital Processing of Random Signals*. Prentice Hall, New Jersey, 1994.

[15] P. Radeva, J. Serrat, and E. Martí. "A snake for model-based segmentation". *in Proc. of the Intern. Conf. on Comp. Vis. – ICCV'95*, pp. 816–821, Cambridge, 1995.

[16] J. Rissanen. "A universal prior for integers and estimation by minimum description length". *Annals of Statistics*, vol. 11, pp. 416–431, 1983.

[17] J. Rissanen. "Minimum-description-length principle". *in Encyclopedia of Statistical Sciences, Vol. 5*, pp. 523–527, New York, 1987. Wiley.

[18] R. Ronfard. "Region-based strategies for active contour models". *Intern. Jour. of Comp. Vis.*, vol. 13, pp. 229–251, 1994.

[19] L. Staib and J. Duncan. "Boundary finding with parametrically deformable models". *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. 14, pp. 1061–1075, 1992.

[20] G. Storvik. "A Bayesian approach to dynamic contours through stochastic sampling and simulated annealing". *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. PAMI-16, pp. 976–986, 1994.

[21] D. Youla. "Generalized image restoration by the method of alternating projections". *IEEE Trans. on Circuits and Syst.*, vol. 25, pp. 694–702, 1978.