Adaptive Parametrically Deformable Contours

Mário A. T. Figueiredo[1]*, José M. N. Leitão[1], and Anil K. Jain[2]

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

[2] Department of Computer Science.
Michigan State University,
East Lansing, MI 48824, U.S.A.

Abstract. We introduce a fully unsupervised contour estimation strategy based on parametrically deformable models. The problem is formulated in a statistical parameter estimation framework with both the contour and observation model (likelihood) parameters are considered unknown. Although other choices could fit in our formulation, we focus on Fourier and spline contour descriptors. To estimate the optimal parametrization order (e.g. number of Fourier coefficients) we adopt the *minimum description length* (MDL) principle. The result is a parametrically deformable contour with adaptive smoothness and which also autonomously estimates the observation model parameters.

1 Introduction and Previous Work

Image segmentation and contour estimation are among the most important, interesting, and challenging problems in image analysis and low-level computer vision. When no special assumptions are made concerning the morphological structure of the objects/regions to be estimated in the observed image, we are in the presence of an image segmentation problem, in the common meaning of the term. When the problem is more confined to that of finding some individual image region, it is commonly referred to as a contour estimation problem; a typical example is organ boundary estimation in medical images.

1.1 *Snakes* and Related Approaches

Having its roots in the seminal work of Kass, Witkin, and Terzopoulos [18], *snake*-type approaches constitute one of the most successful class of approaches to contour estimation. In its original version [18], snakes work by minimizing an *energy* function composed of an (internal) elastic-type term which penalizes the contour deformations, and an (external) attraction potential linking it with image features of interest. The result is a desired compromise between contour *smoothness*, on one hand, and adequacy to the observed data, on the other hand.

* Phone: +351 1 8418464. Fax: +351 1 8418472. E-mail: mtf@it.ist.utl.pt

The major drawbacks of conventional *snake*-type techniques are: lack of adaptiveness (all parameters have to be set a priori); inability to reparametrize itself during the deformation process; use of information strictly along the boundary, making it highly "myopic" and sensitive to initialization.

In recent years, several significant improvements, modifications, and reformulations have been proposed to overcome the referred limitations of the traditional *snake*-type models; see [4], [5], [6], [7], [14], [24], [28], and further references therein). Active contours with the ability to grow (or shrink) in order to accommodate larger (or smaller) objects, or even change their topology, have been proposed [22]. A particularly elegant and successful topology-independent formulation is the one recently proposed in [3] and [21]. This, however, is not very well suited to low-quality images (e.g., noisy medical images) since (like the original snake model) it relies on image gradients.

1.2 Deformable Templates/Models

Deformable templates and models constitute another important approach to contour and object estimation. Here, global models generally described by a small number of parameters are used (in contrast with snakes, which typically use explicit contour descriptions); these models may directly describe (contour/object) shapes or a deformation suffered by basic template. The model parameters are then estimated in the presence of the observed data. Fundamental work on deformable templates is that of Grenander and his collaborators; see [2], [12], and [13] and references therein. For other important references on this type of approach see (this is by no means an exhaustive list): [16], [17], [30], [33], [34], and other references therein. As above, one of the main difficulties when using deformable templates is their lack of adaptiveness, with parameters having to be set a priori.

1.3 The Bayesian Viewpoint

From a Bayesian estimation view-point, active contours 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 [11], [31]. Deformable templates may not include a deformation energy since the parametrization may itself guarantee regularity of the represented shape [6]; nevertheless, when approached from a Bayesian viewpoint, it is common to include a prior to bias the estimate towards some preferred/predicted shape [16], [17], [30], [33], [34].

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 [10], [11]. The main difficulty in this approach is still 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 [11], we have proposed an adaptive Bayesian approach for a ventricular contour estimation problem. A technique which adaptively estimates the observation model parameters was proposed in [10]. Recent work in [35] also presents an adaptive *snake*-related scheme and contributes to the unification of energy-minimizing and Bayesian approaches.

1.4 Solving the Minimization Problem

Regardless of their theoretical/conceptual setting, both classical *snake*-type approaches and deformable templates/models lead to difficult minimization problems. A diverse set of approaches has been proposed to solve them: deterministic iterative energy minimization schemes (see many references in [6]); dynamic programming [1], [10]; multiresolution algorithms [16]; and stochastic methods including *simulated annealing* [31], the Metropolis algorithm [17], and simple (Langevin) diffusion or jump-diffusion processes [12], [2], [13].

1.5 Proposed Approach

In this paper, we propose a fully adaptive contour estimation strategy based on parametrically deformable models. The problem is formulated in a statistical estimation framework where both the contour parameters and the observation model parameters are considered unknown. Although other choices fit in our formulation, we focus on Fourier and spline parametric representations. An issue arising in parametrically deformable contours is the choice of the parametrization order (i.e., the number of parameters); to deal with it, we adopt the *minimum description length* (MDL) principle [25], [26]. The final result is a parametrically deformable contour with adaptive smoothness and which autonomously estimates the observation model parameters (e.g., noise variance, in additive noise models). Another aspect stressed in this paper is that care should be taken in deriving the observation model; in particular, for low quality images (such as medical images) gradient-based external energies may be completely useless.

2 Proposed Technique

2.1 Parametrically Deformable Contours

General Formulation. Deformable contours are (regular) shapes defined by a small number of parameters. Classical examples include Fourier and spline descriptors. Since the search space is already confined by the fact that the shapes are described by a small number of parameters, the elastic energy which penalizes deformations (as in *snake*-type models) is not needed. For example, a curve described by a small number of Fourier coefficients is automatically smooth.

Let $\boldsymbol{\theta}_{(K)}$ denote the set of parameters (parameter vector) defining the shape of a continuous closed contour. This parameter vector belongs to $\Theta_{(K)}$, the set of allowable configurations. The subscript $\cdot_{(K)}$ is used to indicate a K-order parametrization, e.g., a Fourier description with K terms or a spline with K control points. The closed contour (on the image plane) represented by $\boldsymbol{\theta}_{(K)}$ is a continuous periodic vector function $\mathbf{v}(t) = [x(t)\ y(t)]$, of period 2π, i.e., of unit fundamental angular frequency[3]; its N-points discrete version is a $(N \times 2)$ vector

$$\mathbf{v} = \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_{N-1} \end{bmatrix} = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ \vdots & \vdots \\ x_{N-1} & y_{N-1} \end{bmatrix},$$

where $\mathbf{v}_i \equiv \mathbf{v}(i2\pi/(N-1))$, for $i = 0, 1, ..., N-1$. Given $\boldsymbol{\theta}_{(K)}$, the explicit contour representation $\mathbf{v}$ is obtained by some deterministic operator $\mathcal{V}_{(K)}$ defined on $\Theta_{(K)}$, i.e., we write $\mathbf{v} = \mathcal{V}_{(K)}\boldsymbol{\theta}_{(K)}$.

Fourier Descriptors. The complex Fourier series description of a continuous closed curve $\mathbf{v}(t) = [x(t)\ y(t)]^T$ (of unit fundamental angular frequency[3]) is

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

where the complex Fourier coefficients are

$$[c_k\ d_k] = \frac{1}{2\pi} \int_0^{2\pi} [x(t)\ y(t)]\, e^{-jkt} dt \tag{2}$$

(see [15], [30]). The discrete complex Fourier series representation is

$$\mathbf{v}_i = [x_i\ y_i] = \sum_{k=0}^{N-1} [e_k\ f_k]\ e^{jki2\pi/N} \tag{3}$$

with

$$[e_k\ f_k] \equiv \mathbf{g}_k = \frac{1}{N} \sum_{i=0}^{N-1} [x_i\ y_i]\ e^{-jki2\pi/N}. \tag{4}$$

By truncating the series in Eq. (3) (or in Eq. (1)) to K terms (with $K < N$), a smoothed version of the curve is obtained. The above defined parameter vector $\boldsymbol{\theta}_{(K)}$ contains, in this case, $2K$ complex Fourier coefficients (i.e., $4K$ real parameters)

$$\boldsymbol{\theta}_{(K)} = [\mathbf{g}_0^T\ \mathbf{g}_1^T\ ...\ \mathbf{g}_{K-1}^T]^T = \begin{bmatrix} e_0 & e_1 & \cdots & e_{K-1} \\ f_0 & f_1 & \cdots & f_{K-1} \end{bmatrix}^T. \tag{5}$$

Thus, the order of the parametrization defines the degree of smoothness of the contours. Operator $\mathcal{V}_{(K)}$ is, in this case, simply a matrix product

$$\mathbf{v} = \mathcal{V}_{(K)}\, \boldsymbol{\theta}_{(K)} = \mathbf{F}_{(K)} \boldsymbol{\theta}_{(K)},$$

[3] If t is understood as the arc length of the contour, it can be normalized such that the total length equals 2π.

where $\mathbf{F}_{(K)}$ is the $(N \times K)$ matrix representing the K-term truncated Fourier series, i.e., $\left[\mathbf{F}_{(K)}\right]_{ik} = e^{jki2\pi/N}$, for $i = 0, 1, ..., N-1$ and $k = 0, 1, ..., K-1$. It is worth noticing that, if there are no further constraints, i.e. $\Theta_{(K)}$ is the Euclidean space $\mathbb{R}^{4K}$, then the set of all contours (described by the elements of $\Theta_{(K)}$) is itself a linear space; it is the range $\mathcal{R}(\mathcal{V}_{(K)})$ of the linear operator $\mathcal{V}_{(K)}$.

Spline Descriptors. We consider cubic B-splines [15], widely used to represent contours with a small number of parameters. Spline representations in *snake*-type models have been explored by several authors (see [6] and references therein); the key idea has been that since splines minimize a deformation-like energy, the search can be confined to a set of such functions under the action of the image potential. B-splines provide a parametric curve representation of the form

$$\mathbf{v}(t) = [x(t)\ y(t)] = \sum_{k=0}^{K-1} [\alpha_k^x\ \alpha_k^y]\, B_k(t) \tag{6}$$

where $[\alpha_k^x,\ \alpha_k^y] = \boldsymbol{\alpha}_k \in \mathbb{R}^2$ are the *control points* and the $B_k(t)$ are the (cubic B-spline) *basis functions*, also known as *blending functions*. The key feature of this representation, which is uniquely defined by the control points, is: the basis functions are such that the cubic B-spline minimizes, among all functions passing by the control points, the second-order deformation energy [6]. Discretization is obtained by taking N equispaced samples of $\mathbf{v}(t)$; here, we assume that $N > K$. Accordingly, and interpreting the set of K control points as the parameter vector

$$\boldsymbol{\theta}_{(K)} = [\boldsymbol{\alpha}_0^T, \boldsymbol{\alpha}_1^T, ..., \boldsymbol{\alpha}_{K-1}^T]^T = \begin{bmatrix} \alpha_0^x & \alpha_1^x & \cdots & \alpha_{K-1}^x \\ \alpha_0^y & \alpha_1^y & \cdots & \alpha_{K-1}^y \end{bmatrix}^T, \tag{7}$$

the operator $\mathcal{V}_{(K)}$ is, in this case, also a matrix product

$$\mathbf{v} = \mathcal{V}_{(K)}\, \boldsymbol{\theta}_{(K)} = \mathbf{B}_{(K)} \boldsymbol{\theta}_{(K)} \tag{8}$$

where the elements of $(N \times K)$ matrix $\mathbf{B}_{(K)}$ are given by $\left[B_{(K)}\right]_{ik} = B_k(2\pi i/(N-1))$. Note that, also in this case, in the absence of constraints on the control points, that is if $\Theta_{(K)} = \mathbb{R}^{2K}$, the set of all splines with K control points is itself a linear space.

2.2 The Observation Model

Although this is an often overlooked aspect, great care should be takin in defining the observation model. For specific applications (e.g., finding organ boundaries in medical images), all the available knowledge about the image acquisition process should be included [10], [11]. Not doing so may result in poor results, specially on very low quality images (see [10]).

The observed image $\mathbf{I}$ is an indirect, imperfect, and random function of the parameter vector $\boldsymbol{\theta}_{(K)}$, from which it has to be estimated. This observation model has a set of specific parameters $\boldsymbol{\phi}$, i.e., we write $\mathbf{I} = \mathcal{S}\left(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\right)$. The complete observation model is split into two blocks as follows:

$$\boldsymbol{\theta}_{(K)} \quad \stackrel{\mathcal{V}_{(K)}}{\longrightarrow} \quad \mathcal{V}_{(K)}\, \boldsymbol{\theta}_{(K)} \equiv \mathbf{v} \qquad \text{(the ideal contour)} \tag{9}$$

$$\mathbf{v} \stackrel{\text{image model}}{\longrightarrow} \mathbf{I} \qquad \text{(the observed image)} \tag{10}$$

where the first step depends on the type of parametrization chosen, and the second step captures the image generation/acquisition model. We now make the following assumptions concerning the observed image model, i.e. Eq. (10):

Conditional independence: given the contour $\mathbf{v} = \mathcal{V}\boldsymbol{\theta}_{(K)}$, the image pixels are independently distributed.

Region 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 pixels inside (resp. outside) have a common distribution characterized by a parameter vector $\boldsymbol{\phi}_{\text{in}}$ (resp. by $\boldsymbol{\phi}_{\text{out}}$), with $\boldsymbol{\phi} = [\boldsymbol{\phi}_{\text{in}}, \boldsymbol{\phi}_{\text{out}}]$.

From these assumptions, the likelihood function is written as

$$p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) = \left(\prod_{(i,j)\in\mathcal{I}(\mathbf{v})} p\left(I_{(i,j)}|\boldsymbol{\phi}_{\text{in}}\right) \right) \left(\prod_{(i,j)\in\mathcal{O}(\mathbf{v})} p\left(I_{(i,j)}|\boldsymbol{\phi}_{\text{out}}\right) \right) \tag{11}$$

where $I_{(i,j)}$ denotes pixel (i,j) of image $\mathbf{I}$, while $\mathcal{I}(\mathbf{v})$ and $\mathcal{O}(\mathbf{v})$ are the inside and outside regions of this contour, 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. This is a region-based model [10], [11], [14], [28] which, unlike gradient-based techniques, uses all the data in the image, and not only a narrow stripe along the contour. Moreover, it makes sense in situations where gradients do not characterize a region; e.g., when the two regions have the same mean but different variances.

2.3 The Estimation Criterion

The goal of an unsupervised scheme is clearly to estimate $\boldsymbol{\theta}_{(K)}$ and $\boldsymbol{\phi}$ from the observed image $\mathbf{I}$ based on the observation model (the likelihood function), i.e., to obtain the maximum likelihood (ML) estimate

$$\left(\widehat{\boldsymbol{\theta}}_{(K)}, \widehat{\boldsymbol{\phi}}\right)_{\text{ML}} = \arg \max_{\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}} p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}). \tag{12}$$

However, since K is unknown, this maximization suffers from a model order problem which can be stated as follows (assuming, for simplicity, known $\boldsymbol{\phi}$):

– The parameter spaces $\{\Theta_{(K)},\ K = 1, 2, ...\}$ are *nested* in the following sense: for each $\boldsymbol{\theta}_{(K)} \in \Theta_{(K)}$, there is some $\boldsymbol{\theta}'_{(K+1)} \in \Theta_{(K+1)}$ such that

$$p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) = p(\mathbf{I}|\boldsymbol{\theta}'_{(K+1)}, \boldsymbol{\phi}). \tag{13}$$

In the Fourier descriptor case, $\boldsymbol{\theta}_{(K)} = [\mathbf{g}_0^T, \mathbf{g}_1^T, ..., \mathbf{g}_{K-1}^T] \in \Theta_{(K)}$ and $\boldsymbol{\theta}'_{(K+1)} = [\mathbf{g}_0^T, \mathbf{g}_1^T, ..., \mathbf{g}_{K-1}^T, [0,0]] \in \Theta_{(K+1)}$ describe the same contour. For the spline parametrization, it is also clear that $\boldsymbol{\theta}_{(K)} = [\boldsymbol{\alpha}_0^T, \boldsymbol{\alpha}_1^T, ..., \boldsymbol{\alpha}_{K-1}^T] \in \Theta_{(K)}$ and $\boldsymbol{\theta}'_{(K+1)} = [\boldsymbol{\alpha}_0^T, \boldsymbol{\alpha}_1^T, ..., \boldsymbol{\alpha}_{K-1}^T, [0,0]] \in \Theta_{(K+1)}$ correspond to same curve.

- Consequently, K can not be estimated directly by maximizing the likelihood function; in fact, $p(\mathbf{I}|\widehat{\boldsymbol{\theta}}_{(K)}, \boldsymbol{\phi})$, where $\widehat{\boldsymbol{\theta}}_{(K)}$ is the ML estimate of $\boldsymbol{\theta}_{(K)}$ given K, is a non-decreasing function of K [23].

To overcome this difficulty, we adopt Rissanen's *minimum description length* (MDL) criterion [25], [26], [27]. MDL is an information-theoretical principle which, simply put, states that the best model is the one allowing the shortest joint description of the observed data and the model itself. Formally,

$$\left(\widehat{\boldsymbol{\theta}_{(K)}}, \widehat{\boldsymbol{\phi}}\right)_{\mathrm{MDL}} = \arg\min \left\{ -\log p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) + L(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) \right\}, \qquad (14)$$

with $\widehat{\boldsymbol{\theta}_{(K)}}$ standing for the joint estimates of K and $\boldsymbol{\theta}_{(K)}$, i.e., $\widehat{\boldsymbol{\theta}_{(K)}} \equiv \widehat{\boldsymbol{\theta}}_{(\widehat{K})}$. The first term of the objective function in the right-hand side of Eq. (14) is simply the Shannon codelength[4] obtained when coding $\mathbf{I}$ based on the probabilistic model $p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi})$ [8]. The second term is the codelength of the parameters which can be split as $L(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) = L(\boldsymbol{\theta}_{(K)}) + L(\boldsymbol{\phi})$. Being independent of K, $L(\boldsymbol{\phi})$ can be omitted from the objective function. As shown in [25], [27], the optimal codelength for each real-valued parameter is $\frac{1}{2}\log M$, where M is the number of data points to be encoded (in our case the number of image pixels). Accordingly,

$$L(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) = \begin{cases} 2K\log M, & \text{for the Fourier parametrization,} \\ K\log M, & \text{for the spline parametrization,} \end{cases} \qquad (15)$$

because a K-order Fourier parametrization involves $2K$ complex parameters, i.e. $4K$ real ones, while a spline with K control point involves $2K$ real parameters. Notice that, if K is known, the MDL and ML estimates coincide; this is an important feature of the MDL principle.

2.4 A Bayesian Perspective

The MDL principle was not proposed by Rissanen with a Bayesian interpretation in mind [27]. However, Eq. (14) has a clear Bayesian interpretation as a *maximum a posteriori* (MAP) estimator,

$$\begin{aligned}\left(\widehat{\boldsymbol{\theta}_{(K)}}, \widehat{\boldsymbol{\phi}}\right)_{\mathrm{MAP}} &= \arg\max \left\{ p\left(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\,|I\right) \right\} = \arg\max \left\{ p\left(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\right)\, p\left(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\right) \right\} \\ &= \arg\min \left\{ -\log p(\mathbf{I}|\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) - \log p\left(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\right) \right\},\end{aligned}$$

with the prior

$$p(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}) \propto \begin{cases} \exp\left\{-2K\log M\right\}, & \text{for the Fourier parametrization,} \\ \exp\left\{-K\log M\right\}, & \text{for the spline parametrization.} \end{cases} \qquad (16)$$

[4] In bits or *nats* for binary or natural logarithms, respectively [8].

Since K is the number of terms in the Fourier description of the contour, or the number of control points in the spline model, these are basically implicit smoothing priors. The smaller K is, the simpler (smoother) the contours will be. These priors have the advantage of avoiding the shrinkage associated with smoothing priors explicitly expressed on the contour coordinates [30]. Finally, we stress that the criterion in Eq. (14) does not require the previous specification of parameters, thus constituting a fully unsupervised estimator.

3 Solving the Optimization Problem

3.1 Introduction

We are now left with the difficult task of solving

$$\left(\widehat{\boldsymbol{\theta}_{(K)}}, \widehat{\boldsymbol{\phi}}\right)_{\text{MDL}} = \arg\min E\big(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\big) \tag{17}$$

where the objective (energy) function is given by

$$E\big(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\big) = -\sum_{(i,j)\in\mathcal{I}(\mathbf{V})} \log p\left(I_{(i,j)}|\boldsymbol{\phi}_{\text{in}}\right) - \sum_{(i,j)\in\mathcal{O}(\mathbf{V})} \log p\left(I_{(i,j)}|\boldsymbol{\phi}_{\text{out}}\right) + K\beta \tag{18}$$

with $\mathbf{v} = \mathcal{V}_{(K)}\,\boldsymbol{\theta}_{(K)}$, while $\beta = 2\log M$ or $\beta = \log M$, for the Fourier or spline descriptors, respectively.

The optimization problem will be dealt with in a hierarchical way: in a lower (inner) level, the energy is minimized with respect to $\boldsymbol{\theta}_{(K)}$ and $\boldsymbol{\phi}$, keeping K constant; in an upper (outer) level, the resulting function is minimized with respect to K,

$$\min_{K}\left\{\min_{\boldsymbol{\theta}_{(K)},\boldsymbol{\phi}} E\big(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\big)\right\}. \tag{19}$$

3.2 Minimizing the Energy for Fixed K

Introduction. If K is known, the problem (17)-(18) is closely related to classical deformable (known order) template matching, which can still only be achieved by iterative schemes. Several two-alternating-steps schemes have been proposed for this kind of objective functions; see [6] for references and an elucidative review. Our particular problem, however, has two important specificities: (i) the observation parameters $\boldsymbol{\phi}$ are unknown; (ii) the region nature of the data term does not allow the energy to be written as a sum (or integral) of elementary energies, one for each template point, as is required in [6].

An equivalent constrained problem. To rewrite the problem in terms of the explicit contour $\mathbf{v}$, we have to constrain the solution to the space of those that can be obtained as $\mathbf{v} = \mathcal{V}_{(K)}\,\boldsymbol{\theta}_{(K)}$. Let this space (which is the range of $\mathcal{V}_{(K)}$) be denoted $\Omega_{(K)} = \mathcal{V}_{(K)}\,\Theta_{(K)}$. Then, a constrained problem equivalent to the unconstrained original one is

$$\left(\widehat{\mathbf{v}}, \widehat{\boldsymbol{\phi}}\right)_{\mathrm{MDL}} = \text{solution of} \begin{pmatrix} \text{minimize: } E'(\mathbf{v}, \boldsymbol{\phi}) \\ \text{subject to: } \mathbf{v} \in \Omega_{(K)} \end{pmatrix} \tag{20}$$

where $E'(\mathbf{v}, \boldsymbol{\phi})$ is given by (18) without the βK term. From $\widehat{\mathbf{v}} \in \Omega_{(K)}$, the parameter estimate is obtained as $\widehat{\boldsymbol{\theta}}_{(K)} = \mathcal{V}_{(K)}^{-1}\widehat{\mathbf{v}}$. Notice that the inverse $\mathcal{V}_{(K)}^{-1}$ is well defined on $\Omega_{(K)}$ (the range of $\mathcal{V}_{(K)}$) by the fact that (in both the Fourier and spline cases) $\mathcal{V}_{(K)}$ is linear and its null space only contains the null vector [19]. This is basically a form of the *invariance property* of ML estimation [23].

The Algorithm. To solve the constrained minimization problem (20), we use a form of the *gradient projection method* [20], [29]. In this technique, which is simply a modification of classical gradient descent, the (negative) gradient is projected onto the constraint space thus assuring that the updated solutions always belong to this space [20]. Concerning $\boldsymbol{\phi}$, there are no constraints, and we assume that the energy can be exactly minimized with respect to it (see comments below). Formally, the proposed algorithm works as follows:

Fixed-K Algorithm

Step 0: Initialization: assume some initial estimate $\widehat{\mathbf{v}}^0 \in \Omega_{(K)}$. Let $n = 0$.
Step 1 Update the estimate of $\boldsymbol{\phi}$

$$\widehat{\boldsymbol{\phi}}^{n+1} = \arg\min_{\boldsymbol{\phi}} E'(\widehat{\mathbf{v}}^n, \boldsymbol{\phi})\,. \tag{21}$$

Step 2 Compute a small step in the direction opposed to the gradient of the energy with respect to the contour (at its present location)

$$\delta\mathbf{v} = -\varepsilon \;\mathrm{vsgn}\left(\nabla E'\left(\mathbf{v}, \widehat{\boldsymbol{\phi}}^{n+1}\right)\Big|_{\mathbf{v}=\widehat{\mathbf{v}}^n}\right) \tag{22}$$

where $\mathrm{vsgn}(\cdot)$ denotes a coordinate-wise sign function.
Step 3 Compute the projection of $\delta\mathbf{v}$ onto $\Omega_{(K)}$, denoted $\mathcal{P}_{(K)}\delta\mathbf{v}$, and update the contour estimate as

$$\widehat{\mathbf{v}}^{n+1} = \widehat{\mathbf{v}}^n + \mathcal{P}_{(K)}\delta\mathbf{v}. \tag{23}$$

Step 4 If some stopping criterion is met, stop, otherwise go back to **Step 1**.

Relation to other schemes. The projection operation guarantees that $\widehat{\mathbf{v}}^{n+1} \in \Omega_{(K)}$ since $\widehat{\mathbf{v}}^{n} \in \Omega_{(K)}$ and $\Omega_{(K)}$ is a linear space. Moreover, since $\mathcal{P}_{(K)}\widehat{\mathbf{v}}^{n} = \widehat{\mathbf{v}}^{n}$, and the projection operator is linear, the update (23) can be rewritten as

$$\widehat{\mathbf{v}}^{n+1} = \mathcal{P}_{(K)}\left(\widehat{\mathbf{v}}^{n} + \delta\mathbf{v}\right); \tag{24}$$

this reveals the similarity of this algorithm with the two-steps schemes described in [6], if we leave aside the parameter estimation performed in **Step 1**. Our **Step 2** corresponds to what is termed *deformation*, in [6], while our projection step (**Step 3**) corresponds to what is there called *model fitting*. Since any linear space is a convex set, this algorithm can also be seen as having some relation with the *projection onto convex sets* (POCS) technique [32]. Finally, some resemblance with the *expectation-maximization* (EM) scheme of Dempster *et al* [9] may be noticed; our **Step 1** plays the role of M-step while **Step 2** and **Step 3** represent the E-step.

3.3 Some comments.

Concerning Step 1. Implementing **Step 1** consists in obtaining the ML estimate of $\boldsymbol{\phi} = [\boldsymbol{\phi}_{\text{in}}\ \boldsymbol{\phi}_{\text{out}}]$, considering a fixed contour $\widehat{\mathbf{v}}^{n}$, from the likelihood function (11). This depends on the particular image model assumed which, in the experiments presented ahead, will be:

- **Gaussian.** All pixels are independent and Gaussian distributed with means μ_{in} and μ_{out} and variances σ^2_{in} and σ^2_{out}, for the inside and outside regions, respectively. In this case, $\boldsymbol{\phi}_{\text{in}} = [\mu_{\text{in}}\ \sigma^2_{\text{in}}]$, $\boldsymbol{\phi}_{\text{out}} = [\mu_{\text{out}}\ \sigma^2_{\text{out}}]$, and **Step 1** consists simply in computing the (inside and outside) sample mean and variance (which are the ML estimates given independent samples).
- **Rayleigh.** In this case, which adequately models echographic images [10], the pixels are Rayleigh distributed. For the inside pixels, we have

$$p(I_{(i,j)}|\sigma^2_{\text{in}}) = \frac{I_{(i,j)}}{\sigma^2_{\text{in}}} \exp\left\{-\frac{I^2_{(i,j)}}{2\sigma^2_{\text{in}}}\right\} \tag{25}$$

  and a similar expression (with σ^2_{out}) for the outside ones. The parameter vector is now $\boldsymbol{\phi} = [\sigma^2_{\text{in}}\ \sigma^2_{\text{out}}]$ and the respective ML estimates are (see [10]) simply one half of the sample means of squares.

Concerning Step 2. Here, we have to compute the gradient of the energy $E'(\mathbf{v}, \boldsymbol{\phi})$ with respect to each coordinate of the explicit contour representation $\mathbf{v}$. Since the coordinates represent pixel locations in a digital image, they are intrinsically discrete (in fact integer-valued) and the gradient is approximated by discrete differences relatively to each contour coordinate. It is possible to show that this gradient is always normal to the contour; this is exactly true for a continuous representation and a good approximation for a fine enough discretization. Parameter ε should be kept small to avoid instabilities near the minima of the objective function.

Concerning Step 3. This step requires finding the point in $\Omega_{(K)}$ which is closest to $\delta\mathbf{v}$,

$$\mathcal{P}_{(K)}\left(\widehat{\mathbf{v}}^n + \delta\mathbf{v}\right) = \mathcal{V}_{(K)}\left(\arg\min_{\boldsymbol{\theta}_{(K)}} \| \mathcal{V}_{(K)}\boldsymbol{\theta}_{(K)} - \left(\widehat{\mathbf{v}}^n + \delta\mathbf{v}\right) \|^2\right) \tag{26}$$

where $\|\cdot\|$ denotes Euclidean norm; in other words, we have to look for the $\boldsymbol{\theta}_{(K)}$ which best fits $\widehat{\mathbf{v}}^n + \delta\mathbf{v}$ in a mean squared error sense. The two parameterizations considered have to be studied separately:

Fourier. In this case, the elements of $\boldsymbol{\theta}_{(K)}$ represent coordinates in an orthogonal basis; then, all that has to be done is to compute the Fourier series (according to (3)) truncated to the first K terms.

Splines In this case, what has to be solved is the least squares fit expressed in (26), with $\mathcal{V}_{(K)} = \mathbf{B}_{(K)}$, which involves the pseudo-inverse of matrix $\mathbf{B}_{(K)}$,

$$\mathcal{P}_{(K)}\left(\widehat{\mathbf{v}}^n + \delta\mathbf{v}\right) = \mathbf{B}_{(K)}\left(\mathbf{B}_{(K)}^T\mathbf{B}_{(K)}\right)^{-1}\mathbf{B}_{(K)}^T\left(\widehat{\mathbf{v}}^n + \delta\mathbf{v}\right). \tag{27}$$

Of course, for each K, this operator can be computed before running the algorithm.

3.4 Solving with respect to K

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

Unknown-K Algorithm

Step A: Let $K = K_{\max}$ and take some initial contour estimate $\widehat{\mathbf{v}}^0 \in \Omega_{(K)}$.

Step B: Run the **Fixed-K Algorithm** with the current value of K, and taking $\widehat{\mathbf{v}}^0 \in \Omega_{(K)}$ as initial contour estimate. Store the final values of $\widehat{\boldsymbol{\theta}}_{(K)} = \mathcal{V}_{(K)}^{-1}\widehat{\mathbf{v}}$ and $E\left(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\right)$.

Step C: If $K > 1$, let $K = K - 1$, let $\widehat{\mathbf{v}}^0 \in \Omega_{(K)} = \mathcal{V}_{(K+1)}\boldsymbol{\theta}_{(K+1)}$, and go back to **Step B** (i.e. take the result of the previous **Fixed-K Algorithm** as the initial estimate for the next one).

Step D: Find the minimum of all stored values of $E\left(\boldsymbol{\theta}_{(K)}, \boldsymbol{\phi}\right)$ and take corresponding estimates as the final ones.

4 Experimental Results

5 Concluding Remarks

We have presented a new approach to unsupervised deformable contour estimation. The problem is formulated as a statistical parameter estimation problem; since the number of parameters is unknown (the order of the contour parametrization), the MDL principle was invoked. The observation model parameters are also considered unknown and estimated simultaneously with the contour. After showing that, for a given order, the resulting optimization problem can be formulated as a constrained one, we apply a form of the gradient projection algorithm. This, in turn, is inserted into an outer loop which takes care of the order estimation part.

Examples were presented, using both synthetic and medical ultrasound images, showing the ability of the proposed method to estimate contours in an unsupervised manner, i.e. adapting to unknown smoothness and unknown observation parameters. In the case of the synthetic images, the good adequacy between the estimated and true parameter values testifies for the good performance of our approach.

References

1. A. Amini, T. Weymouth, and R. Jain. "Using dynamic programming for solving variational problems in vision". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 855–867, 1990.
2. Y. Amit, U. Grenander, and M. Piccioni. "Structural image restoration through deformable templates". *Jour. of the American Statist. Assoc.*, vol. 86, pp. 376–387, 1991.
3. V. Caselles, R. Kimmel, and G. Sapiro. "Geodesic active contours". *in Proc. of the Intern. Conf. on Computer Vision – ICCV'95*, pp. 694–699, Cambridge, 1995.
4. A. Chakraborty, L. Staib, and J. Duncan. "Deformable boundary finding influenced by region homogeneity". *in Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition – CVPR'94*, pp. 624–627, Seattle, 1994.
5. L. Cohen. "On active contour models and baloons". *Computer Vision, Graphics, and Image Processing (CVGIP): Image Understanding*, vol. 53, pp. 211–218, 1991.
6. L. Cohen. "Auxiliary variables and two-step iterative algorithms in computer vision problems". *Intern. Jour. of Computer Vision*, vol. 6, pp. 59–83, 1996.
7. L. Cohen and I. Cohen. "Finite-element methods for active contour models and baloons for 2D and 3D images". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1131–1147, 1993.
8. T. Cover and J. Thomas. *Elements of Information Theory.* John Wiley & Sons, New York, 1991.
9. A. Dempster, N. Laird, and D. Rubin. "Maximum likelihood estimation from incomplete data via the EM algorithm". *Jour. of the Royal Statistical Society B*, vol. 39, pp. 1–38, 1977.
10. J. Dias and J. Leitão. "Wall position and thickness estimation from sequences of echocardiograms images". *IEEE Trans. on Medical Imaging*, vol. 15, pp. 25-38, 1996.
11. M. Figueiredo and J. Leitão. "Bayesian estimation of ventricular contours in angiographic images". *IEEE Trans. on Medical Imaging*, vol. 11, pp. 416–429, 1992.
12. U. Grenander. *General Pattern Theory: A Mathematical Study of Regular Structures.* Oxford University Press, Oxford, 1993.
13. U. Grenander and M. Miller. "Representation of knowledge in complex systems". *Jour. of the Royal Statistical Society B*, vol. 56, pp. 1–33, 1994.

14. J. Ivins and J. Porrill. "Active region models for segmenting medical images". *in Proc. of the IEEE Intern. Conf. on Image Processing – ICIP'94*, pp. 227–231, Austin, 1995.
15. A. Jain. *Fundamentals of Digital Image Processing.* Prentice Hall, Englewood Cliffs, 1989.
16. A. Jain, Y. Zhong, and S. Lakshmanan. "Object matching using deformable templates". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 267–277, 1996.
17. M. Dubuisson Jolly, S. Lakshmanan, and A. Jain. "Vehicle segmentation and classification using deformable templates". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 293–308, 1996.
18. M. Kass, A. Witkin, and D. Terzopoulos. "Snakes: Active contour models". *Intern. Jour. of Computer Vision*, vol. 1, pp. 259–268, 1987.
19. E. Kreyszig. *Introductory Functional Analysis with Applications.* John Wiley & Sons, New York, 1978.
20. D. Luenberger. *Linear and Nonlinear Programming.* Addison Wesley Publishing Company, Reading, Massachusetts, 1984. 2^{nd} Edition.
21. R. Malladi, J. Sethian, and B. Vemuri. "Shape modeling with front propagation". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 158–175, 1995.
22. T. McInerney and D. Terzopoulos. "Topologically adaptable snakes". *in Proc. of the Intern. Conf. on Computer Vision – ICCV'95*, pp. 840–845, Cambridge, 1995.
23. B. Porat. *Digital Processing of Random Signals.* Prentice Hall, New Jersey, 1994.
24. P. Radeva, J. Serrat, and E. Martí. "A snake for model-based segmentation". *in Proc. of the Intern. Conf. on Computer Vision – ICCV'95*, pp. 816–821, Cambridge, 1995.
25. J. Rissanen. "A universal prior for integers and estimation by minimum description length". *Annals of Statistics*, vol. 11, pp. 416–431, 1983.
26. J. Rissanen. "Minimum-description-length principle". *in Encyclopedia of Statistical Sciences, Vol. 5*, pp. 523–527, New York, 1987. Wiley.
27. J. Rissanen. *Stochastic Complexity in Stastistical Inquiry.* World Scientific, Singapore, 1989.
28. R. Ronfard. "Region-based strategies for active contour models". *Intern. Jour. of Computer Vision*, vol. 13, pp. 229–251, 1994.
29. J. Rosen. "The gradient projection method of nonlinear programming. Part I, linear constraints ". *SIAM Journal*, vol. 8, pp. 181-217, 1960.
30. L. Staib and J. Duncan. "Boundary finding with parametrically deformable models". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 1061–1075, 1992.
31. G. Storvik. "A Bayesian approach to dynamic contours through stochastic sampling and simulated annealing". *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 976–986, 1994.
32. D. Youla. "Generalized image restoration by the method of alternating projections". *IEEE Trans. on Circuits and Systems*, vol. 25, pp. 694–702, 1978.
33. A. Yuille. "Generalized deformable models, statistical physics, and the matching problem". *Neural Computation*, vol. 2, pp. 1–24, 1990.
34. A. Yuille and P. Hallinan. "Deformable templates". in A. Blake and A. Yuille, editors, *Active Vision*, pp. 21–38. MIT Press, 1992.

35. S. Zhu, T. Lee, and A. Yuille. "Region competition: Unifying snakes, region growing, energy/Bayes/MDL for multi-band image segmentation". *in Proc. of the Intern. Conf. on Computer Vision – ICCV'95*, pp. 416–423, Cambridge, 1995.