Bayesian Estimation of Ventricular Contours in Angiographic Images

Mário Teles de Figueiredo, *Student Member, IEEE,* and José M. N. Leitão

***Abstract*—This paper presents a new method for left ventricular contour determination in digital angiographic images. The problem is formulated in a Bayesian framework, adopting as estimation criterion the *maximum a posteriori* probability (MAP). The *true* contour is modeled as a one-dimensional noncausal Gauss–Markov random field and the observed image is described as the superposition of an *ideal image* (deterministic function of the *real* contour) with white Gaussian noise. The proposed algorithm estimates simultaneously the contour and the model parameters by implementing an adaptive version of the *iterated conditional modes* algorithm (AICM). The convergence of this scheme is proved and its performance evaluated on both synthetic and real angiographic images. The method exhibits robustness against image artifacts and the contours obtained are considered good by expert clinicians. Being completely data-driven and fast, the proposed algorithm is suitable for routine clinical use.**

I. Introduction

ANGIOGRAPHY is an invasive medical imaging modality that allows the radiographic visualization of vessels and cavities, through the introduction of a contrast agent. In digital angiography, X-rays are converted into visible light by an image intensifier and acquired by an analog video camera. The camera output signal is then converted to digital format and stored in a computer.

A common procedure in digital angiography is image subtraction. An image acquired before contrast injection is subtracted from all subsequent images, the result being only the structures one wishes to visualize. This is the origin of the common designation of DSA (digital subtraction angiography). All images used in the present work are DSA images.

Left ventricular angiography is an important procedure in the evaluation of many cardiac diseases, providing valuable information concerning anatomic and functional aspects [24], [32]. Computer analysis of angiographic images, aiming at objective and quantitative study of left ventricular function, has deserved special attention in recent years [20], [23]–[26], [29], [32]. Most of the work has been directed towards the automatic assessment of functional aspects like ejection fraction or wall motion. For any of these purposes, the automatic determination of the ventricle boundary is a necessary first step.

Most of the ventricle contour detection algorithms (for angiography or other imaging modalities) described in the literature (see, e.g., [3], [6], [11], [17], [24], [29], [30], [31]) are obtained in a more or less empirical way. In recent literature on image processing and computer vision, systematic approaches based on Bayesian estimation criteria have been described for problems such as image segmentation and restoration (fundamental references are [2] and [9]). This is also a trend in some fields of medical imaging, namely image reconstruction ([10], [12], and [27] are just a few examples).

Adopting also a Bayesian framework, this article proposes a method for ventricular contour estimation. The type of reasoning is not specific of this application and can easily be adapted to other problems, such as vessel border determination, or to other imaging modalities.

To model the contour continuity and smoothness, a one-dimensional noncausal Markov random field is used. The model is established by exploring the equivalence between Markov and Gibbs descriptions (Hammersley–Clifford theorem [1], [9]). In the Gibbs formulation, the *a priori* probability of the contour is written in terms of an energy-like function which is the sum of a set of local *potential* functions tailored to penalize the configurations considered less probable.

The observed image is modeled as the result of corrupting an *ideal image* (which is a deterministic function of the *true contour*) with additive white Gaussian noise (AWGN), not independent of the *true contour.* This approach leads to an estimator that uses the entire image and not a strictly local *edge-detector-like* algorithm.

The *a posteriori* probability of the contour given the image, computed according to Bayes law, is

$$p(\boldsymbol{r}|\boldsymbol{I}) = \frac{p(\boldsymbol{I}|\boldsymbol{r})p(\boldsymbol{r})}{p(\boldsymbol{I})} \tag{1}$$

where $p(\boldsymbol{r})$ is the *a priori* probability of the contour $\boldsymbol{r}$, and $p(\boldsymbol{I}|\boldsymbol{r})$ is the conditional probability of the observed image $\boldsymbol{I}$, given the contour. The *maximum a posteriori* (MAP) probability estimate is defined as

$$\begin{aligned}\hat{\boldsymbol{r}}^{\mathrm{MAP}} &= \arg\max_{\boldsymbol{r}} \{p(\boldsymbol{r}|\boldsymbol{I})\} \\ &= \arg\max_{\boldsymbol{r}} \{p(\boldsymbol{I}|\boldsymbol{r})p(\boldsymbol{r})\}. \end{aligned} \tag{2}$$

Note that for a given observed image, the probability $p(\boldsymbol{I})$ is a constant. The maximization (2) is carried out in this work by an adaptive version of the *iterated conditional modes* (ICM) algorithm [2], herein called AICM. This algorithm performs simultaneous parameter and contour estimation. We prove that AICM converges to the so called *partial optimal solutions,* which verify a suboptimal criterion [16].

Manuscript received June 20, 1991; revised January 10, 1992. This work was supported in part by Junta Nacional de Investigação Cientifica e Tecnológica (JNICT), project 87.342 MIC.

The authors are with Centro de Análise e Processamento de Sinais, Instituto Nacional de Investigação Cientifica and the Departamento de Engenharia Electroténica e de Computadores, Instituto Superior Técnico, Lisbon, Portugal.

IEEE Log Number 9107442.

This paper is related to the work of Friedland and Adam [8]: although not explicitly adopting a Bayesian framework, a one-dimensional Markov random field is also used in [8] to model the ventricular contour smoothness. There are, however, some fundamental differences: in [8], the interaction between the image and the contour (corresponding to $p(\boldsymbol{I}|\boldsymbol{r})$ in the Bayesian framework) is supported upon a strictly local edge detector rather than derived from a model of the observed data; moreover, the estimation procedure in [8] uses simulated annealing (which is a stochastic relaxation algorithm much more time consuming than ICM) and the model parameters are empirically defined by trial and error instead of estimated directly from the data. If wanted, temporal information, as proposed in [8], can also be incorporated in the method herein proposed by using a two-dimensional, rather than a one-dimensional Markov random field.

This paper is organized as follows. In Section II the noncausal Markov random field model of the contour is presented. The observed image model is described in Section III. Section IV contains the specification of the *a priori* contour model and the derivation of the *a posteriori* probability. In Section V the estimation algorithm is presented in two steps: first, a simpler version assuming knowledge of all the necessary parameters is described; next, the completely data-driven scheme, including simultaneous parameter estimation, is introduced. Tests performed on several real and synthetic images are presented in Section VI.

II. Noncausal Markov Random Field Contour Model

Noncausal Markov random fields have been used to describe prior knowledge in several image processing problems like restoration or segmentation [2], [9], [16], [21], [33]. This Section presents a brief overview of this formulation and its adequation to the problem under study.

A. Coordinate System

For the sake of analytical and computational simplicity it is convenient to describe the ventricle border by a one-dimensional scalar function. The particular type of description adopted depends on the analysis procedure to be implemented with the extracted contour.

Let the ventricular contour be described by discrete polar coordinates, as seen in Fig. 1,

$$\boldsymbol{R} = (R_i = R(\theta_i), 1 = 1, 2, \cdots, M) \quad \text{for } \theta_i = \theta_{\min} + (i-1)\xi \tag{3}$$

where $\xi = (\theta_{\max} - \theta_{\min})/(M-1)$. Each coordinate R_i is confined to a set of possible values Λ. For digital images, Λ is a finite set of discrete values although, for analytical convenience, it is sometimes assumed to be the set of real numbers.

The polar coordinate system herein adopted is adequate to two quantification methods called *center of mass method* (CMM) and *Stanford method* (SM) [20]. It should be noted that some ventricle contours can not be described in polar coordinates; these are highly pathological cases which are usually excluded from automatic quantitative analysis.

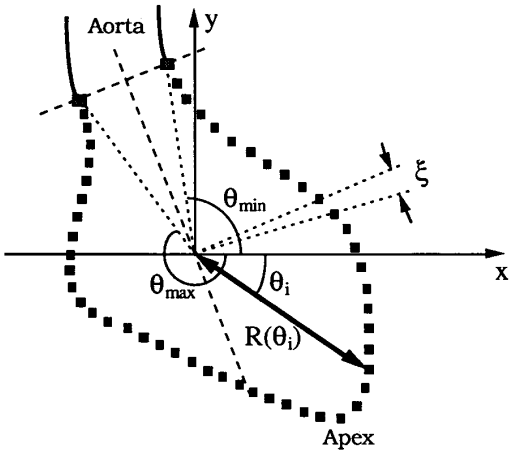


Fig. 1. Ventricular contour in discrete polar coordinates.

Other coordinate systems can be adopted, like the one used by the *long axis method* (LAM) [20] or the one introduced in [17]. These are still one-dimensional scalar descriptions of the contour and the method described in this paper could easily be adapted to them.

If other objects are to be estimated, more adequate coordinate systems have to be used, e.g., vessels can be described by a one-dimensional vector field of center position and width [22]; continuity and smoothness can still be modeled by a one-dimensional noncausal Markov random field.

B. One-dimensional Markov Random Field

For each contour point (site) i, a set of points is termed the *neighborhood* of i, V_i, if

$$i \notin V_i \quad \text{and} \quad j \in V_i \Leftrightarrow i \in V_j \quad \forall_{i,j=1,2,\cdots,M}. \tag{4}$$

The set of all possible configurations (contours) is $\Omega = \{\boldsymbol{R} = (r_1, r_2, \cdots, r_M) : r_i \in \Lambda, i = 1, 2, \cdots, M\}$.

It is assumed that each contour is a sample of a one-dimensional non-causal Markov random field, with respect to the neighborhood system $\mathcal{V} = \{V_i, i = 1, 2, \cdots, M\}$, that is

$$P(r_i|\{r_j : j \neq i\}) = P(r_i|\{r_j : j \in V_i\}) \quad \text{for } i = 1, 2, \cdots, M. \tag{5}$$

This equation should be seen as referring to probability density functions, if the variables are continuous, or to probability masses if the variables are discrete.

Note that the sites near the contour extrema, e.g., $i = 1$ and $i = M$, need special rules since they have neighbors from outside the contour. The problem of assigning values to the off-contour sites will be addressed in Section IV.

It is not trivial to verify the consistency of a given family of conditional probabilities like (5) [9]. By consistency is meant that Bayes law and the Kolmogorov–Chapman equations should be verified. The equivalence between Markov and Gibbs fields expressed by the Hammersley–Clifford theorem

[1], [9], presented in the sequel, allows the circumvention of this problem.

The Gibbs field definition is based on the concept of *clique.* A *clique* is a set either with only one site or else with mutually neighbor sites. Let $\mathcal{C}$ be the set of all cliques associated to a neighborhood system $\mathcal{V}$. A probability measure defined on Ω is called a Gibbs measure, with respect to a neighborhood system $\mathcal{V}$, if it has the form

$$p(\boldsymbol{r}) = \frac{1}{Z}\exp\left\{-\sum_{C\in\mathcal{C}} V_C(\boldsymbol{r})\right\}, \tag{6}$$

where

$$Z = \sum_{\boldsymbol{r}\in\Omega}\exp\left\{-\sum_{C\in\mathcal{C}} V_C(\boldsymbol{r})\right\}, \tag{7}$$

called the *partition function,* is a normalizing constant [9]. Each *clique potential* $V_C(\boldsymbol{r})$ depends only on the sites of $\boldsymbol{r}$ that belong to C, i.e., $V_C(\boldsymbol{r}) = V_C(\{r_i : i \in C\})$. If the variables are continuous the sum over Ω in (7) should be modified to an integral.

Hammersley–Clifford theorem states that a field is Markovian with respect to a neighborhood system $\mathcal{V}$ if and only if its joint probability has the Gibbs form (6), with respect to $\mathcal{V}$ [1], [9]. The local conditional probability is obtained from the Gibbs clique potentials by

$$P(r_i|\{r_j : j \in V_i\}) = \frac{1}{Z_i}\exp\left\{-\sum_{C:i\in C} V_C(\boldsymbol{r})\right\} \tag{8}$$

with

$$Z_i = \sum_{r_i\in\Lambda}\exp\left\{-\sum_{C:i\in C} V_C(\boldsymbol{r})\right\}. \tag{9}$$

The specification of a set of clique potentials poses no consistency problems [9].

III. Image Generation Mechanism

As seen above, the MAP estimator assumes knowledge of both the prior probability of the contour $p(\boldsymbol{r})$ and the probabilistic description of how an image is formed, i.e., the conditional probability of the image given a contour, $p(\boldsymbol{I}|\boldsymbol{r})$.

Assume the image expressed in the same polar coordinate system used to describe the contours, $I(\theta_i, \rho)$. Since a digital image is intrinsically rectangular, $I(\theta_i, \rho)$ should be understood as $I_d(\rho\cos\theta_i, \rho\sin\theta_i)$ where I_d represents a rectangularly addressable digital image. For digital images, $I(\theta_i, \rho)$ can only be defined for a finite and discrete set of values $\rho \in \Lambda = \{1, 2, \cdots, L\}$, which is the same that was previously assumed to contain the contour coordinates R_i. Figs. 2(a) and 3(a) show examples of intensities along radial lines belonging to the image of Figs. 10 and 12 (note in Fig. 3(a) the presence of a surgical stitch seen as a very abrupt variation just after pixel 40).

Let $\boldsymbol{I} = \{I(\theta_i, \rho) : i \in \{1, 2, \cdots M\}, \rho \in \Lambda\}$ be the observed image and $\boldsymbol{I_0} = \{I_0(\theta_i, \rho) : i \in \{1, 2, \cdots M\}, \rho \in \Lambda\}$ be an ideal image of a ventricle, with perfectly homogeneous contrast agent mixture and acquired without noise. The function $I_0(\theta_i, \rho, r_i)$, sketched in Fig. 4, would verify

$$I_0(\theta_i, \rho, r_i) = \begin{cases} A_i \Leftarrow \rho < r_i - \Delta \\ B_i \Leftarrow \rho > r_i + \Delta, \end{cases} \tag{10}$$

and have some shape between $r_i - \Delta$ and $r_i + \Delta$ describing the transition from the inside to the outside of the ventricle. In (10), A_i and B_i model the image intensity inside and outside

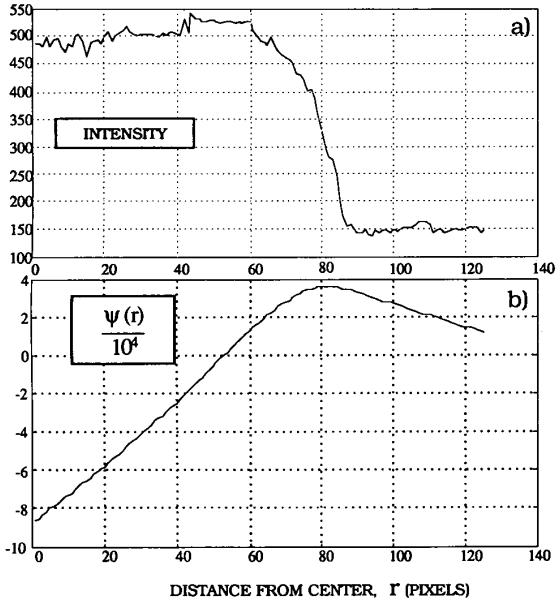


Fig. 2. (a) Image intensity along a radial line of the ventricle of Figs. 10 and 11. (b) Likelihood function of the border position.

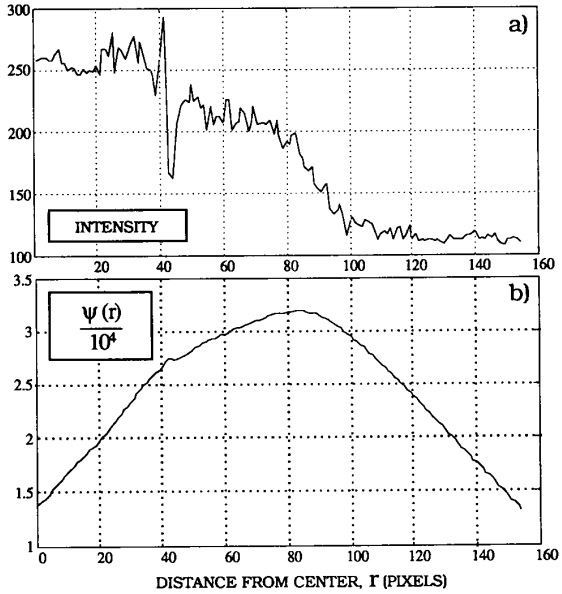


Fig. 3. (a) Image intensity along a radial line of the ventricle of Figs. 12 and 13. (b) Likelihood function of the border position.

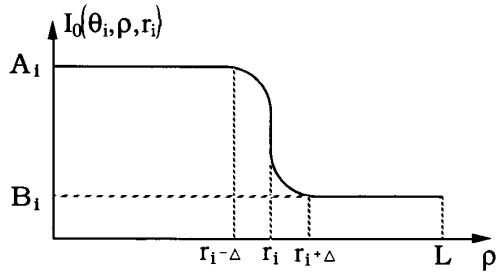


Fig. 4. Function $I_0(\theta_i, \rho, r_i)$; ideal intensity profile along a radial section of the ventricle image.

the ventricle, respectively, for the angle θ_i. It is assumed that $A_i > B_i$. For the sake of algorithmic simplicity an abrupt transition is assumed, i.e.,

$$I_0(\theta_i, \rho, r_i) = \begin{cases} A_i \Leftarrow \rho < r_i \\ B_i \Leftarrow \rho > r_i. \end{cases} \tag{11}$$

To complete the observed image model, noise has to be taken into account. In radiographic images there are three main sources of noise: quantum noise, scattered radiation and thermal noise [15], [24, p. 46].

- Quantum noise, present in all photoelectronic systems, is due to the fact that the process of photon detection obeys a Poisson law. This is particularly important in radiographic systems given the high energy of X-ray photons. As the mean and the variance of a Poisson distribution are equal, this kind of noise is signal-dependent. Since we are dealing with Poisson distributions of high mean, which can be well approximated by Gaussian functions, quantum noise is commonly modeled as signal-dependent additive white Gaussian noise (AWGN) [13, pp. 273–274].
- Thermal noise, mainly due to the video camera and subsequent electronic equipment, can be adequately modeled as signal-independent AWGN [13, pp. 273–274].
- Scattered radiation is described as a deterministic decrease of contrast plus another quantum noise term [7], [15].

In conclusion, the random perturbation can be taken as being additive white Gaussian noise, with variance dependent on the image intensity.

Given the image intensity function $I_0(\theta_i, \rho, r_i)$ [see (11)], the observed image can then be modeled as

$$I_0(\theta_i, \rho, r_i) = \begin{cases} I_0(\theta_i, \rho, r_i) + n_{A_i} \Leftarrow \rho \leq r_i \\ I_0(\theta_i, \rho, r_i) + n_{B_i} \Leftarrow \rho > r_i \end{cases} \tag{12}$$

where $n_{A_i}(\rho)$ and $n_{B_i}(\rho)$ are white Gaussian sequences with variances $\sigma^2_{A_i}$ and $\sigma^2_{B_i}$, respectively. To avoid notational complexity, we define the following vectors:

$$\boldsymbol{f}_i = [I_0(\theta_i, 1, r_i), I_0(\theta_i, 2, r_i), \cdots, I_0(\theta_i, L, r_i)]^T \tag{13}$$

$$\boldsymbol{g}_i = [I(\theta_i, 1, r_i), I(\theta_i, 2, r_i), \cdots, I(\theta_i, L, r_i)]^T. \tag{14}$$

The likelihood function is found to be

$$p(\boldsymbol{g}_i|r_i) = (2\pi)^{-\frac{L}{2}} \sigma_{A_i}^{-r_i} \sigma_{B_i}^{-L+r_i}$$

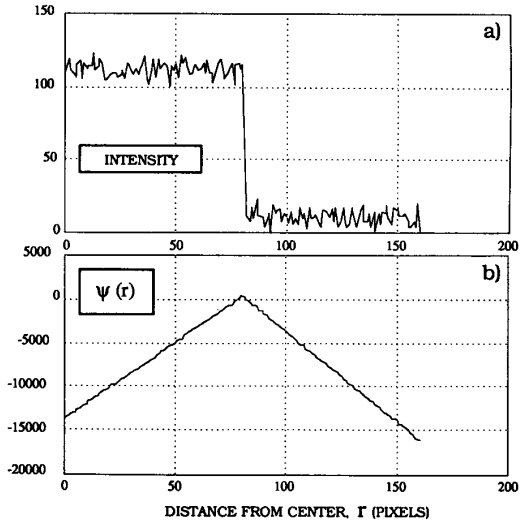


Fig. 5. (a) Ideal step edge with additive white Gaussian noise, (b) Likelihood function of the step location.

$$\cdot \exp\left\{-\frac{1}{2}[\boldsymbol{f}_i - \boldsymbol{g}_i]^T \boldsymbol{Q}^{-1}[\boldsymbol{f}_i - \boldsymbol{g}_i]\right\} \tag{15}$$

where $\boldsymbol{Q}$ is a $L \times L$ diagonal matrix,

$$\boldsymbol{Q} = \operatorname{diag}\underbrace{(\sigma^2_{A_i}, \sigma^2_{A_i}, \cdots, \sigma^2_{A_i}}_{r_i \text{ elements}}, \underbrace{\sigma^2_{B_i}, \sigma^2_{B_i}, \cdots, \sigma^2_{B_i})}_{L-r_i \text{ elements}}. \tag{16}$$

Note that $p(\boldsymbol{g}_i|r_i)$ also depends on r_i through vector $\boldsymbol{f}_i$ and matrix $\boldsymbol{Q}$. Introducing (11) into (15) and taking logarithms yields

$$\begin{aligned} \ln p(\boldsymbol{g}_i|r_i) &\propto 2\boldsymbol{v}_i^T \boldsymbol{g}_i - r_i\left(\frac{A_i^2}{\sigma^2_{A_i}} - \frac{B_i^2}{\sigma^2_{B_i}} + \ln\frac{\sigma^2_{A_i}}{\sigma^2_{B_i}}\right) \\ &\quad - \boldsymbol{g}_i^T \boldsymbol{Q}^{-1} \boldsymbol{g}_i \\ &\equiv \psi_i(r_i), \end{aligned} \tag{17}$$

where $\boldsymbol{v}_i$ is a vector of dimension L given by

$$\boldsymbol{v}_i = \underbrace{\left[\frac{A_i}{\sigma^2_{A_i}}, \frac{A_i}{\sigma^2_{A_i}}, \cdots, \frac{A_i}{\sigma^2_{A_i}}\right.}_{r_i \text{ elements}}, \underbrace{\left.\frac{B_i}{\sigma^2_{B_i}}, \frac{B_i}{\sigma^2_{B_i}}, \cdots, \frac{B_i}{\sigma^2_{B_i}}\right]^T}_{L-r_i \text{ elements}}. \tag{18}$$

Fig. 5(a) shows an ideal step edge ($A = 110$ and $B = 10$) contaminated by AWGN ($\sigma_A = \sigma_B = 10$). In Fig. 5(b) the likelihood function of the step position, $\psi(r)$ as given by (17), is presented. Monte–Carlo simulations were performed with this ideal step function; in all of the 200 runs executed the maximum of $\psi(r)$ was located at the true point, i.e., $r = 80$.

Figs. 2(b) and 3(b) show the likelihood functions associated to the real intensity profiles of Figs. 2(a) and 3(a), respectively. It should be noted that the abrupt variation of intensity seen in Fig. 3(a), due to a surgical stitch, does not affect much the likelihood function; the only effect of that artifact is a very weak local maximum at its location. This example illustrates the robustness of the adopted likelihood function when compared to what would be obtained with a local edge detection operator.

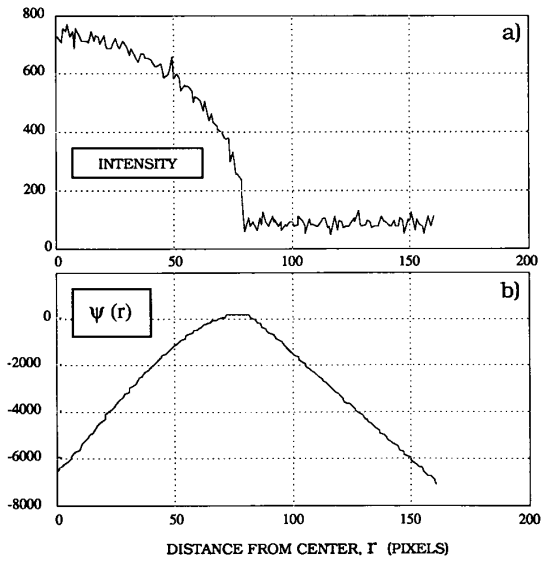


Fig. 6. (a) Simulated intensity function of the projection of a round object, contaminated by additive white Gaussian noise. (b) Likelihood function $\psi(r)$.

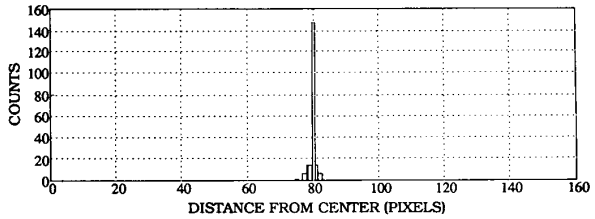


Fig. 7. Histogram of the position of the maximum of $\psi(r)$, relative to the simulated projection of a round object, for 200 Monte–Carlo runs.

To further study the behavior of this likelihood function under model mismatch conditions the following test was performed. A simulated intensity profile of the projection of a round object, contaminated by AWGN ($\sigma = 20$) was generated [Fig. 6(a)]. The histogram of the positions of the maxima of the likelihood functions, obtained by Monte–Carlo simulation (200 runs), is presented in Fig. 7. In Fig. 6(b) an example of $\psi(r)$ for this function is depicted. It can be concluded from this study that under this model mismatch condition the ML estimator is not as robust but its behavior is still very reasonable.

The set of image values along a radial line at a given angle depends only on the contour position for that angle. This is expressed by the conditional independence property

$$p(\boldsymbol{g}_i, \boldsymbol{g}_j | r_i, r_j) = p(\boldsymbol{g}_i | r_i) \cdot p(\boldsymbol{g}_j | r_j). \tag{19}$$

This allows writing

$$\begin{aligned} \ln p(\boldsymbol{I}|\boldsymbol{r}) &= \sum_{i=1}^{M} \ln p(\boldsymbol{g}_i | r_i) \\ &= \sum_{i=1}^{M} \psi_i(r_i), \end{aligned} \tag{20}$$

as the observed image $\boldsymbol{I}$ coincides with the set of all $\boldsymbol{g}_i$.

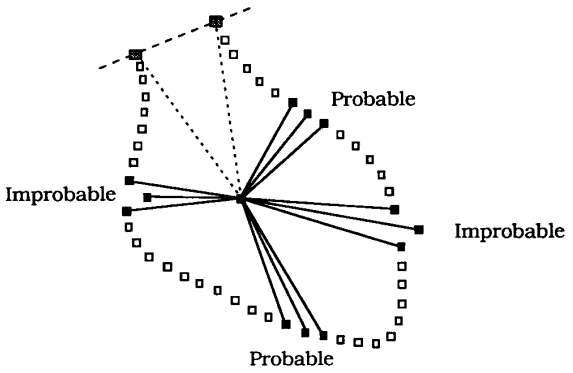


Fig. 8. Probable and improbable configurations.

IV. The *A Posteriori* Probabilty $P(\boldsymbol{r}|\boldsymbol{I})$

The MAP contour estimate is

$$\begin{aligned} \hat{\boldsymbol{r}}^{\mathrm{MAP}} &= \arg\max_{\boldsymbol{r}} \left\{ \ln p(\boldsymbol{I}|\boldsymbol{r}) - \sum_{C \in \mathcal{C}} V_C(\boldsymbol{r}) \right\} \\ &= \arg\min_{\boldsymbol{r}} \left\{ -\sum_{i=1}^{M} \psi_i(r_i) + \sum_{C \in \mathcal{C}} V_C(\boldsymbol{r}) \right\}. \end{aligned} \tag{21}$$

As shown in Section II, the clique potentials that define the *a priori* probability of the contour can be tailored to express prior knowledge about its characteristics. Some *a priori* knowledge comes from the fact that the ventricular contour is, with high probability, a smooth line without discontinuities or very high curvature segments. In Fig. 8, four configurations are represented, two being probable and the other two improbable. Squares of second-order differences,

$$V_{C_i} = K^2 (r_i - 2r_{i-1} + r_{i-2})^2, \tag{22}$$

can be adopted as they penalize the improbable configurations which are characterized by a high deviation from collinearity. In general, the adoption of a nth order difference (computed from $n+1$ values) as clique potential assumes neighborhoods that contain $2n$ points, $V_i = \{i-n, i-n+1, \cdots, i-1, i+1, \cdots, i+n-1, i+n\}$. In particular the second-order difference implies neighborhoods with 4 points $V_i = \{i-2, i-1, i+1, i+2\}$.

As mentioned in Section II, the off-contour neighbors of the extreme sites need special rules. It is convenient, from the analytical and computational points of view, to have a zero mean field. This is achieved by the procedure next described. Consider that the contour extrema R_0 and R_{M+1} are previously known values. The coordinates R_i, $i = 1, 2, \cdots, M$ of a given contour, illustrated in Fig. 9(a), evolve around the straight line $R_i = R_0 + i(R_{M+1} - R_0)/(M+1)$. With the change of coordinates

$$S_i = R_i - \left(R_0 + i \frac{R_{M+1} - R_0}{M+1} \right), \quad i = 1, 2, \cdots, M \tag{23}$$

the new values S_i evolve as seen in Fig. 9(b). The field

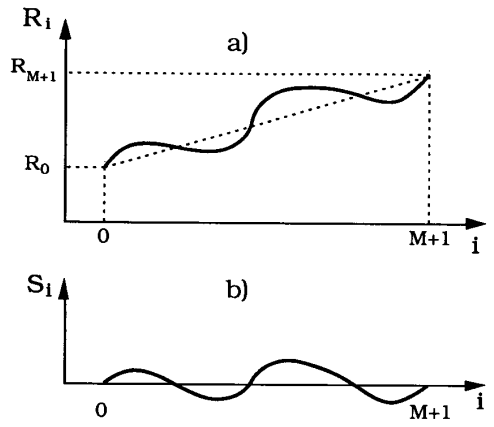


Fig. 9. (a) Contour coordinates R_i. (b) Modified contour coordinates S_i.

$\boldsymbol{S} = (S_i, i = 1,2,\cdots,M)$ can be given *free boundary conditions* by assigning zero value to all off-contour sites. The actual coordinates R_i can easily be recovered from the S_i using (23). Note that the values of the variables S_i are confined to a new set $\Lambda' = \{-L, -L+1, \cdots, 0 \cdots, L\}$.

The problem can now be shifted from finding $\hat{\boldsymbol{r}}^{\mathrm{MAP}}$ to finding $\hat{\boldsymbol{s}}^{\mathrm{MAP}}$. Once $\hat{\boldsymbol{s}}^{\mathrm{MAP}}$ is found, $\hat{\boldsymbol{r}}^{\mathrm{MAP}}$ can be obtained by inverting (23).

The functions $\psi_i(r_i)$ are computed as $\psi_i\big(s_i + R_0 + i\frac{R_{M+1}-R_0}{M+1}\big)$. To keep a shorter notation a new function is defined:

$$\varphi_i(s_i) \equiv \psi_i\left(s_i + R_0 + i\frac{R_{M+1} - R_0}{M+1}\right). \tag{24}$$

The conditional probabilities of the Markov formulation can be obtained from the clique potentials via (8) and (9). The following fact, concerning the local conditional probability of a field with clique potentials that are squares of Nth order differences can be derived from (8) and (9), [4]:

Let $\boldsymbol{S}$ be a Markovian, unidimensional field relatively to the Nth order neighborhood system $\mathcal{V} = \{V_i\}$ with $V_i = \{i-N, i-N+1, \cdots, i-1, i+1, \cdots, i+N-1, i+N\}$. The larger cliques associated with this neighborhood system are $C_i = \{i, i-1, i-2, \cdots, i-N\}$. Let the potentials of these larger cliques be squares of Nth order differences, that is,

$$V_{C_i}(s_i, s_{i-1}, \cdots, s_{i-N}) = K^2\left(\sum_{k=0}^{N} \alpha_k s_{i-k}\right)^2, \tag{25}$$

with

$$\alpha_k = (-1)^k \binom{N}{k} \quad \text{for } k = 0,1,\cdots,N, \tag{26}$$

and zero for all the other (smaller) cliques. Suppose that the set Λ' which contains the possible values of variables S_i is the set of real numbers and that the off-contour neighbors of the extreme sites are equal to zero. Then $\boldsymbol{S}$ is conditionally Gaussian:

$$P(s_i | s_j, j \in V_i) = \frac{1}{\sqrt{2\pi\lambda^2}} \cdot \exp\left\{-\frac{1}{2\lambda^2}\left(s_i - \sum_{k \in V_i} \beta_{ik} s_k\right)^2\right\} \tag{27}$$

with variance

$$\lambda^2 = \frac{1}{2K^2}\frac{(N!)^2}{(2N)!} \tag{28}$$

and

$$\beta_{ik} = \frac{(N!)^2(-1)^{|i-k|-1}}{(N+|i-k|)!(N-|i-k|)!} \quad \text{for } k = i \pm 1, i \pm 2, \cdots, i \pm N. \tag{29}$$

The joint probability of all the field sites is zero-mean Gaussian [1]

$$p(\boldsymbol{s}) = \frac{\sqrt{\det \boldsymbol{A}}}{\sqrt{(2\pi\lambda^2)^M}} \exp\left(-\frac{1}{2\lambda^2}\boldsymbol{s}^T\boldsymbol{A}\boldsymbol{s}\right) \tag{30}$$

where $\boldsymbol{A}$ (inverse of the covariance matrix) is an $M \times M$ symmetric matrix with elements given by

$$\boldsymbol{A}_{ij} = \begin{cases} 1 \Leftarrow i = j \\ -\beta_{ij} \Leftarrow i \neq j \end{cases} \tag{31}$$

It is clear that $\boldsymbol{A}$ is positive definite. Note that (30) can also be written as (6) and that the clique potentials given by (25) are squares of differences. As the extreme sites are set to zero, the clique potentials can only be all zero if all the site values are also zero.

It is interesting to note that the mean of (27), with the β_{ik} given by (29), is equal to the Lagrange polynomial interpolation of order $2N-1$, based on the $2N$ neighbors [4].

If the variables s_i are confined to a finite and discrete set (like $\Lambda' = \{-L, \cdots, 0, \cdots, L\}$), similar results can be derived, with the conditional and joint distributions being not Gaussian but discrete distributions with a Gaussian envelope.

In conclusion, note the following.

a) The contour is modeled by a one-dimensional noncausal Markov random field, with respect to the second order neighborhood system $\mathcal{V} = \{V_i = \{i-2, i-1, i+1, i+2\}, i = 1,2,\cdots,M\}$.

b) The contour model is decomposed into a zero mean Gauss–Markov random field (with free, i.e., zero, boundary conditions) $\boldsymbol{S}$, plus a deterministic part which is a function of the contour extrema R_0 and R_{M+1}:

$$R_i = S_i + R_0 + i\frac{R_{M+1} - R_0}{M+1}, \quad i = 1,2,\cdots,M. \tag{32}$$

Since the contour coordinates are known to belong to a finite discrete set of values, the Gaussian assumption is obviously not correct. This is, however, a common approximation when the number of possible values is large and the discretization step is small [2].

c) The values, R_0 and R_{M+1}, defined by the user, are assumed, as far as the estimation problem is concerned, as previously known constants.

d) There are three kinds of cliques associated with the adopted neighborhood system, $C_i^2 = \{i+1, i, i-1\}$, $C_i^1 = \{i, i-1\}e\ C_i^0 = \{i\}$. The clique potentials are

$$V_{C_i^2}(\boldsymbol{s}) = \frac{1}{12\lambda^2}(s_{i+1} - 2s_i + s_{i-1})^2 \tag{33}$$

$$V_{C_i^1}(\boldsymbol{s}) = 0 \tag{34}$$

$$V_{C_i^0}(\boldsymbol{s}) = 0 \tag{35}$$

where the factor $\frac{1}{12\lambda^2}$ was introduced by using (27) and (28) so that the conditional probability of a contour point given its neighbors is Gaussian with variance λ^2,

$$p(s_i|s_{i-2}, s_{i-1}, s_{i+1}, s_{i+2}) = \frac{\exp\left\{-\frac{1}{2\lambda^2}\left(s_i - \frac{1}{6}(-s_{i-2} + 4s_{i-1} + 4s_{i+1} - s_{i+2})\right)^2\right\}}{\sqrt{2\pi\lambda^2}}, \tag{36}$$

e) The MAP estimate of field $\boldsymbol{S}$ is

$$\hat{\boldsymbol{s}}^{\text{MAP}} = \arg\max_{\boldsymbol{s}} \{\ln p(\boldsymbol{s}|\boldsymbol{I})\} = \arg\min_{\boldsymbol{s}} \left\{-\sum_{i=1}^{M} \varphi_i(s_i) + \sum_{C \in \mathcal{C}} V_C(\boldsymbol{s})\right\}. \tag{37}$$

and the MAP estimate of field $\boldsymbol{R}$ is given by

$$\hat{r}_i^{\text{MAP}} = \hat{s}_i^{\text{MAP}} + R_0 + i\frac{R_{M+1} - R_0}{M+1}, \quad i = 1, 2, \cdots, M. \tag{38}$$

f) The functions $\varphi_i(s_i)$ can be seen as one site clique potentials. Thus, the *a posterior* probability is also Markov–Gibbs with the clique potentials given by

$$V_{C_i^2}^{\boldsymbol{p}}(\boldsymbol{s}) = \frac{1}{12\lambda^2}(r_i - 2r_{i-1} + r_{i-2})^2 \tag{39}$$

$$V_{C_i^1}^{\boldsymbol{p}}(\boldsymbol{s}) = 0 \tag{40}$$

$$V_{C_i^0}^{\boldsymbol{p}}(\boldsymbol{s}) = -\varphi_i(s_i) \tag{41}$$

and the *a posteriori* local conditional probability is as given by (42) below.

V. Estimation Algorithm

As expressed in (37), the computation of the MAP estimate involves the minimization of a nonconvex function of many variables. The main difficulty arises from the functions $\varphi_i(s_i)$ which are completely unknown *a priori.* Simulated annealing [9], [14] could be adopted at the cost of an immense computational burden, not compatible with routine use. As an alternative, we will adopt the deterministic relaxation algorithm known as *iterated conditional modes* (ICM), proposed by Besag [2]. This algorithm belongs to the class of optimization techniques known as *coordinate descent (ascent) methods* [19, pp. 227–230)].

The function to be minimized also depends on a set of unknown parameters that must be estimated if a completely data-driven algorithm is to be obtained. Simultaneous estimation of parameters (studied in, e.g., [2] and [16]) is a very important feature of a system designed for routine users with no previous knowledge about its values. This is even more true in the problem under study since a contour with M points possesses $4M + 1$ parameters.

Let $\Theta = \{A_i, B_i, \sigma_{A_i}, \sigma_{B_i}, i = 1, 2, \cdots, M\}$ be the set of unknown parameters of $p(\boldsymbol{I}|\boldsymbol{r})$ (see (20), and (17)). These parameters will be considered unknown but deterministic. Parameter λ of $p(\boldsymbol{s})$ [or $p(\boldsymbol{r})$], being the variance of the contour [see (30)], expresses prior knowledge about its degree of smoothness. Its estimation criterion must thus show preference for low values, corresponding to smooth contours. An adaptive process should not freely estimate λ from the data but rather make a compromise between adequation to the data and obedience to the preference for low values. Had we placed the estimation problem under a regularization framework this would be the issue of choosing the regularization parameter [28]. Under a Bayesian framework, a natural method is to impose an *a priori* distribution on the parameter, instead of considering it completely unknown. We assume the prior probability

$$p(\lambda) \propto \exp\{-a\lambda^2\} \tag{43}$$

where a is a parameter that specifies the degree of preference for low values. Other distributions could be considered; this one is simple, leads to a computationally easy estimator and to good results. The role of parameter a implies that it should not be estimated from the data but rather used to tune the behavior of the algorithm.

The simultaneous estimation of the contour and the parameters can be expressed by the criterion [16]

$$(\boldsymbol{s}^*, \Theta^*, \lambda^*) = \arg\max_{\boldsymbol{s},\Theta,\lambda} p(\boldsymbol{s}, \boldsymbol{I}, \lambda|\Theta). \tag{44}$$

Since this optimality criterion is extremely difficult to implement, a weaker one,

$$\boldsymbol{s}^* = \arg\max_{\boldsymbol{s}} p(\boldsymbol{s}, \boldsymbol{I}|\Theta^*, \lambda^*) \tag{45}$$

$$\Theta^* = \arg\max_{\Theta} p(\boldsymbol{s}^*, \boldsymbol{I}|\Theta, \lambda^*) \tag{46}$$

$$\lambda^* = \arg\max_{\lambda} p(\boldsymbol{s}^*, \boldsymbol{I}|\Theta^*, \lambda^*)p(\lambda), \tag{47}$$

$$p(s_i|s_{i-2}, s_{i-1}, s_{i+1}, s_{i+2}, \boldsymbol{I}) \propto \exp\left\{-\frac{\left(s_i - \frac{1}{6}(-s_{i-2} + 4s_{i-1} + 4s_{i+1} - s_{i+2})\right)^2}{2\lambda^2} + \varphi_i(s_i)\right\}. \tag{42}$$

called the *partial optimal solution,* is adopted [16]. The solutions to this optimization problem have the following properties:

- $\boldsymbol{s}^*$ is a MAP estimate based on $\boldsymbol{I}$, Θ^*, and λ^*. Note that $p(\boldsymbol{s},\boldsymbol{I}|\Theta^*,\lambda^*) = p(\boldsymbol{s}|\boldsymbol{I},\Theta^*,\lambda^*)\ p(\boldsymbol{I}|\Theta^*,\lambda^*)$, and $p(\boldsymbol{I}|\Theta^*,\lambda^*)$ does not depend on $\boldsymbol{s}$.
- Θ^* is a ML estimate based on $\boldsymbol{S}^*$ and $\boldsymbol{I}$ as is clear from (46).
- λ^* is a MAP estimate based on $\boldsymbol{S}^*$. In fact,

$$\begin{aligned}\lambda^* &= \arg\max_{\lambda} p(\boldsymbol{s}^*,\boldsymbol{I}|\Theta^*,\lambda)p(\lambda)\\ &= \arg\max_{\lambda} p(\boldsymbol{s}^*|\Theta^*,\lambda)p(\boldsymbol{I}|\boldsymbol{s}^*,\Theta^*,\lambda)p(\lambda)\\ &= \arg\max_{\lambda} p(\boldsymbol{s}^*|\lambda)p(\lambda) = \arg\max_{\lambda} p(\lambda|\boldsymbol{s}^*)\end{aligned} \quad (48)$$

since $p(\boldsymbol{I}|\boldsymbol{s}^*,\Theta^*,\lambda)$ does not depend on λ, and $p(\boldsymbol{s}^*|\Theta^*,\lambda)$ does not depend on Θ^*.

An adaptive version of ICM (AICM), similar to the one proposed by Besag [2], will be applied. As we prove in the Appendix, the only stationary points of this iterative process are the above mentioned *partial optimal solutions.*

The algorithm will be presented in two steps: 1) contour estimation assuming known parameters (ICM); 2) modification of ICM to include simultaneous parameter estimation (AICM).

A. ICM with Known Parameters

Consider a visiting schedule to the field sites $\{n_1, n_2, n_3, \cdots, n_t, \cdots\}$. Typical schedules include the cyclic visit to all the contour sites and the cyclic visit with alternating directions. In this last case, which is the one adopted herein, the algorithm coincides with the *Aitken double sweep method* [19, pp. 227–230]. At time t, variable s_k (for $k = n_t$) is replaced by the value that maximizes the local *a posteriori* conditional distribution (8)

$$\begin{aligned}s_k(t+1) &= \arg\max_{s_k} P(s_k|\{s_j(t) : j \in V_k\},\boldsymbol{I})\\ &= \arg\max_{s_k}\left\{\frac{1}{Z_k}\exp\left(-\sum_{C:k\in C} V_C^p(\boldsymbol{s})\right)\right\}\end{aligned} \quad (49)$$

which does not depend on the local partition function Z_k. Introducing (42), we get (50) below. Any line search technique can be used to solve (50). In the implementation here described we adopted the *golden section* search algorithm [19, pp. 199–200].

The algorithm is initialized, as proposed in [2], with the maximum likelihood (ML) estimate

$$\boldsymbol{s}(t=0) = \hat{\boldsymbol{s}}^{\mathrm{ML}} = \arg\max_{\boldsymbol{s}} \sum_{i=1}^{M} \varphi_i(s_i), \quad (51)$$

which can be computed separately for each variable s_i.

Writing $P(\boldsymbol{s}|\boldsymbol{I})$ as

$$\begin{aligned}P(\boldsymbol{s}|\boldsymbol{I}) = P(s_i|\{s_j(t) : j \neq i\},\boldsymbol{I})\\ \cdot\ P(\{s_j : j \neq i\}|\boldsymbol{I})\end{aligned} \quad (52)$$

where the second factor is the joint conditional (on $\boldsymbol{I}$) probability of all the sites other than i, and knowing, from the Markovian property, that

$$P(s_i|\{s_j(t) : j \neq i\},\boldsymbol{I}) = P(s_i|\{s_j(t) : j \in V_i\},\boldsymbol{I}), \quad (53)$$

it is clear that each ICM step increases $P(\hat{\boldsymbol{s}}|\boldsymbol{I})$ until a coordinatewise (possibly local) maximum is reached [2].

B. ICM with Parameter Estimation—AICM

Consider a visiting schedule to the field sites.

Step 0) Initialization:

Step 0.1) Parameter Initialization: Note that it is not necessary to provide an initial estimate of λ because the ML estimate of the contour does not depend on that parameter. As for $\{A_i, B_i, \sigma_{A_i}, \sigma_{B_i}, i = 1, 2, \cdots, M\}$, two values $r_D \in \Lambda$ and $r_F \in \Lambda$ can be defined such that

$$\begin{aligned}&\text{Probability of } (r_i > r_D) \approx 1 \quad \forall_i\\ &\text{Probability of } (r_i < r_F) \approx 1 \quad \forall_i.\end{aligned}$$

Reasonable initial estimates can be computed using only image values located to the left of r_D and to the right of r_F (adequate values are $r_D = 15$ and $r_F = L - 15$):

$$\hat{A}_i^{\mathrm{ML}}(\boldsymbol{I}) = \frac{1}{r_D}\sum_{\rho=1}^{r_D} g_i(\rho) \quad (54)$$

$$\hat{B}_i^{\mathrm{ML}}(\boldsymbol{I}) = \frac{1}{L - r_F + 1}\sum_{\rho=r_F}^{L} g_i(\rho) \quad (55)$$

$$\left(\widehat{\sigma_{A_i}^2}\right)^{\mathrm{ML}}(\boldsymbol{I}) = \frac{1}{r_D}\sum_{\rho=1}^{r_D}\left(g_i(\rho) - \hat{A}_i^{\mathrm{ML}}\right)^2 \quad (56)$$

$$\left(\widehat{\sigma_{B_i}^2}\right)^{\mathrm{ML}}(\boldsymbol{I}) = \frac{1}{L - r_F + 1}\sum_{\rho=r_F}^{L}\left(g_i(\rho) - \hat{B}_i^{\mathrm{ML}}\right)^2. \quad (57)$$

Step 0.2) Contour Initialization: Based on the preceding parameter values, determine the ML contour estimate according to (51).

Step 1) Update the parameter estimates according to

$$\hat{\Theta} = \arg\max_{\Theta}\{p(\hat{\boldsymbol{s}},\boldsymbol{I}|\Theta)\} \quad (58)$$

$$s_k(t+1) = \arg\max_{s_k}\left\{\varphi_k(s_k) - \frac{\left(s_k - \frac{1}{6}(-s_{k-2}(t) + 4s_{k-1}(t) + 4s_{k+1}(t) - s_{k+2}(t))\right)^2}{2\lambda^2}\right\} \quad (50)$$

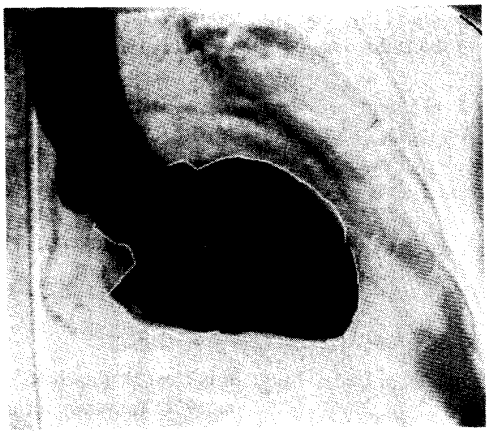


Fig. 10. Left ventricle ML contour estimate used to initialize the AICM algorithm.

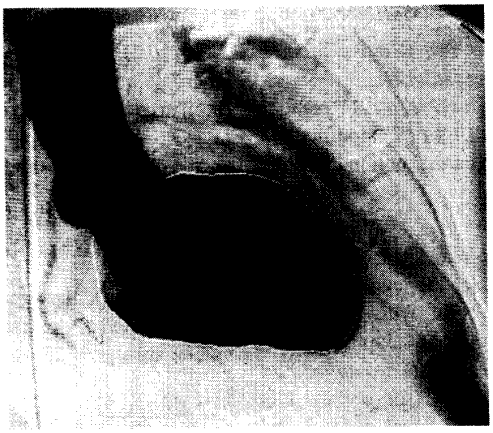


Fig. 11. Contour obtained after 8 AICM.

and

$$\hat{\lambda} = \arg\max_{\lambda} \{p(\hat{\boldsymbol{s}}|\lambda)p(\lambda)\} \tag{59}$$

where $\hat{\boldsymbol{s}}$ is the present contour estimate. Equation (17), together with the conditional independence property (19), leads to the following exact ML estimates:

$$\hat{A}_i^{\mathrm{ML}}(\hat{\boldsymbol{s}}, \boldsymbol{I}) = \frac{1}{\hat{r}_i} \sum_{\rho=1}^{\hat{r}_i} g_i(\rho) \tag{60}$$

$$\hat{B}_i^{\mathrm{ML}}(\hat{\boldsymbol{s}}, \boldsymbol{I}) = \frac{1}{L - \hat{r}_i} \sum_{\rho=\hat{r}_i(t)+1}^{L} g_i(\rho) \tag{61}$$

$$\left(\widehat{\sigma_{A_i}^2}\right)^{\mathrm{ML}}(\hat{\boldsymbol{s}}, \boldsymbol{I}) = \frac{1}{\hat{r}_i} \sum_{\rho=1}^{\hat{r}_i} \left(g_i(\rho) - \hat{A}_i^{\mathrm{ML}}\right)^2 \tag{62}$$

$$\left(\widehat{\sigma_{B_i}^2}\right)^{\mathrm{ML}}(\hat{\boldsymbol{s}}, \boldsymbol{I}) = \frac{1}{L - \hat{r}_i} \sum_{\rho=\hat{r}_i+1}^{L} \left(g_i(\rho) - \hat{B}_i^{\mathrm{ML}}\right)^2. \tag{63}$$

From (30),

$$p(\boldsymbol{s}|\lambda) = \frac{\sqrt{\det \boldsymbol{A}}}{\sqrt{(2\pi\lambda^2)^M}} \exp\left\{-\frac{1}{2\lambda^2}\boldsymbol{s}^T\boldsymbol{A}\boldsymbol{s}\right\}. \tag{64}$$

This leads to

$$\begin{aligned}\hat{\lambda} &= \arg\max_{\lambda} p(\hat{\boldsymbol{s}}|\lambda)p(\lambda) \\ &= \arg\max_{\lambda} \left\{-\frac{M}{2}\ln\lambda^2 - \frac{1}{2\lambda^2}\hat{\boldsymbol{s}}^T\boldsymbol{A}\hat{\boldsymbol{s}} - a\lambda^2\right\}.\end{aligned} \tag{65}$$

Solving for λ^2 leads to

$$\widehat{\lambda^2} = \frac{-M + \sqrt{M^2 + 8a\hat{\boldsymbol{s}}^T\boldsymbol{A}\boldsymbol{s}}}{4a}. \tag{66}$$

The second derivative, at the point given by (66), is negative proving that it is in fact a maximum.

Step 2) Perform one ICM iteration using the present contour parameter estimates. One iteration of ICM is a full sweep over all the contour sites.

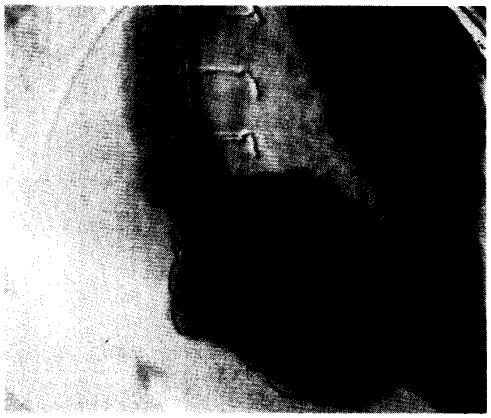


Fig. 12. Left ventricle ML contour estimate used to initialize the AICM algorithm.

Fig. 13. Contour obtained after 9 iterations of AICM.

Step 3) Check stopping condition; if not met, go back to Step 1). In our implementation, the stopping condition is

$$\sum_{i=1}^{M} |s_i(t+1) - s_i(t)| < \mu \tag{67}$$

where μ is the *stopping threshold.*

VI. Implementation and Results

The algorithm above described was implemented on a SIEMENS DIGITRON digital angiography equipment using Intel's FORTRAN-86 and Siemens User's Library. In all of the following examples the contours are defined by 100 points, on 512×512 pixels, 10 bits per pixel, images. Parameter a was set to 10^8 and the stopping condition constant μ was set to 1.

The program starts by asking the user to define the contour extrema and the approximate position of center of the ventricle. From these 3 points the coordinate system can be set up and all the related parameters (R_0, R_{M+1}, $\theta_{\min}$, $\theta_{\max}$, and ξ) computed. The initialization procedure, including image selection by the user, takes about 30 s. The computing time per iteration is about 0.5 s; convergence has always been attained after a maximum of $8 \sim 10$ iterations ($4 \sim 5$ s).

Fig. 10 presents an initial maximum likelihood contour estimate of a left ventricle. Given the good quality of this image the initial estimate is reasonable, with serious errors in just a few points. The result after 8 AICM iterations is presented in Fig. 11. In the example of Fig. 12, the initial estimate is very irregular. A serious error due to a surgical stitch is seen. Fig. 13 contains the contour estimate obtained after 9 iterations. Fig. 14 shows the evolution of $\widehat{\lambda^2}$ along the AICM iterations of the previous examples, showing the convergence behavior of the algorithm.

Two tests performed with synthetic images are next presented. In the first one, an image formed by an ellipse of intensity 150 over a background of intensity 100 was corrupted by AWGN with standard deviation $\sigma = 10$. The ML contour estimate, presented in Fig. 15, is good, and no changes occur after the first iteration, shown in Fig. 16. As the real values

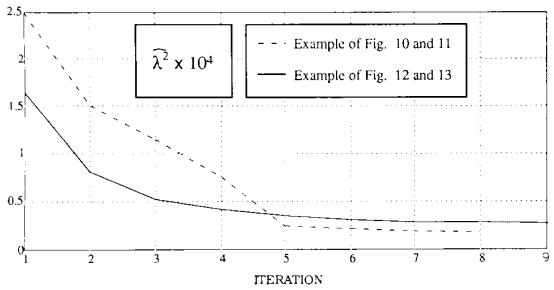


Fig. 14. Evolution of $\widehat{\lambda^2}$ along the AICM iterations that led to the estimates of Figs. 11 and 13.

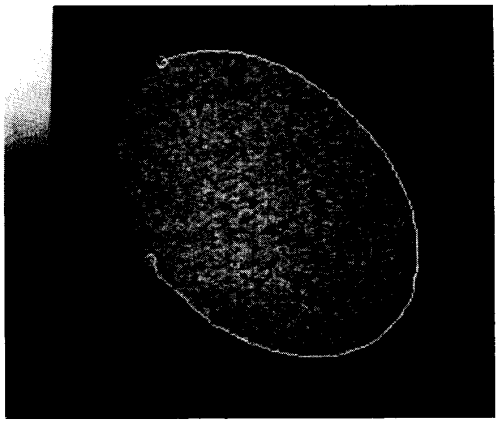


Fig. 15. Synthetic image (ellipse intensity = 150, background intensity = 100) corrupted by AWGN of standard deviation $\sigma = 10$. ML contour estimate used to initialize the AICM algorithm.

of the intensities and noise standard deviations are known, it is interesting to examine the estimated values of these parameters. In Fig. 17 the estimates $\hat{A}_i$ and $\hat{B}_i$ are represented along the contour sites $i = 1, \cdots, 100$. As expected, these estimates are distributed around the true values which are 150 and 100, respectively. Fig. 18 shows the estimates of σ_{A_i} and σ_{B_i}, distributed around the actual value $\sigma = 10$.

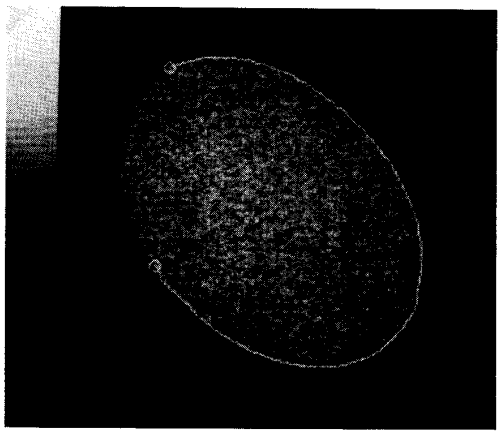


Fig. 16. Synthetic image (ellipse intensity = 150, background intensity = 100) corrupted by AWGN of standard deviation $\sigma = 10$. Contour estimate after just one iteration.

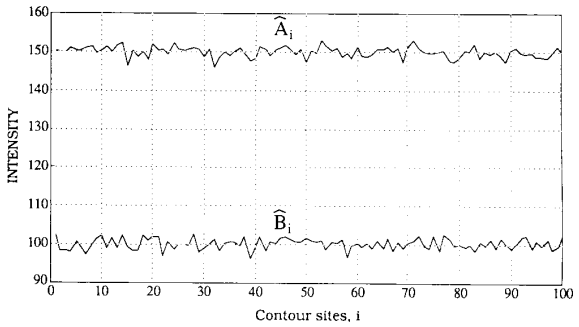


Fig. 17. Estimates of the mean intensities A_i and B_i along the contour sites of Fig. 16.

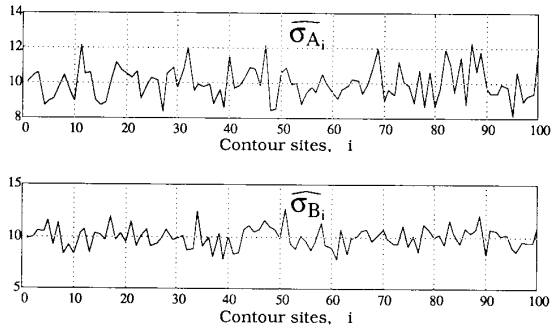


Fig. 18. Estimates of the standard deviations σ_{A_i} and σ_{B_i} along the contour sites of Fig. 16.

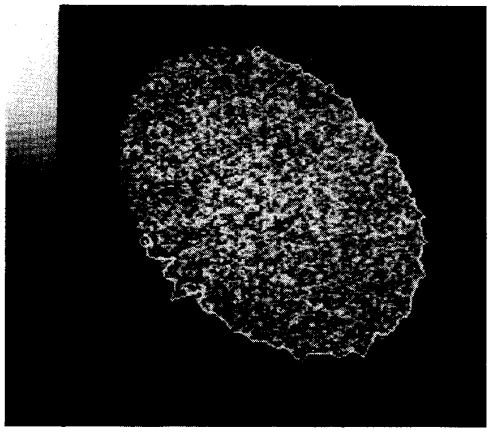


Fig. 19. Synthetic image corrupted by AWGN of standard deviation $\sigma = 25$. ML contour estimate used to initialize the AICM algorithm.

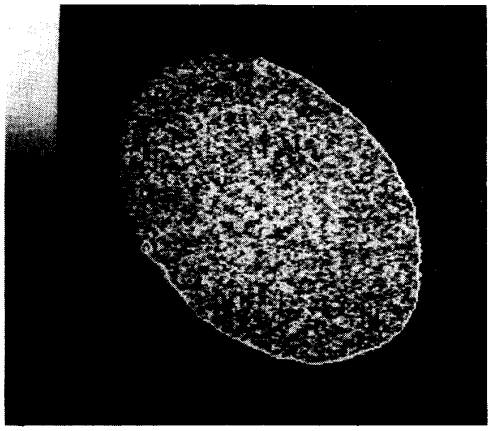


Fig. 20. Synthetic image corrupted by white Gaussian noise of standard deviation $\sigma = 25$. Contour obtained after 8 iterations of AICM.

The second test was performed with noise standard deviation $\sigma = 25$. Fig. 19 presents the ML contour estimate, showing some degree of irregularity. The contour resulting from 8 AICM iterations is shown in Fig. 20. The estimates $\hat{A}_i$ and $\hat{B}_i$, seen in Fig. 21, are distributed around the true values (150 and 100, respectively) but with greater dispersion than in the previous example. This is not surprising as it is well known that the standard deviation of the sample mean of a Gaussian variable is proportional to the standard deviation of the underlying distribution. Shown in Fig. 22 are the estimates of the noise variance along the contour, still distributed around the true value ($\sigma = 25$) but also with higher spread.

Finally, three more tests performed on real DSA images are presented in Figs. 23–25. Shown in Figs. 23(a), 24(a), and 25(a) are the ML estimates used to initialize the AICM algorithm. Figs. 23(b), 24(b), and 25(b) present the contours obtained after 4, 10, and 9 iterations, respectively. These examples witness the behavior of the algorithm; notice the catheter visible in all images and the faint contrast at the ventricle border.

VII. Conclusion

In this paper a new method for left ventricular contour determination in digital angiographic images was presented. The problem was formulated in a Bayesian estimation framework based on statistical models of prior knowledge and of the observation mechanism. The *a priori* smoothness constraint on the contour was expressed in terms of the Gibbs potentials of a one-dimensional noncausal Markov random field. The observation mechanism model was used to derive a likelihood

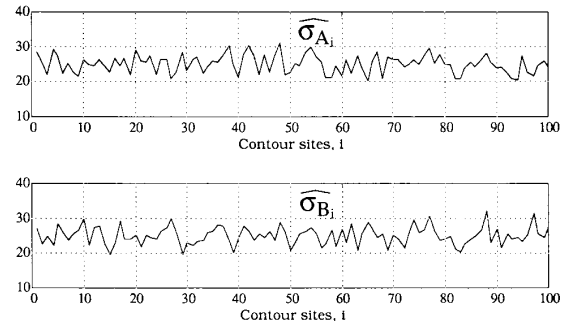


Fig. 21. Estimates of the mean intensities A_i and B_i along the contour sites of Fig. 20.

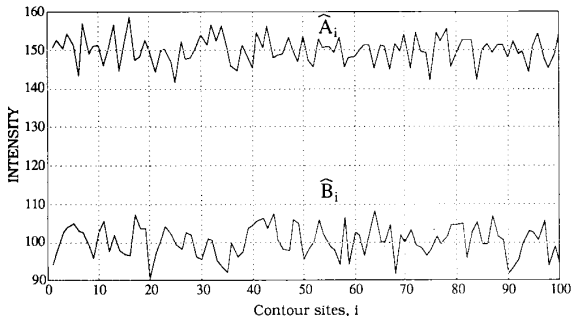
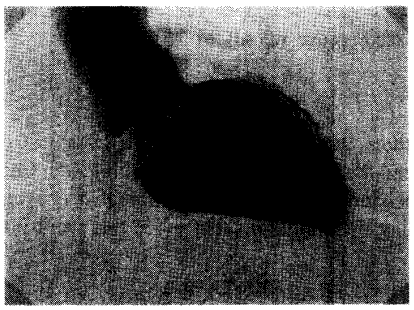
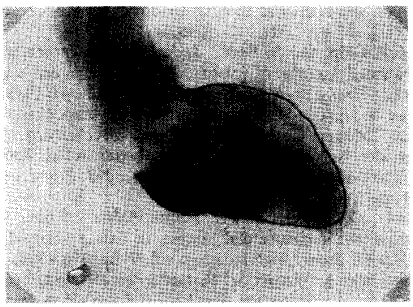


Fig. 22. Estimates of the standard deviations σ_{A_i} and σ_{B_i} along the contour sites of Fig. 20.

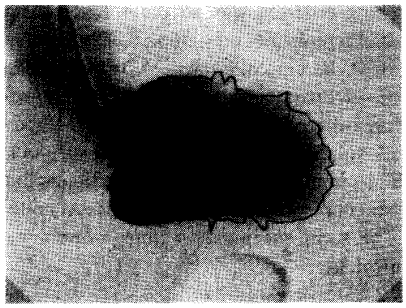


(a)

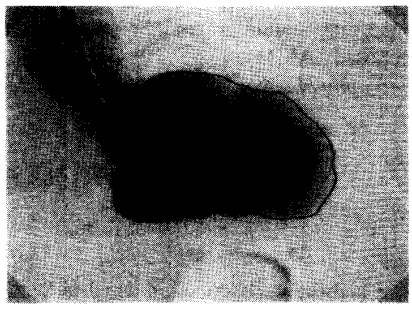


(b)

Fig. 23. (a) ML contour estimate. (b) Contour after 4 iterations of the AICM algorithm.

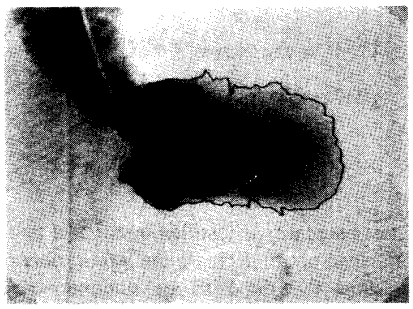


(a)

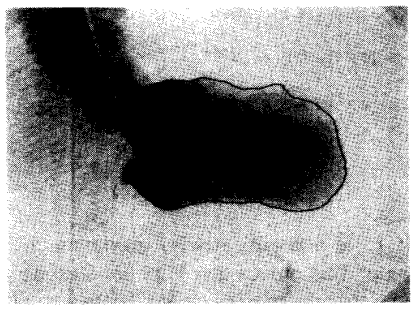


(b)

Fig. 24. (a) ML contour estimate. (b) Contour after 10 iterations of the AICM algorithm.



(a)



(b)

Fig. 25. (a) ML contour estimate. (b) Contour after 9 iterations of the AICM algorithm.

function of the boundary position along each radial section of the ventricle. This function was shown to be robust against model mismatch and image artifacts. Based on these two models (the Gibbs prior and the observation mechanism) a

MAP estimation criterion was adopted. The maximization of the *a posteriori* probability is implemented by an adaptive version of ICM, herein called AICM, that performs simultaneous contour and parameter estimation. The algorithm is thus completely data-driven. We proved that this method converges to a class of solutions called *partial optimal solutions*. The method exhibits immunity against image artifacts like catheters or surgical stitches and the results were found good by expert clinicians. The short computation times are compatible with routine clinical use.

APPENDIX

In this Appendix it is shown that *partial optimal solutions* (POS) are the only stationary points of the adaptive version of ICM proposed (AICM).

Let *partial optimal solutions* be defined as verifying

$$\Theta^* = \arg\max_{\Theta} p(\boldsymbol{s}^*, \boldsymbol{I}|\Theta, \lambda^*) \tag{A.1}$$

$$\lambda^* = \arg\max_{\lambda} p(\boldsymbol{s}^*, \boldsymbol{I}|\Theta^*, \lambda)p(\lambda) \tag{A.2}$$

$$\boldsymbol{s}^* = \arg\max_{\boldsymbol{s}} p(\boldsymbol{s}, \boldsymbol{I}|\Theta^*, \lambda^*). \tag{A.3}$$

The AICM algorithm, described in detail in Section V, can be rewritten as follows.

- Assume some initial estimates $\hat{\Theta}^{(0)}$, $\hat{\boldsymbol{s}}^{(0)}$, and $\hat{\lambda}^{(0)}$.
- At each iteration update the estimates according to

$$\hat{\Theta}^{(i+1)} = \arg\max_{\Theta} p\left(\hat{\boldsymbol{s}}^{(i)}, \boldsymbol{I}|\Theta, \hat{\lambda}^{(i)}\right) \tag{A.4}$$

$$\hat{\lambda}^{(i+1)} = \arg\max_{\lambda} p\left(\hat{\boldsymbol{s}}^{(i)}, \boldsymbol{I}|\hat{\Theta}^{(i)}, \lambda\right)p(\lambda) \tag{A.5}$$

$$\hat{\boldsymbol{s}}^{(i+1)} \text{ such that } p\left(\boldsymbol{s}^{(i+1)}, \boldsymbol{I}|\Theta^{(i)}, \lambda^{(i)}\right) \geq p\left(\boldsymbol{s}^{(i)}, \boldsymbol{I}|\Theta^{(i)}, \lambda^{(i)}\right) \tag{A.6}$$

where the last inequality is justified by the fact that each ICM step is guaranteed not to decrease the *a posteriori* probability.

We first show that the POS are stationary points of the algorithm. Assume that at some iteration n, the process is at a POS, i.e., $\left(\hat{\boldsymbol{s}}^{(n)}, \hat{\Theta}^{(n)}, \hat{\lambda}^{(n)}\right) = (\boldsymbol{s}^*, \Theta^*, \lambda^*)$ that verifies (A.1), (A.2), and (A.3). Then

$$\hat{\Theta}^{(n+1)} = \arg\max_{\Theta} p(\boldsymbol{s}^*, \boldsymbol{I}|\Theta, \lambda^*) = \Theta^* = \hat{\Theta}^{(n)} \tag{A.7}$$

$$\hat{\lambda}^{(n+1)} = \arg\max_{\lambda} p(\boldsymbol{s}^*, \boldsymbol{I}|\Theta^*, \lambda)p(\lambda) = \lambda^* = \hat{\lambda}^{(n)} \tag{A.8}$$

and

$$\hat{\boldsymbol{s}}^{(n+1)} = \boldsymbol{s}^* = \hat{\boldsymbol{s}}^{(n)} \tag{A.9}$$

because if $\boldsymbol{s}^*$ is a coordinatewise maximum, ICM can not go further up. This proves that POS are stationary points of AICM. Note that the maximizations relative to Θ and λ are global since exact ML estimates are computed; however, the maximum relative to $\boldsymbol{s}$ is coordinatewise (possibly local) since it is a result of ICM. Recall that ICM performs coordinate ascent on the posterior probability.

To prove that all stationary points of AICM are POS we start by assuming that the iterative process reaches a stationary point at some iteration n, i.e., $\hat{\Theta}^{(n+1)} = \hat{\Theta}^{(n)}$, $\hat{\lambda}^{(n+1)} = \hat{\lambda}^{(n)}$, and $\hat{\boldsymbol{s}}^{(n+1)} = \hat{\boldsymbol{s}}^{(n)}$. This can only be so if

$$\hat{\Theta}^{(n+1)} = \hat{\Theta}^{(n)} = \arg\max_{\Theta} p\left(\hat{\boldsymbol{s}}^{(n)}, \boldsymbol{I}|\Theta, \hat{\lambda}^{(n)}\right) \tag{A.10}$$

$$\hat{\lambda}^{(n+1)} = \hat{\lambda}^{(n)} = \arg\max_{\lambda} p\left(\hat{\boldsymbol{s}}^{(n)}, \boldsymbol{I}|\hat{\Theta}^{(n)}, \lambda\right)p(\lambda) \tag{A.11}$$

$$\hat{\boldsymbol{s}}^{(n+1)} = \hat{\boldsymbol{s}}^{(n)} = \arg\max_{\boldsymbol{s}} p\left(\boldsymbol{s}, \boldsymbol{I}|\hat{\Theta}^{(n)}, \hat{\lambda}^{(n)}\right) \tag{A.12}$$

where again the maximum w.r.t. the contour $\boldsymbol{s}$ is a coordinatewise (possibly local) maximum. Conditions (A.10), (A.11), and (A.12), being similar to (A.1), (A.2), and (A.3), define a POS.

Some resemblance with the well-known expectation–maximization (EM) algorithm (see [5] or [18]) can be found. EM is also an iterative and adaptive algorithm aiming at ML parameter estimation from incomplete data. In each iteration, the data are completed by computing the conditional expectation of the missing part based on the observed part and on the present parameter estimate (E step); the complete data is then used to compute a ML estimate of the parameters (M step). This process is shown to converge to the ML estimate of the parameters given the observed data [5], [18]. Interpreting the image as the observed data and the true contour as the missing data makes the relation between EM and AICM evident. There is, however, a fundamental difference in the fact that the contour (missing data) estimate produced at each iteration is not a conditional expectation. The conditional expectation nature of the E step is a fundamental part of the proof of convergence of EM; hence, this resemblance can not be used to prove convergence of AICM.

ACKNOWLEDGMENT

The authors would like to thank Dr. Seabra-Gomes, Dr. T. Real, and Dr. J. Baptista from Instituto do Coração and Hospital de Santa Cruz for all the valuable cooperation and for providing the conditions that made this work possible.

REFERENCES

[1] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Royal Statist. Soc. B*, vol. 36, pp. 192–225, 1974.

[2] ——, "On the statistical analysis of dirty pictures," *J. Royal Statist. Soc. B*, vol. 48, pp. 259–302, 1986.

[3] C. Chu, E. Delp, and A. Buda, "Detecting left ventricular endocardial and epicardial boundaries by digital two-dimensional echocardiography," *IEEE Trans. Med. Imaging*, vol. 7, pp. 81–90, 1988.

[4] M. Teles de Figueiredo, "Prior knowledge in ventricular contour determination," Masters thesis, Instituto Superior Téchnico, Lisbon, 1990.

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

[6] P. Detmer, G. Bashein, and R. Martin, "Matched filter identification of left-ventricular endocardial borders in transesophageal echocardiograms," *IEEE Trans. Med. Imaging*, vol. 9, pp. 396–404, 1990.

[7] H. Fahimi and A. Macovski, "Reducing the effects of scattered radiation in X-ray projection imaging," *IEEE Trans. Med. Imaging*, vol. 8, pp. 56–63, 1989.

[8] N. Friedland and D. Adams, "Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing," *IEEE Trans. Med. Imaging*, vol. 8, pp. 344–353, 1989.

[9] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, 1984.

[10] S. Geman and D. McClure, "Statistical methods for tomographic image reconstruction," in *Proc. 46th Session Int. Statist. Inst.,* Bulletin of the ISI, vol. 52, pp. 353–356, 1987.
[11] P. Grattoni and R. Bonamini, "Contour detection of the left ventricular cavity from angiographic images," *IEEE Trans. Med. Imaging,* vol. MI-4, pp. 72–78, 1985.
[12] P. Green, "Bayesian reconstruction from emission tomography data using a modified EM algorithm," *IEEE Trans. Med. Imaging,* vol. MI-9, pp. 84–93, 1990.
[13] A. Jain, *Fundamentals of Digital Image Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1989.
[14] S. Kirkpatrick, C. Gelatt, and M. Vecchi, "Optimization by simulated annealing," *Sci.,* vol. 220, pp. 671–680, 1983.
[15] R. Kruger, "X-Ray digital cineangiography," in *Cardiac Imaging and Image Process,* S. Collins and D. Skorton, Ed. McGraw-Hill, New York: 1986, pp. 206–238.
[16] S. Lakshmanan and H. Derin, "Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 11, pp. 799–813, 1989.
[17] P. Lilly, J. Jenkins, and P. Bourdillon, "Automatic contour definition on left ventriculograms by image evidence and a multiple template-based model," *IEEE Trans. Med. Imaging,* vol. 8, pp. 173–185, 1989.
[18] R. Little and D. Rubin, *Statistical Analysis with Missing Data.* New York: Wiley, 1987.
[19] D. Luenberger, *Linear and Nonlinear Programming.* Reading, MA: Addison-Wesley, 1984, 2nd ed.
[20] E. Marcus, P. Lorente, E. Bartha, R. Beyar, D. Adam, and S. Sideman, "A comparative study of quantitative methods for characterization of left ventricular contraction," in *Proc. Comput. Cardiol.—1985,* pp. 145–148.
[21] J. Marroquin, S. Mitter, and T. Poggio, "Probabilisitic solution of ill-posed problems in computacional vision," *J. Amer. Statist. Assoc.,* vol. 82, pp. 76–89, 1987.
[22] T. Pappas and J. Lim, "A new method for estimation of coronary artery dimensions in angiograms," *IEEE Trans. Acoust., Speech, Signal Processing,* vol. 36, pp. 1501–1513, 1988.
[23] O. Ratib and A. Righetti, "Multiharmonic fourier analysis of ventricular wall motion for the detection of regional alterations in contraction and relaxation," in *Proc. Comput. Cardiol.—1983,* pp. 423–426.
[24] J. Reiber, P. Serruys, and C. Slager, *Quantitative Coronary and Left Ventricular Cineangiography.* Dordrecht, The Netherlands: Martinus Nijhoff Publishers, 1986.
[25] A. Rickards, R. Seabra-Gomes, and P. Thurston, "The assessment of regional abnormalties of the left ventricle by angiography," *Europ. J. Cardiol.,* no. 5, 1977.
[26] R. Seabra-Gomes, "Left ventricular segmental analysis—Methods and clinical applications," in *VI Europ. Congr. Cardiol.,* 1987.
[27] L. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emisison tomography," *IEEE Trans. Med. Imaging,* vol. MI-1, pp. 113–122, 1982.
[28] A. Thompson, J. Brown, J. Kay, and D. Titterington, "A study of methods of choosing the smoothing parameter in image restoration by regularization," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 13, pp. 326–339, 1991.
[29] K. Toraichi, K. Katagishi, and R. Mori, "A left ventricular function analyzer and its application," *IEEE Trans. Biomed. Eng.,* vol. BE-34, pp. 317–328, 1987.
[30] M. Unser, G. Pelle, P. Brun, and M. Eden, "Automated extraction of serial myocardial borders from M-mode echocardiograms," *IEEE Trans. Med. Imaging,* vol. 8, pp. 96–103, 1989.
[31] H. Wollschaeger, R. Waltenspiel, U. Solzbach, A. Zeiher, and H. Just, "Reliable automatic frame-by-frame contour detection of digitized L.V. cine-angiograms," in *Proc. Comput. Cardiol.—1988,* pp. 353–356.
[32] W. Wong, R. Kirkeeid, and L. Gould, "Computer applications in angiography," in *Cardiac Imaging and Image Processing,* S. Collins and D. Skorton, Ed. New York: McGraw-Hill, 1986, pp. 206–238.
[33] J. Woods, "Two-dimensional discrete Markovian fields," *IEEE Trans. Inform. Theory,* vol. IT-18, pp. 232–240, 1972.