Model-based Deformable Surface Finding for Medical Images

Lawrence H. Staib,* *Member, IEEE,* and James S. Duncan, *Senior Member, IEEE*

Abstract— This paper describes a new global shape parametrization for smoothly deformable three-dimensional (3D) objects, such as those found in biomedical images, whose diversity and irregularity make them difficult to represent in terms of fixed features or parts. This representation is used for geometric surface matching to 3D medical image data, such as from magnetic resonance imaging (MRI). The parametrization decomposes the surface into sinusoidal basis functions. Four types of surfaces are modeled: tori, open surfaces, closed surfaces and tubes. This parametrization allows a wide variety of smooth surfaces to be described with a small number of parameters. Extrinsic model-based information is incorporated by introducing prior probabilities on the parameters. Surface finding is formulated as an optimization problem. Results of the method applied to synthetic images and 3D medical images of the heart and brain are presented.

I. Introduction

THIS work describes an approach to finding object boundaries in three-dimensional (3D) medical images based on parametrically deformable shape models using prior shape information. Boundary finding is enhanced both by considering the bounding surface as a whole and by using model-based shape information.

Boundary finding using only local image information has often been frustrated by poor-contrast boundary regions due to objects in close proximity, adverse viewing conditions, noise and imaging artifacts. Imperfect image data can be augmented with the extrinsic information that a geometric shape model provides. The bounding surface of an object in three dimensions can be profitably considered as a whole, rather than as surface segments, because it tends to result in a more consistent solution overall. These models are best suited for objects whose diversity and irregularity of shape make them poorly represented in terms of fixed features or parts. The objects are expected to have a tendency toward some average shape with a continuum of possibilities near that shape. This kind of information is available whenever there is a sample of images of the structure of interest available, as is true for many research and clinical applications. Such smoothly deformable objects do not necessarily have an obvious decomposition that can be exploited. A uniform mathematical shape representation that describes the entire shape is therefore needed and it should describe a broad class of shapes.

L. Staib and J. Duncan are with the Departments of Diagnostic Radiology and Electrical Engineering, Yale University, 333 Cedar Street, New Haven, CT, 06520-8042 USA. (e-mail: lawrence.staib@yale.edu).

For a representation to be useful for modeling it should be concise. Methods based on explicitly listing points or patches on the surface are verbose because of the implicit redundancy. Parametric representations capture the overall shape in a small number of parameters. This means that the optimization of a match measure between data and a model can occur in a lower dimensional space. Parametric representations allow the direct analytic computation of object properties, such as surface area or curvature. This form of representation also allows the incorporation of prior probabilistic information.

In this work, boundary finding is formulated as an optimization problem using parametric Fourier models of surfaces. The model is matched to the image by optimizing in the parameter space the match between the model and a boundary measure applied to the image. Probability distributions on the parameters of the representation are incorporated to bias the model to a particular overall shape while allowing for deformations. This leads to a maximum *a posteriori* objective function. This work builds on previous work on a model-based system for the analysis of curves in two-dimensional images [1].

II. Related Work in Boundary Finding

Local edge detectors applied to real images produce spurious edges and gaps. These problems can be overcome only by the incorporation of information from higher scale organization of the image and models of the objects sought. Contextual information has been used for boundary determination via grouping, relaxation labeling and scale-space methods. These methods, by themselves, will not necessarily find complete boundaries. Pixel search methods, which find an optimal path through a two-dimensional image, do not generalize obviously to three dimensions because there is no natural ordering of voxels in a surface. An alternative method for boundary analysis is the Hough transform [2]. The Hough approach is similar to the current method in that it finds shapes by looking for maxima in a parameter space. However, the storage and computational complexity of the Hough method are a great disadvantage, especially if deformations are envisaged.

Other investigators have considered whole-boundary methods that adjust a tentative surface mesh in order to match to the image. By considering the boundary as a whole, a structure is imposed on the problem that bridges gaps and results in overall consistency. These active surfaces are a direct generalization of "snakes"[3] and have been used for finding surfaces in 3D images [4], [5], [6]. The parameters of mesh-based representations are coordi-

nates and are thus extremely local and ultimately flexible but of limited expressive power.

Snakes that can adjust their topology have been developed using level sets [7] and have been extended to surfaces [8], [9]. This allows an even greater flexibility for general boundary finding but is not necessary when specific prior information is known.

Other whole-boundary methods optimize in a parameter space. Parametric representations are useful for modeling because they capture the overall shape concisely. This means that the optimization of a match measure between data and a model can occur in a lower dimensional space. Widrow [10], and later Yuille et al. [11], used parametrized templates to model objects where the parameters are sizes and relationships between simple subparts. These methods describe the overall shape of the structure using very few parameters. However, the object must have sufficient structure to be represented in terms of parts and a new model must be developed for each new object. Work has also been done developing deformable templates based on Markov models of curves incorporating knowledge of shape from statistical features [12]. In the next section, we will discuss parametrizations for surfaces in more detail.

Pentland and others [13], [14] have developed physically-based methods for analyzing shape. Shapes are represented by the low-order frequency displacement eigenvectors corresponding to the free vibration modes of the object. This is a Fourier-type representation using computed basis functions that are derived from a physical analogy. While such analogous models may be appropriate for motion computations, it is not clear that this computation is worthwhile in boundary finding where the physical analogy is less apparent. In fact, recent work suggests that even a small amount of statistical information results in an improvement over modal analysis [15].

A number of methods have been developed specifically to identify structures from medical images. For example, Bomans et al. [16] use a boundary-finding method based on a 3D version of the Marr-Hildreth edge operator to find surfaces in MR brain images. Morphological operators are used to remove small holes and thin connections. Similarly, Higgins et al. [17] developed a semi-automated method for extracting the endocardium from high-speed 3D computed tomography (CT) cardiac images. This method uses a combination of manual editing, mathematical morphology, maximum-homogeneity filtering and adaptive 3D thresholding. A number of techniques for segmentation rely on clustering techniques for voxel classification. Cline et al. [18] use multispectral voxel classification, in conjunction with connectivity, to segment 3D MR brain images. This method is limited by the assumption of normality in the probability distributions of the tissues. Interactive and semi-automated methods are a compromise between hand tracing and fully automated methods. Kennedy et al. [19] describe a number of semi-automated methods for segmenting MR images of the brain. Hohne and Hanson [20] use mathematical morphology, connected components and thresholding to interactively segment 3D images with feedback from rendered displays.

The advantage of incorporating prior information into the boundary finding process has begun to be recognized. Collins et al. [21] segment the brain using an elastic registration to an average brain based on a hierarchical local correlation. The average brain provides strong prior information about the expected image data. Cootes et al. [22] augment a snake-like model with statistics to model structures in medical images in order to locate them. They represent objects using a principle component analysis of a sample of shapes. They also include information about the expected gray levels around model points as additional prior information. This approach, because it relies on coordinates, requires consistent sampling and labeling of point features. In addition, possible deformations are strictly limited according to the statistics of the training samples. The use of principle components has also been adapted for Fourier representations by Szekély et al. [23].

III. Surface Representations

Shape parametrizations facilitate the application of geometric models to the problem of boundary finding by allowing the formation of a mathematical objective function to be optimized. Other kinds of models incorporate implicit or explicit geometric constraints, such as generalized cylinders [24], but do not lend themselves to a mathematical formulation. An arbitrary surface can be represented explicitly by three functions of two index parameters: $x(u,v)$, $y(u,v)$ and $z(u,v)$. A surface is indexed by the two parameters (u,v). While a curve's points are naturally ordered by arclength, there is no natural ordering of points on an arbitrary surface.

Second degree algebraic surfaces have been used extensively as parametric models because of their simplicity and conciseness. Ellipses and ellipsoids have been particularly popular, and used, for example, in the work of Blokland et al. to model the heart [25]. Their conciseness, however, greatly limits their expressiveness, although extensions are possible. Nuyts et al. [26] forms a piece-wise elliptical model of the heart by using a model whose radial cross-sections consist of two elliptic curves. Higher order polynomial surfaces can be expressed using implicit representations [27].

Superquadrics are an extension of quadrics using an exponent that allows the shape to vary from an ellipsoid to a rectangular parallelepiped [28]. The basic shape can be altered by such operations as twisting, bending and tapering [29], as can any explicit representation. The main disadvantage of superquadrics is that even with these altering operations, superquadrics are limited by their doubly symmetric cross-section and thus still only represent a very limited family of shapes, without resorting to composition. Superquadrics have also been augmented by deformations according to a variety of models. In spline model deformations [30], [31], the superquadric itself is reduced to a means for initialization and the refinement is modeled by

the deformation in a manner similar to other snake surface techniques. Deformations based on strain modes [32] (see discussion above) and wavelet bases [33] have also been introduced. Wavelets are another very promising representation in that they have the ability to range from local to global description.

Spherical harmonics have been used as a type of surface representation for radial or stellar surfaces $(r(\theta, \phi))$ [2], [34]. The surface is represented as a weighted sum of spherical harmonics which are orthogonal over the sphere. This is a type of Fourier representation, as defined below, but is restricted to radial deformations of a sphere, and thus models a limited class of objects. Recently, however, spherical harmonics have been extended to more general shapes by representing a surface using three functions of θ and ϕ [35]. An appropriate mapping is determined by iteratively solving a constrained optimization problem based on the diffusion equation. This representation can then be used in deformable boundary finding [23]. However, the mapping does not maintain this optimal form during the deformation process without a recomputation of the above optimization at each iteration.

IV. FOURIER SURFACES

Smoothly deformable biomedical objects do not necessarily have an obvious decomposition that can be exploited. A uniform shape representation that describes the entire shape is therefore needed and it should describe a relatively broad class of shapes.

The prior information available is a flexible bias towards more likely shapes. This sort of model can be achieved by using a generic parametrization with probability distributions defined on the parameters. That is, the parametrization itself is expressive enough to represent any potential shape of a given geometric type, but the associated probability distributions introduce a bias towards an expected range of shapes. The spread in the distributions is due to variability among instances of the object. This kind of parametrization represents a stronger use of prior information than methods that use only simple shape characteristics.

Fourier representations are those that express the function in terms of an orthonormal basis. The motivation for a basis representation is that it allows us to express any object as a weighted sum of a set of known functions. An orthonormal set is desirable because it makes the parameters (weights) distinct. For example, to express the one-dimensional function $f(t)$ on the interval (a, b) in terms of the basis $\phi_k(t)$, we use

$$f(t) = \sum_{k=1}^{\infty} p_k \phi_k(t) \qquad \text{where} \qquad p_k = \int_a^b f(t)\phi_k(t)\, dt. \quad (1)$$

The coefficients p, the projections of the function onto the k basis functions, are the parameters of the representation. In order to use this representation the sum is truncated. In most such representations, the higher indexed basis functions represent higher spatial variation. Therefore, if the function to be represented is expected to have limited spatial variation, as is the case for most real object boundaries, the series can be truncated and still accurately represent the function. However, objects of arbitrary complexity can be represented by including sufficiently high harmonics. Because it is desirable to use a concise representation, the minimum number of harmonics necessary to represent a class of shapes can be determined given an error tolerance and a sample of the shapes to be modeled. The number of harmonics can be decreased until the average error between the true boundary and the boundary from the truncated representation exceeds the error tolerance.

The usual basis functions are the sinusoids [36], although others, such as orthogonal polynomials or spherical harmonics in two dimensions, are possible. The sinusoids have the advantage of representing the familiar notion of frequency. A limitation of this choice is that it does not represent sharp transitions or very complex boundaries very well because this requires a large number of parameters.

A surface in three dimensions can be described explicitly by three coordinate functions of two surface parameters, $\mathbf{x}(u, v) = (x(u, v), y(u, v), z(u, v))$, where u and v vary over the surface. Thus, in order to represent surfaces, a basis for functions of two variables is needed; the following can be used [37]:

$$\phi = \{1, \cos mu, \sin mu, \cos lv, \sin lv, \cos mu \cos lv,$$
$$\sin mu \cos lv, \cos mu \sin lv, \sin mu \sin lv, \ldots$$
$$(m, l = 1, 2, \ldots) \quad \}. \quad (2)$$

Each function is then represented by

$$f(u, v) =$$
$$\sum_{m=0}^{K_1-1} \sum_{l=0}^{K_2-1} \lambda_{m,l} \left[a_{m,l} \cos mu \cos lv + b_{m,l} \sin mu \cos lv + \right.$$
$$\left. c_{m,l} \cos mu \sin lv + d_{m,l} \sin mu \sin lv \right] \quad (3)$$

where

$$\lambda_{m,l} = \begin{cases} 1 & \text{for} \quad m = 0,\, l = 0 \\ 2 & \text{for} \quad m > 0,\, l = 0 \quad \text{or} \quad m = 0,\, l > 0 \\ 4 & \text{for} \quad m > 0,\, l > 0 \end{cases}$$

and the series is truncated at $K_1 - 1$ and $K_2 - 1$ terms. This basis allows the specification of even functions using the cosine terms and odd functions using the sine terms. There are three sets of parameters corresponding to the three coordinate functions: $\mathbf{p}_x = \{a_x, b_x, c_x, d_x\}$, $\mathbf{p}_y = \{a_y, b_y, c_y, d_y\}$, $\mathbf{p}_z = \{a_z, b_z, c_z, d_z\}$. These will be referred to collectively as the parameter vector $\mathbf{p} = \{\mathbf{p}_x, \mathbf{p}_y, \mathbf{p}_z\}$.

The complex basis is useful for computational purposes because the parameters can be computed in a single transform:

$$\phi = \{1, e^{i(mu+lv)}, \ldots \quad (m, l = \pm 1, \pm 2, \ldots) \quad \}. \quad (4)$$

Each function is then represented by

$$f(u, v) = \sum_{m=-(K_1-1)}^{K_1-1} \sum_{l=-(K_2-1)}^{K_2-1} g_{m,l}\, e^{i(mu+lv)}. \quad (5)$$
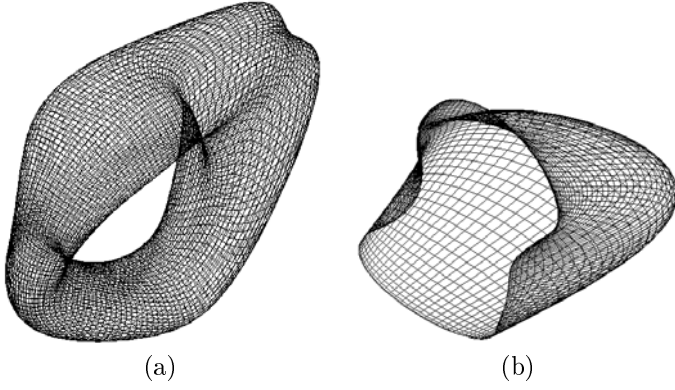
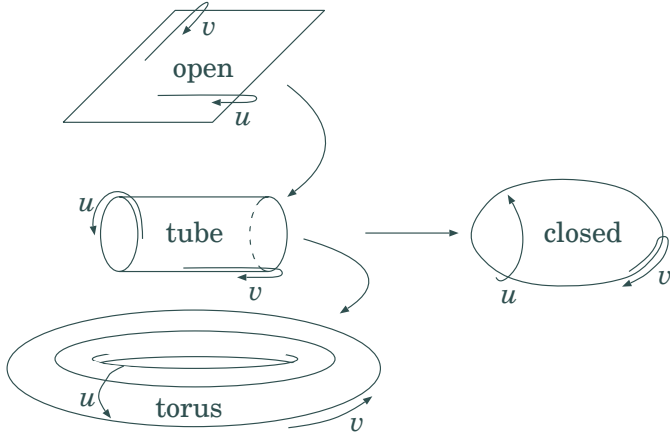Fig. 1. An example torus surface (a) and an example open surface (b) using up to fourth order harmonics.

Fig. 2. Surface types: This shows the relationship between the surface types and the surface parameters u and v.

Using Euler's formula, $e^{ix} = \cos x + i \sin x$, we can derive the conversion between the sine-cosine basis parameters and the complex basis parameters.

There are four classes of simple surfaces in three dimensions that will be described: tori (closed tubes), open surfaces (with one edge), tubes (open surfaces with two edges) and closed surfaces (no edges). The torus is formed using the entire basis shown in Equation 2. The result is a torus because both surface parameters are forced to be periodic. An example torus surface using this parametrization is shown in Fig. 1. The other three types of surfaces can be described using subsets of the above basis which flatten out or constrain the torus in different ways as described below and shown schematically in Fig. 2.

A. Open Surfaces

Representing open surfaces with the basis in Equation 2 is complicated by the periodicity property. Since the surface is open, a straightforward representation of the surface would result in discontinuities at the boundary. Thus, these discontinuities can be avoided by having the two surface parameters start at one side of the surface, trace along the surface to the other end, and then exactly retrace the surface in the opposite direction to create a closed path.

This is analogous to the representation of open curves [36], [1]. Note that while the curve traced by the surface parameter has a slope discontinuity at the ends where it retraces its path, the corresponding coordinate function has no such discontinuity because the parameter slows down to a stop as it approaches the end ($\frac{\partial x(0)}{\partial u} = 0$, etc.).

This results in a function $x(u, v)$ that is even and thus only the purely even terms, $a_{x,0,0}$, $a_{x,m,0}$, $a_{x,0,l}$ and $a_{x,m,l}$ are nonzero. This also holds for $y(u, v)$ and $z(u, v)$. The converse is also true; that is, any expansion (with any number of terms) with only those terms nonzero for all l and m results in an even function and thus describes an open surface. We are therefore effectively restricting the basis to include only even functions of both u and v:

$$\phi_{\text{open}} = \{1, \cos mu, \cos lv, \dots, \cos mu \cos lv, \dots$$
$$(m, l = 1, 2, \dots)\}. \quad (6)$$

Open surfaces are useful for a wide variety of structures, including objects with one prominent opening, the bounding surface between two touching objects, and flat objects. An example open surface using this parametrization is shown in Fig. 1.

B. Tube Surfaces

Tubes require the open representation along one of the surface parameters and the closed representation along the other. This results in the following basis which is even in v and unrestricted in u:

$$\phi_{\text{tube}} = \{1, \cos lv, \sin mu, \cos mu, \dots, \cos mu \cos lv,$$
$$\sin mu \cos lv, \dots \quad (m, l = 1, 2, \dots)\}. \quad (7)$$

Thus, the only nonzero terms are $a_{x,0,0}$, $a_{x,0,l}$, $a_{x,m,0}$, $b_{x,m,0}$, $a_{x,m,l}$ and $b_{x,m,l}$ and the corresponding y and z terms. Tubes are an extension of generalized cylinders [24] but in parametric form.

Tubes are useful for elongated hollow objects and elongated objects with flat ends. They are also useful for temporal sequences of planar images, where the third dimension is time, and multimodal images, where the third dimension is modality. A spatiotemporal example using a tube model is shown in Fig. 15. Two example tube surfaces using this parametrization are shown in Fig. 3.

C. Closed Surfaces

Closed surfaces are the most difficult to represent because they are most dissimilar to tori. One way to represent closed surfaces is by considering tubes whose ends close up to a point at both ends instead of being open. This is done by using the basis

$$\phi_{\text{closed}} = \{1, \sin lv, \dots, \cos mu \sin lv, \sin mu \sin lv, \dots$$
$$(m, l = 1, 2, \dots)\} \quad (8)$$

thus forcing both functions to constants at $v = 0, \pi, 2\pi$. This, however, forces the ends together too. This means
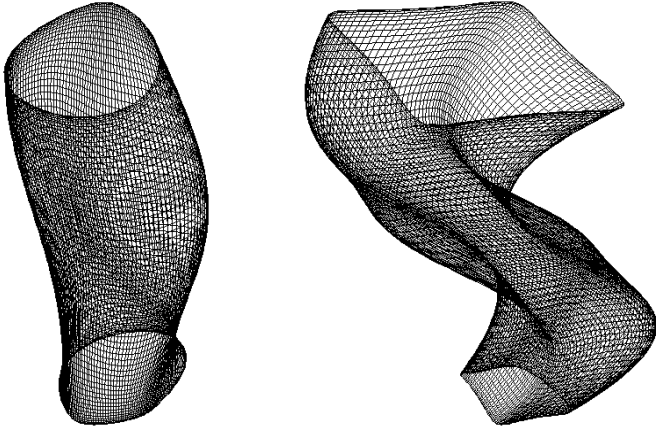
Fig. 3. Two tube surface examples using up to fourth order harmonics.
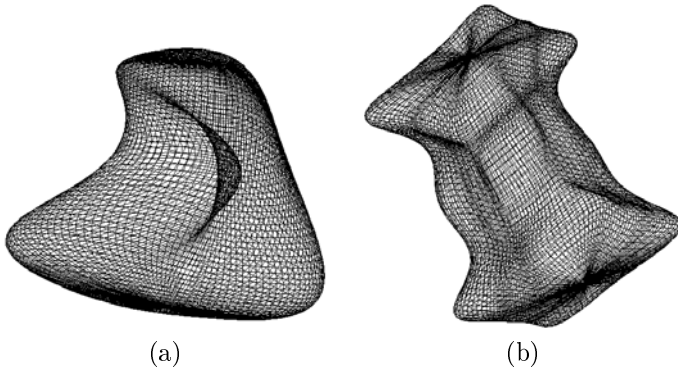

(a)　　　　　　　　　　(b)

Fig. 4. Two closed surface examples using up to fourth order (a) and eighth order (b) harmonics.

that the ends must be separated by adding a weighted term to each coordinate of the form: $\sin(v - \pi/2)$. These three weights are three additional shape parameters.

This representation requires that the values for v be repeated as for an open curve but negated so that

$$x(u,v) + x(u, 2\pi - v) = 2x(0,0),$$
$$y(u,v) + y(u, 2\pi - v) = 2y(0,0),$$
$$z(u,v) + z(u, 2\pi - v) = 2z(0,0). \quad (9)$$

Two examples of closed surfaces using this parametrization are shown in Fig. 4, with terms up to fourth order on the left and eighth order on the right.

D. Surface geometry

The Fourier surface description makes the calculation of geometric surface properties straightforward because a continuous description of the surface is known. Without an analytic description of the surface, curvature can be calculated based on the computation of derivatives from a local surface patch fit, or from a discrete approximation of the derivatives at each point. These methods are dependent on the proper choice of the size of the patch or neighborhood. For Fourier surfaces, partial derivatives of the surface functions can be calculated from the functional description. Curvature is then calculated directly from these
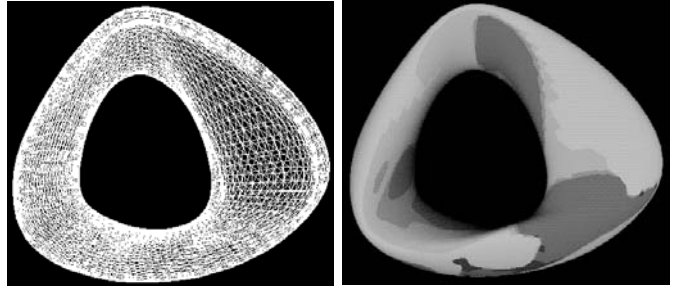

Fig. 5. Curvature calculation example. Wire frame and curvature shown from the surface shown in Figure 1 (a). Bright indicates peaked areas, dark indicates pits and gray indicates saddle areas as determined by the Gaussian and mean curvatures.

partial derivatives [38]. In addition, using the complex basis of Equation 4, the partials have a relatively simple form. For example,

$$\mathbf{x}_u = \frac{\partial x}{\partial u} = \sum_{m=-(K_1-1)}^{K_1-1} \sum_{l=-(K_2-1)}^{K_2-1} im\mathbf{g}_{ml}e^{i(mv+lu)} = \mathcal{F}^{-1}\{im\mathbf{g}\}, \quad (10)$$

where $\mathcal{F}$ is the discrete Fourier transform, $\mathbf{g} = (g_x, g_y, g_z)$ and $\mathbf{x} = (x, y, z)$. Similarly for $\mathbf{x}_v$ and the second partial derivatives,

$$\mathbf{x}_v = \mathcal{F}^{-1}\{il\mathbf{g}\}, \quad \mathbf{x}_{uu} = \mathcal{F}^{-1}\{-m^2\mathbf{g}\}, \quad (11)$$
$$\mathbf{x}_{vv} = \mathcal{F}^{-1}\{-l^2\mathbf{g}\}, \mathbf{x}_{uv} = \mathcal{F}^{-1}\{-ml\mathbf{g}\}. \quad (12)$$

At the edges of tubes and open surfaces and at the poles of the closed surfaces, these formulas break down and the calculation must be treated as a special case.

From these partial derivatives, the surface geometry can be characterized. Gaussian (K) and mean (H) curvature can be computed from these derivatives [38]. Surface curvature properties have been used to classify and characterize shape. For example, surface regions can be classified by the sign of the surface curvatures as peaks, ridges, saddles, valleys, pits and flats. An example curvature calculation based on the Gaussian and mean curvatures is shown in Fig. 5. This type of characterization is useful for shape matching for deformation tracking [39] and shape classification for determining normal and pathologic variation in anatomic shape.

V. Boundary Finding Objective Function

In order to fit one of these models to the image data, a measure of fit is optimized by varying the model parameters. The surface is expected to be distinguishable by some measure of boundary strength (direction can also be used) computed from the image. The sum or integral of the boundary strength image over a given surface indicates the degree of correspondence between them.

Any measure that indicates a change in some property that distinguishes the object from the background could be used as a boundary measure. A natural candidate for many

images is the gray-level gradient. The magnitude is the strength of the boundary and the direction is the normal to the boundary. The gray-level gradient can be calculated by first smoothing with a Gaussian to reduce the effect of noise. The smoothed boundary response will also help in the optimization by attracting the surface from further away. An alternative would be to match to the distance transform of an edge map (such as from Canny's method [40]) calculated from the image.

A. Probabilistic Formulation

In addition to the constraint that the parametric model provides, it is useful to incorporate prior information in the form of probabilities to further constrain the problem. In order to incorporate probabilistic information into the measure of fit, consider the problem of boundary determination as one in which the data is a 3D image, $b(\mathbf{x})$, which could depict the boundary of any object in the parametric representation, and $\mathbf{p}$ is a particular value of the parameter vector. In terms of probabilities, if we want to decide which parameter vector, $\mathbf{p}$, an image, b, corresponds to, we need to evaluate the probability of the template given the image, $\Pr(\mathbf{p}|b)$, and find the maximum over $\mathbf{p}$. This can be expressed using Bayes rule, where

$$\mathbf{p}_{\mathrm{map}} = \arg\max_{\mathbf{p}} \Pr(\mathbf{p}|b) = \arg\max_{\mathbf{p}} \frac{\Pr(b|\mathbf{p})\Pr(\mathbf{p})}{\Pr(b)}. \quad (13)$$

Here, $\mathbf{p}_{\mathrm{map}}$ is the maximum *a posteriori* solution, $\Pr(\mathbf{p})$ is the prior probability of the parameter vector $\mathbf{p}$ and $\Pr(b|\mathbf{p})$ is the conditional probability, or likelihood, of the image given the parameter vector. This expression can be simplified by taking the logarithm and eliminating $\Pr(b)$, the prior probability of the image data, which is equal for all $\mathbf{p}$:

$$\mathbf{p}_{\mathrm{map}} = \arg\max_{\mathbf{p}} M(b, \mathbf{p}) = \arg\max_{\mathbf{p}} \left[\ln \Pr(\mathbf{p}) + \ln \Pr(b|\mathbf{p})\right]. \quad (14)$$

This maximum *a posteriori* objective function shows the tradeoff or compromise that is made between prior information, $\Pr(\mathbf{p})$, and image-derived information, $\Pr(b|\mathbf{p})$. For a uniform prior, this formulation reduces to the maximum likelihood solution.

In order to derive the expression for the likelihood, consider the image b to be a noise-corrupted version of an image template, $t_{\mathbf{p}}$, corresponding to a particular value of the parameter vector with noise that is independent and additive: $b = t_{\mathbf{p}} + n$. This assumption avoids an excessive increase in complexity. Furthermore, Cooper [41] showed, for a related problem, that this assumption did not alter the performance significantly. Then, the likelihood, $\Pr(b|\mathbf{p})$, is equivalent to $\Pr(n = b - t_{\mathbf{p}})$. The noise at each image point, $n(\mathbf{x})$, equals $b(\mathbf{x}) - t_{\mathbf{p}}(\mathbf{x})$ and is governed by the probability density $\Pr(n)$. These events are independent for each point, so the probability for the noise over the entire region $\mathcal{A}$ is just the product of the individual probabilities. The noise is the combined effect of many factors such as signal

degradation, artifacts and boundary measurement which are difficult to model explicitly. We make the assumption that the noise is Gaussian with zero mean and standard deviation σ_n.

The object template, $t_{\mathbf{p}}(\mathbf{x})$, is an image that represents the boundary of the object. The boundary can be embedded into the image template by making $t_{\mathbf{p}}(\mathbf{x})$ constant along the boundary of the object it represents and zero everywhere else. In order to match this template with the image b, consider $b(\mathbf{x})$ to be a boundary measure applied to the raw image data: $b(\mathbf{x}) = b(i(\mathbf{x}))$.

First, it is only necessary to sum along the boundary for terms involving the template, because the template is zero everywhere else. In addition, the magnitude of $t_{\mathbf{p}}(\mathbf{x})$ is taken to be constant (k), over the boundary that it defines. The function M can be simplified further by removing the terms that do not depend on $\mathbf{p}$, such as b^2 which is summed over the entire image. The details of this derivation follow directly from the two-dimensional version [1]. Then, the continuous version of this for a surface is

$$M(b, \mathbf{p}) = \ln \Pr(\mathbf{p}) + \frac{k}{2\sigma_n^2} \iint_{\mathcal{A}} [b(x(\mathbf{p}, u, v), y(\mathbf{p}, u, v), z(\mathbf{p}, u, v))] dA \quad (15)$$

where dA is an area element on the surface $\mathcal{A}$.

Equation 15 is the maximum *a posteriori* objective function for surface finding. The first term is the contribution of the prior probability of the parameter vector. The greater the variance of the prior, the smaller the influence of this term. The second indicates the degree of correspondence between the model and the image.

The probability distributions associated with the parameters are intended to bias the model towards a particular range of shapes. This prior knowledge comes from experience with a sample of images of the object being delineated, when such a sample is available. This is a common situation in medical imaging where many subjects are imaged according to a particular protocol in either an experimental or clinical context. When prior information is not available, uniform distributions are used for the prior probabilities of the parameters and an initial estimate of the boundary must be supplied. The images in a sample will differ due to variability in the object shape and the view of the object. The prior probability distributions can then be estimated from the shapes determined from the sample by decomposing the boundaries into their model parameters and collecting statistics. The boundaries of the sample objects are determined either by manual segmentation or, in a training phase, on a set of images with manual initialization and uniform distributions. The normal distribution could be improved most importantly by relaxing the independence restriction and using the covariance matrix. This would capture a stronger model of the inherent variation.

Of course, when the model is very dissimilar to the data, due to large variability, the model does not help very much. Note that an initialization of the view parameters determining the translation and scaling can be determined sepa-

rately, if necessary because of high variability, using a rigid template and correlation. Such a mismatch can also be aided by having multiple models representing typical configurations of parameters, perhaps corresponding to clinical states. In addition, stronger image features, along with priors, such as region information from gray level or texture [42], can also be used.

An independent, multivariate Gaussian can be used for the parameters. An example distribution is shown in Fig. 11. The middle surface corresponds to the mean parameter values. To the left and right of it are the surfaces corresponding to the mean parameter values plus and minus one standard deviation, respectively.

B. Evaluation

The objective function, Equation 15, can be evaluated by numerical integration. The boundary measure, b, can be evaluated at each point on the surface using linear interpolation. The area element on the surface $\mathcal{A}$ is given by

$$dA = \left| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right| du \, dv \,. \tag{16}$$

The gradient of the objective is necessary for optimization. The derivative for the surface objective is

$$\frac{\partial M}{\partial \mathbf{p}_x} = \frac{\partial \ln \Pr(\mathbf{p})}{\partial \mathbf{p}_x} + \frac{k}{2\sigma_n^2} \iint_{\mathcal{A}} \left[|b(x,y,z)| \frac{\partial}{\partial \mathbf{p}_x} \left| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right| \right.$$
$$\left. + \frac{\partial |b(x,y,z)|}{\partial x} \frac{\partial x(\mathbf{p}, u, v)}{\partial \mathbf{p}_x} \left| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right| \right] du \, dv \tag{17}$$

and similarly for y and z. This expression can also be evaluated by numerical integration. Expressions such as $\frac{\partial |b|}{\partial x}$ can be determined by discrete derivative calculations at each point on the surface, again using trilinear interpolation. The expressions such as $\frac{\partial x}{\partial \mathbf{p}_x}$ can be calculated from the expressions for x, y, and z (shown in Equation 3). The partials $\frac{\partial \mathbf{x}}{\partial u}$ and $\frac{\partial \mathbf{x}}{\partial v}$ can be evaluated analytically. For Gaussian prior distributions,

$$\frac{\partial [\ln \Pr(\mathbf{p})]}{\partial \mathbf{p}_x} = -\frac{\mathbf{p}_x - \mathbf{m}_x}{\sigma_x^2} \,, \tag{18}$$

where $\mathbf{m}_x$ is the mean of the distribution for the parameters governing x and σ_x is the standard deviation. The expression $\frac{\partial}{\partial \mathbf{p}_x} \left| \frac{\partial \mathbf{x}}{\partial u} \times \frac{\partial \mathbf{x}}{\partial v} \right|$ can be calculated analytically from Equation 10. The above follows similarly for $\frac{\partial}{\partial \mathbf{p}_y}$ and $\frac{\partial}{\partial \mathbf{p}_z}$.

VI. BOUNDARY PARAMETER OPTIMIZATION

The problem to be solved is that of maximizing the objective function $M(\mathbf{p})$. The objective function we are solving is not in general convex, but depends ultimately on the gray-level surface shape of the image. However, the prior probability term in the objective function is the logarithm of a Gaussian and is thus quadratic. This term will therefore dominate on the tails of the distributions, making

distant points in the space non-optimal. The starting point for the optimization will be taken to be the maximum of the prior distributions. The global optimum probably will be near the starting point and thus a local optimum is likely to be a global optimum. The degree to which this is true depends on the width of the distributions. Since a local optimization method is likely to be sufficient, although there is still the possibility of converging to a poor local maximum, the excessive computation involved in finding a global optimum is deemed not necessary. Poor convergence can be identified by a corresponding low objective function value and verified visually. Smoothing can also be used to avoid getting trapped in a local maximum.

Conjugate gradient optimization [43] was used to optimize the objective function. This method takes steps based on the direction of the gradient (the direction of greatest increase) until an optimal point is found.

The computing time depends primarily on the length of the parameter vector, due to the gradient computation, and the size of the object, due to the surface integration. In addition, the convergence of the optimization also depends on the length of the parameter vector, as this determines the dimension of the optimization space. As an example, for the case in Fig. 15 the execution time was roughly 10 minutes on a Sun SparcStation 10. The main memory requirements are the 3D image and the discretized surface, with the image being the primary storage.

VII. EXPERIMENTS

A measure of the difference or error between two surfaces must be developed in order to test the system. There is no simple way to determine the best correspondence between points on the two surfaces. Instead, the error is defined as the average distance between each point on the estimated surface and the closest point on the true surface. That is, the error between surfaces S and $\hat{S}$ is

$$e(S, \hat{S}) = \frac{\int_{(u,v) \in \hat{S}} \min_{(u',v') \in S} |S(u',v') - \hat{S}(u,v)| \, dA}{\int_{(u,v) \in \hat{S}} dA} \,. \tag{19}$$

This can be computed discretely using a distance transform of a binary volume representing the true surface [44]. By correlating this with the binary volume representing the estimated surface, we get the sum of the minimum distances from the estimated to the true surface. This is then normalized by the size of the estimated surface.

The deformable surface boundary finding method has been applied to a variety of objects from radiologic images. The example shown in Fig. 6 is a simple synthetic image of a closed surface with Gaussian noise added, resulting in a signal-to-noise ratio (SNR) = 2.5. A probability distribution was simulated and the initial surface was positioned roughly at the target object. The final optimized surface (using up to fourth order terms) matches well with the target. The error, calculated by the above method, is 0.65 pixels.
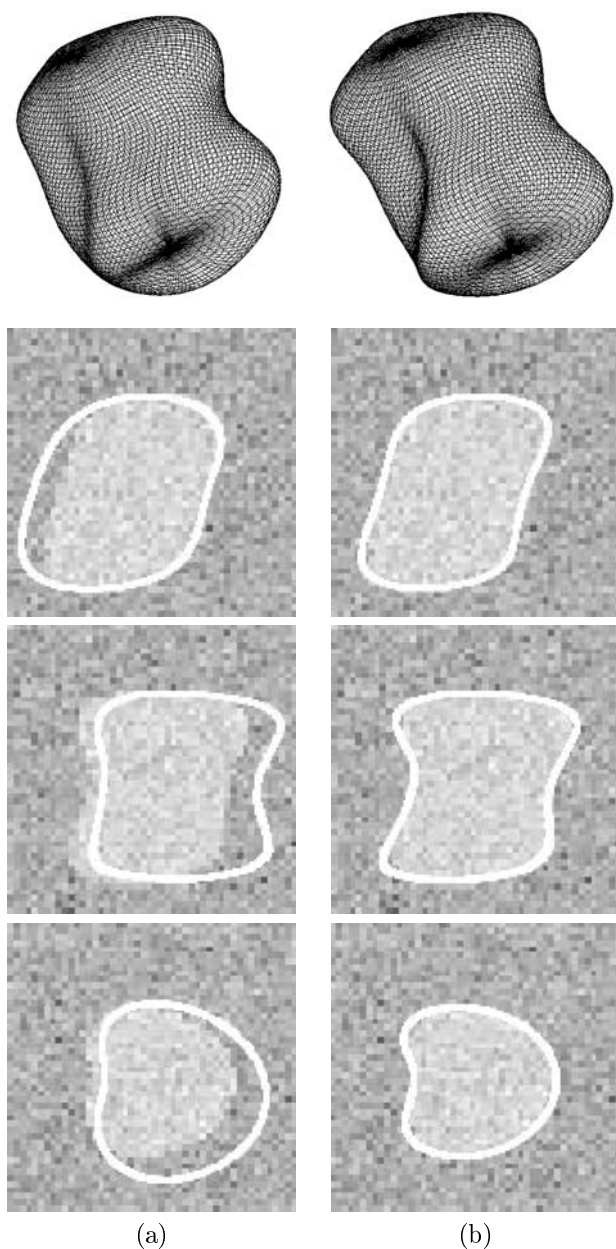
(a)   (b)

Fig. 6. Closed surface synthetic image example. (a) Three perpendicular slices through the 3D image (48 × 48 × 48) are shown with the initial surface using up to four harmonics (69 parameter closed surface) and the wire frame. (b) The same slices shown with the final surface and the wire frame.
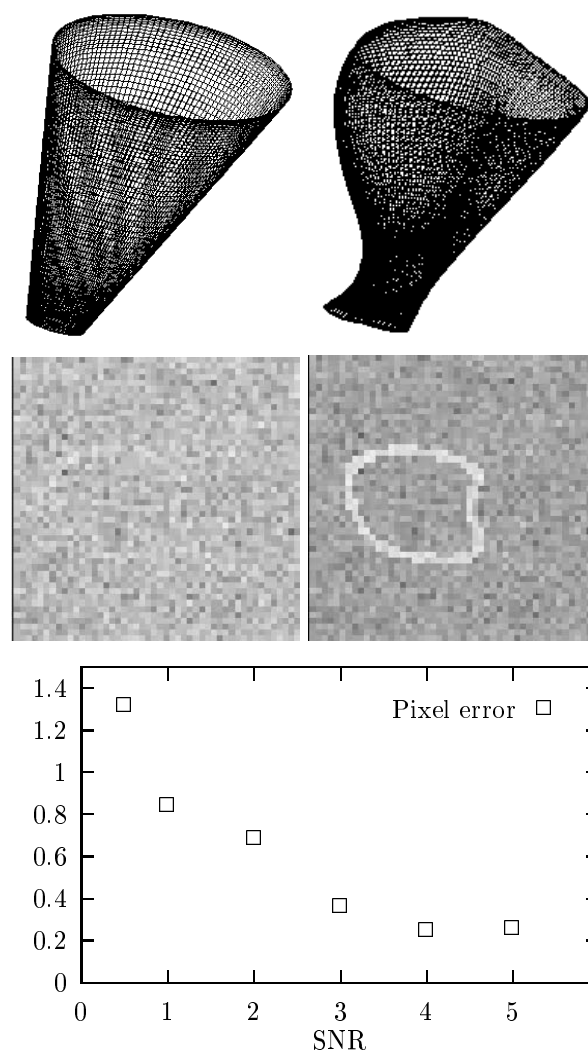

Fig. 7. Noise experiment. Top: Wire frames of initial cone parametrization and exact parametrization using up to four harmonics (84 parameter tube surface). Middle: Slices of synthetic 3D image (48 × 48 × 48) shown with two levels of additive noise: SNR = 1, 5. Bottom: Plot of the average error in pixels as a function of SNR.

Experiments have shown this method to be relatively insensitive to noise. In Fig. 7, an experiment showing the dependence of error on additive noise is shown. A simple synthetic image of a tube surface was degraded by additive noise of varying levels. The resulting error is shown to rise sharply for SNR < 1 but is quite good for higher SNR's.

The effect of the initial values of the parameters on the performance was such that each parameter was found to have a region of success or capture depending on the quality of the image, within which the solution was found reliably. Once the parameters are varied beyond that range, the result will converge to false local minima corresponding to nearby features. False minima can be distinguished,

however, both visually and by the relative value of the objective function. In Fig. 8, the dependence of error on initial position is tested by varying the horizontal shift of the initial parametrization using the image above with SNR = 3. The resulting error is shown to be quite good for shifts less than five pixels.

In Fig. 9, a 3D cardiac image of a dog's heart from the Dynamic Spatial Reconstructor (DSR) is analyzed. The DSR is a dynamic, 3D imaging device based on high-speed x-ray computed tomography capable of imaging the moving heart [45]. As before, the closed surface was matched to the gradient magnitude calculated from the image. The endocardial (inner) wall of the left ventricle is successfully delineated.

In Figs. 10, 11 and 12, one temporal frame of a 3D magnetic resonance image (MRI) sequence of the canine heart cycle is analyzed. A prior was determined from a sample of four images which were delineated manually with a closed
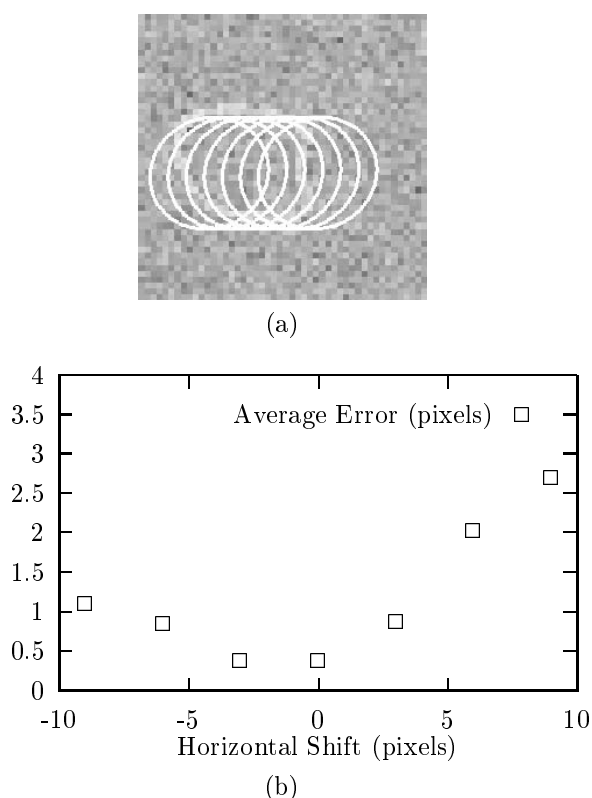
(a)

(b)

Fig. 8. Starting position experiment. (a) Slice from synthetic 3D image with SNR = 3 shown in Figure 7 overlaid with the range of starting positions tested using the same initial surface. (b) Plot of dependence of error on initial shift for this experiment.

surface. The corresponding wire frames are shown in Fig. 10. The prior probability distribution, characterized by the mean and standard deviation, was determined from the parameters of the sample, and is depicted in Fig. 11. The mean of the distribution was used as the initial position for boundary finding of another image. The closed surface was matched to the gradient magnitude calculated from the image. The endocardial border of the left ventricle is successfully delineated. Compared with contours hand-traced by an expert, the average error was 0.64 pixels.

In Figs. 13 and 14, a 3D MRI of the human brain is analyzed. Here, a tube surface was matched to the gradient magnitude calculated from the image. The border of the right caudate nucleus is successfully delineated. The initial position was the mean of a prior probability distribution. The variation about this mean due to individual parameters is shown in Fig. 13. Compared with contours hand-traced by an expert, the average error was 0.15 pixels.

Note that with open or tube surfaces, there can be a tendency to grow at the edges because the additional support will increase the objective function even if it is relatively weak. The objective function can be augmented using normalization by a power, e.g., $1/2$ or 1, of the surface area, $A = \int \int dA$. The power controls the strength of the normalization. The likelihood interpretation of the objective function is lost, however.

Another application of these models is to spatiotempo-

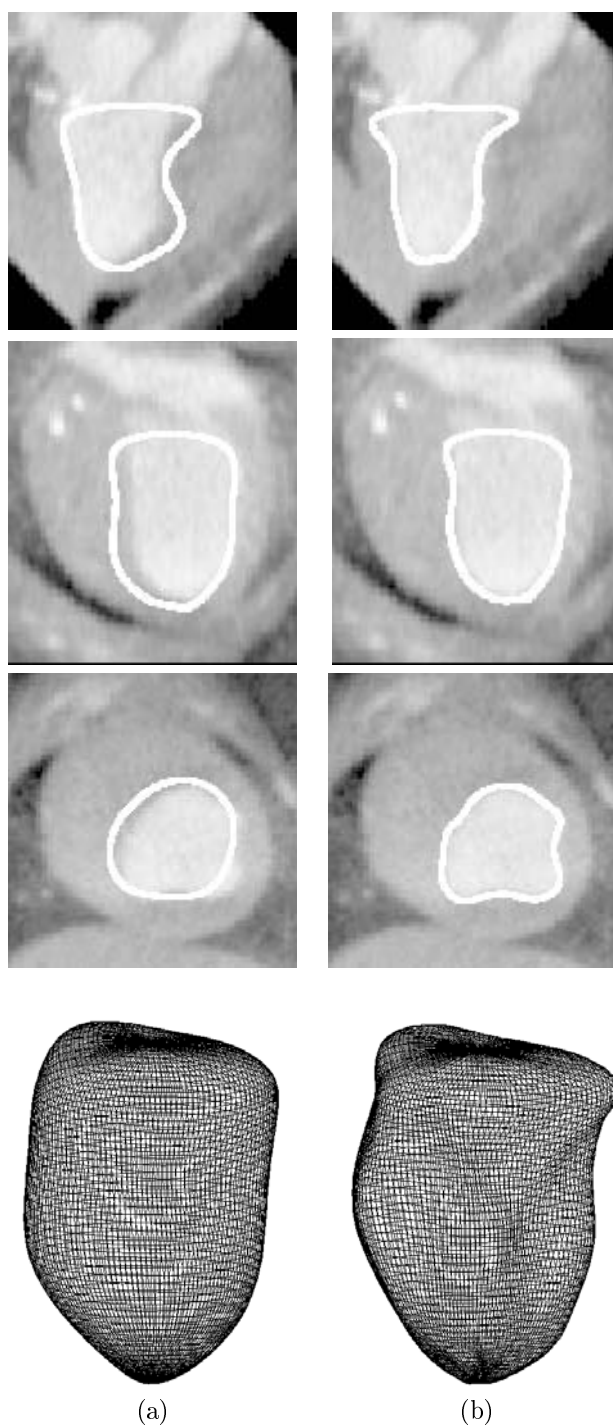(a)   (b)

Fig. 9. Dynamic spatial reconstructor (DSR) cardiac image example. (a) Three perpendicular slices through the 3D image ($98 \times 100 \times 110$) are shown with the initial surface and the wire frame using up to eight harmonics (321 parameter closed surface). (b) The same slices are shown with the final surface at the endocardium of the left ventricle and the corresponding wire frame.

ral images. Spatiotemporal models can be developed and used to measure motion. For two-dimensional objects, the motion can be characterized by the spatiotemporal surface corresponding to the object's moving boundary. These spatiotemporal surfaces would be parametrized using basis functions, as described above. Generalizing further,

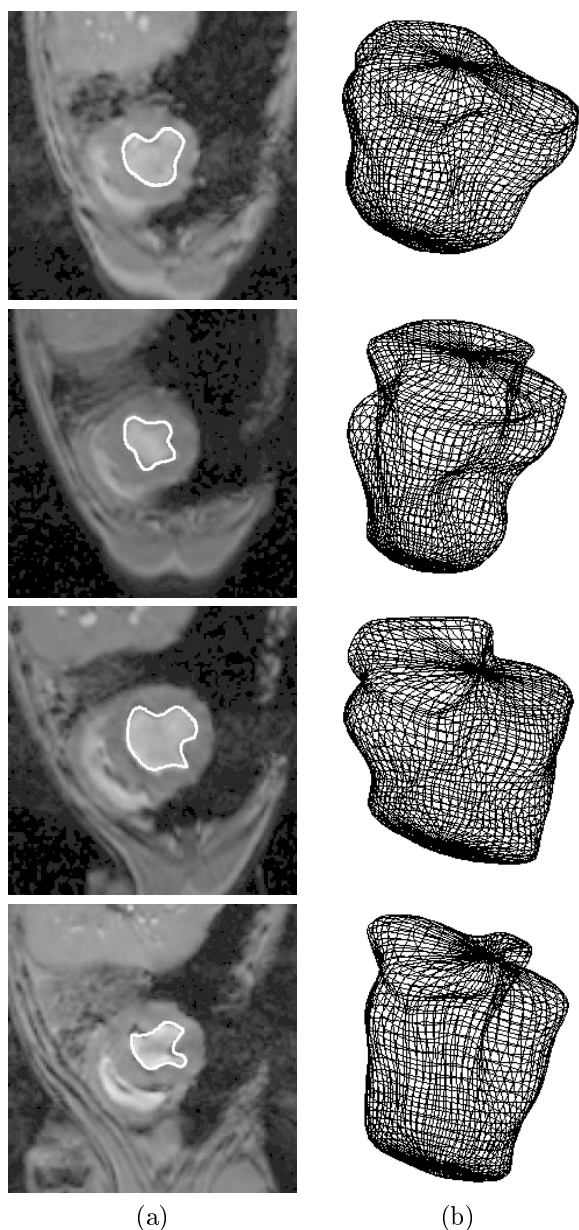(a)   (b)

Fig. 10. Cardiac MRI training set. (a) A slice through each of the training images is shown with the corresponding contour. (b) Wire frames using up to eight harmonics for each of the individuals in the sample used to determine the distribution shown in Fig. 11.
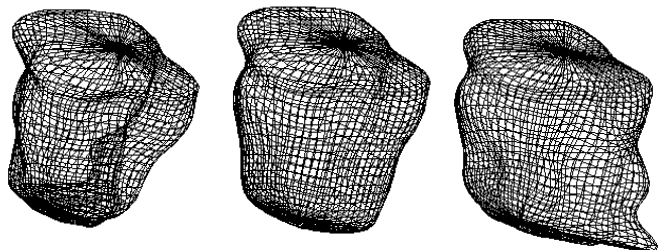

Fig. 11. Cardiac MRI example prior. The mean surface (center) is shown with surfaces corresponding to parameters plus and minus one standard deviation. This distribution is built from the sample shown in 10 and used in Fig. 12.

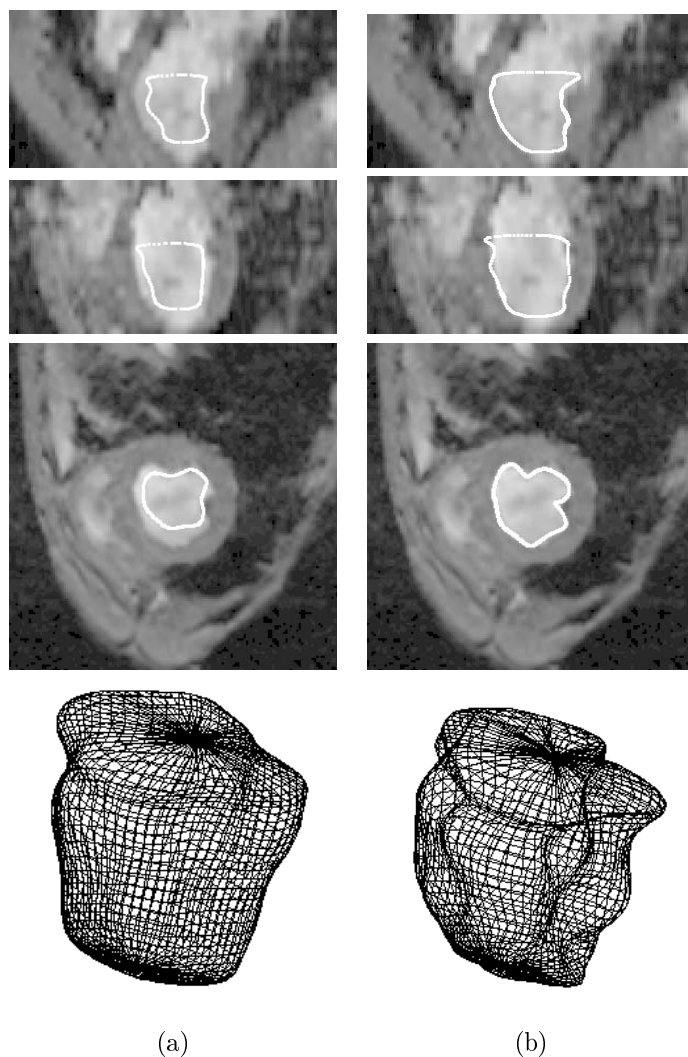(a)   (b)

Fig. 12. Cardiac MRI example. (a) Three perpendicular slices through a 3D image ($150 \times 150 \times 73$) are shown with the initial surface (from the prior shown in Fig. 11) and the corresponding wire frame (321 parameter closed surface). (b) The same slices are shown with the final surface at the endocardium and the corresponding wire frame.

the motion of surfaces could be modeled by a manifold in four dimensions. Note that if the correspondence between points on successive boundaries can also be determined, this represents an approach to non-rigid object motion tracking.

In Fig. 15, a 3D spatiotemporal MRI (two spatial dimensions and one temporal dimension) of a dog's heart at one level through the entire cardiac cycle is analyzed. Here, a tube surface was matched to the gradient magnitude calculated from the image. The endocardial wall of the left ventricle is successfully delineated throughout the cardiac cycle. The initial position was the mean of a prior probability distribution. Compared with contours hand-traced by an expert, the average error was 0.47 pixels.

VIII. Summary

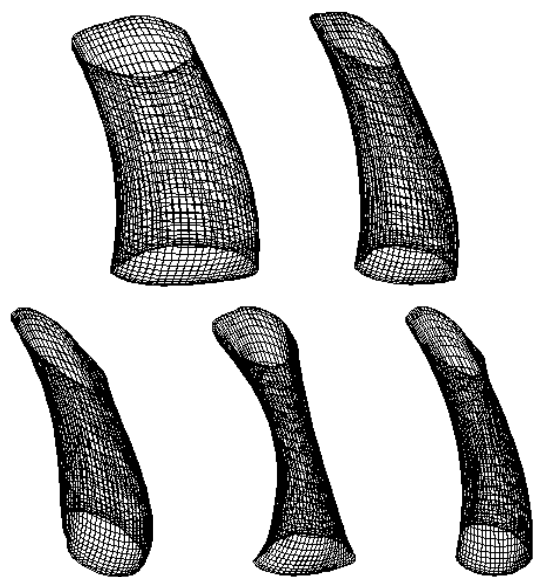This work presents a general boundary finding system for 3D medical images. The goal of this work was to in-

Fig. 13. Brain MRI example prior. Example wire frames showing the variation caused by changing five individual parameters with high variance for the prior shape used in Fig. 14.

corporate both model-based global shape information and prior knowledge of shape into boundary finding for continuously deformable objects. This way of modeling helps constrain the problem and takes advantage of the strong prior information often available in medical imaging. From testing on real and synthetic images, the system was found to perform well at delineating structures and to be relatively insensitive to the problems of broken boundaries, spurious edges from nearby objects, noise and starting position. The flexibility of the model makes this an attractive method for boundary finding. In addition, a new global shape parametrization for surfaces useful as a representation for computer vision and modeling has been described. This parametrization extends the expressibility of previous parametrizations.

REFERENCES

[1]   L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 11, pp. 1061–1075, Nov. 1992.
[2]   D. H. Ballard and C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, 1982.
[3]   M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Computer Vision*, vol. 1, no. 4, pp. 321–331, 1988.
[4]   I. Cohen, L. Cohen, and N. Ayache, "Using deformable surfaces to segment 3-D images and infer differential structures," *Comp. Vision Graphics Image Proc.*, vol. 56, no. 2, pp. 242–263, 1992.
[5]   L. Lifshitz, K. Fogarty, J. Gauch, and E. Moore, "Computer vision and graphics in fluorescence microscopy," in *Visualization Biomed. Comp. 1992, Proc. SPIE 1808*, R. A. Robb, Ed., 1992, pp. 521–534.
[6]   D. Terzopoulos, A. Witkin, and M. Kass, "Symmetry-seeking models and 3D object reconstruction," *Int. J. Computer Vision*, vol. 1, pp. 211–221, 1987.
[7]   R. Malladi, J. Sethian, and B. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 2, pp. 158–175, Feb. 1995.
[8]   H. Delingette, G. Subsol, S. Cotin, and J. Pignon, "A craniofacial surgery simulation testbed," in *Visualization Biomed. Comp. 1994, Proc. SPIE 2359*, R. A. Robb, Ed., 1994.

(a)   (b)

Fig. 14. Brain MRI example. (a) Three perpendicular slices through a 3D image (132 × 165 × 97) are shown with the initial surface and the wire frame (180 parameter tube surface). (b) The same slices are shown with the final surface at the caudate nucleus and the corresponding wire frame.

[9]   R. T. Whitaker, "Algorithms for implicit deformable models," in *Proc. Fifth Int. Conf. Comp. Vision*, June 1995, pp. 822–827.
[10] B. Widrow, "The "rubber mask" technique - I and II," *Pattern Recognition*, vol. 5, pp. 175–211, 1973.
[11] A. L. Yuille, D. S. Cohen, and P. W. Hallinan, "Feature extraction from faces using deformable templates," in *Proc. Comp. Vision Pattern Recog.*, 1989, pp. 104–109.
[12] Y. Chow, U. Grenander, and D. M. Keenan, "Hands: A pattern theoretic study of biological shapes," Tech. Rep., Div. Appl. Math., Brown University, Providence, Rhode Island, 1989.
[13] A. Pentland and S. Sclaroff, "Closed-form solutions for physically based shape modeling and recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 7, pp. 715–729, July
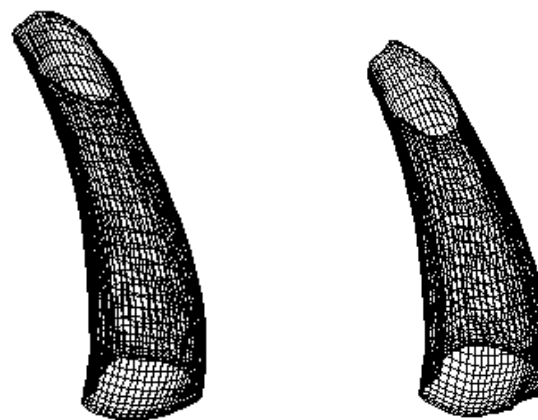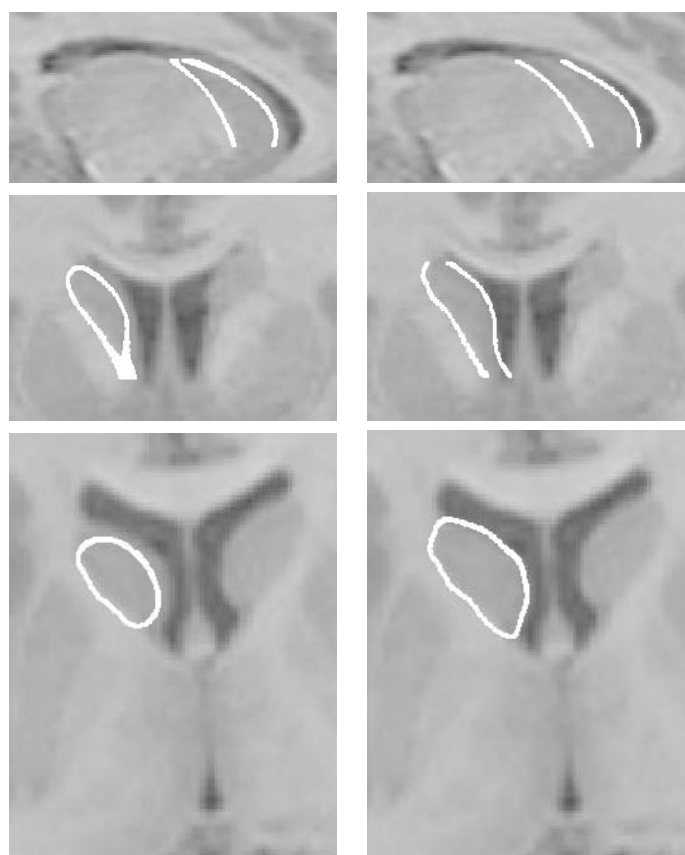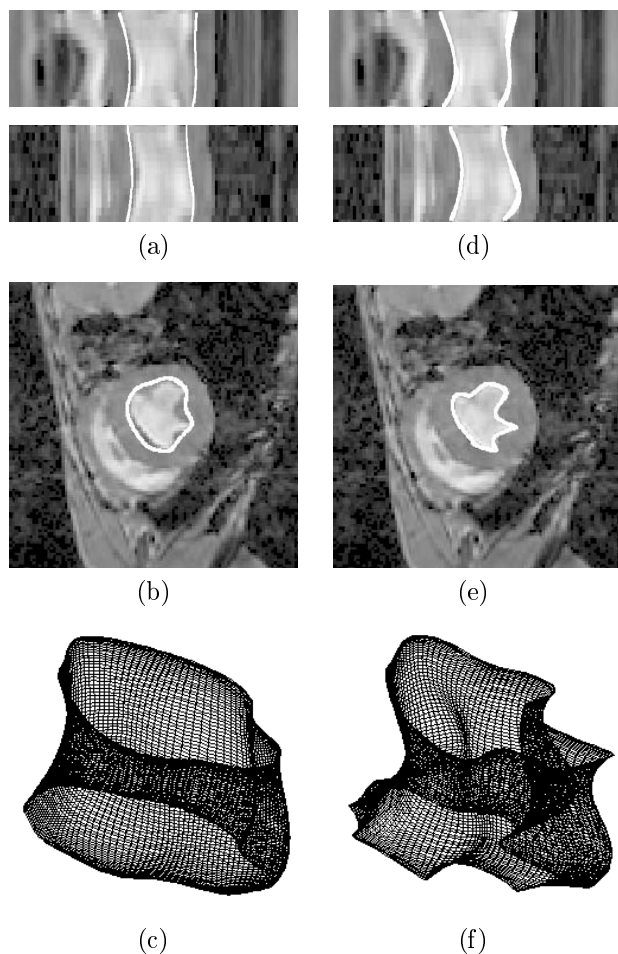
Fig. 15. Cardiac MRI spatiotemporal stack example. (a) Two perpendicular temporal slices through the spatiotemporal image stack ($256 \times 256 \times 16$), and a spatial slice, (b), are shown with the initial surface (180 parameter tube surface) and the wire frame, (c). (d-f) The same slices shown with the final surface at the endocardium and the wire frame.

1991.

[14] C. Nastar and N. Ayache, "Classification of nonrigid motion in 3D images using physics-based vibration analysis," in *Proc. IEEE Workshop Biomedical Image Anal.*, 1994, pp. 61–69.

[15] T. Cootes and C. Taylor, "Combining point distribution models with shape models based on finite element analysis," *Image and Vision Computing*, vol. 13, no. 5, pp. 403–410, June 1995.

[16] M. Bomans, K. Hohne, U. Tiede, and M. Riemer, "3-D segmentation of MR images of the head for 3-D display," *IEEE Trans. Medical Imaging*, vol. 9, no. 2, pp. 177–183, June 1990.

[17] W. E. Higgins, N. Chung, and E. L. Ritman, "Extraction of left-ventricular chamber from 3-D CT images of the heart," *IEEE Trans. Medical Imaging*, vol. 9, no. 4, pp. 384–395, Dec. 1990.

[18] H. E. Cline, W. E. Lorensen, R. Kikinis, and F. Jolesz, "Three-dimensional segmentation of MR images of the head using probability and connectivity," *J. Comp. Assisted Tomogr.*, vol. 14, no. 6, pp. 1037–1045, Nov./Dec. 1990.

[19] D. N. Kennedy, P. A. Filipek, and V. S. Caviness, "Anatomic segmentation and volumetric calculations in nuclear magnetic resonance imaging," *IEEE Trans. Medical Imaging*, vol. 8, no. 1, pp. 1–7, Mar. 1989.

[20] K. H. Hohne and W. A. Hanson, "Interactive 3D segmentation of MRI and CT volumes using morphological operations," *J. Comp. Assisted Tomogr.*, vol. 16, no. 2, pp. 285–294, 1992.

[21] D. Collins, T. Peters, W. Dai, and A. Evans, "Model based segmentation of individual brain structures from MRI data," in *Visualization Biomed. Comp. 1992, Proc. SPIE 1808*, R. A. Robb, Ed., 1992, pp. 10–23.

[22] T. Cootes, A. Hill, C. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," in *Information Proc. Med. Imaging*, H. H. Barrett and A. F. Gmitro, Eds., pp. 33–47. LNCS 687, Springer-Verlag, Berlin, 1993.

[23] G. Szekély, A. Kelemen, C. Brechbüler, and G. Gerig, "Segmentation of 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier surface models," in *Comp. Vision, Virtual Reality and Robotics in Med. (CVRMed '95, LNCS 905)*, N. Ayache, Ed., Berlin, 1995, pp. 495–505, Springer-Verlag.

[24] K. Rao and R. Nevatia, "Computing volume descriptions from sparse 3-D data," *Int. J. Computer Vision*, vol. 2, no. 1, pp. 33–50, 1988.

[25] J. Blokland, A. Vossepoel, A. Bakker, and E. Pauwels, "Delineating elliptical objects with an application to cardiac scintigrams," *IEEE Trans. Medical Imaging*, vol. 6, no. 1, pp. 57–66, Mar. 1987.

[26] J. Nuyts, P. Suetens, A. Oosterlinck, M. De Roo, and L. Mortelmans, "Delineation of ECT images using global constraints and dynamic programming," *IEEE Trans. Medical Imaging*, vol. 10, no. 4, pp. 489–498, Dec. 1991.

[27] D. Keren, J. Subrahmonia, and D. Cooper, "Robust object recognition based on implicit algebraic curves and surfaces," in *Proc. Comp. Vision Pattern Recog.*, 1992, pp. 791–794.

[28] F. Solina and R. Bajcsy, "Recovery of parametric models from range images: The case for superquadrics with global deformations," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, no. 2, pp. 131–147, Feb. 1990.

[29] A. H. Barr, "Global and local deformations of solid primitives," *Computer Graphics*, vol. 18, no. 3, pp. 21–30, July 1984.

[30] D. Terzopoulos and D. Metaxas, "Dynamic 3D models with local and global deformations: Deformable superquadrics," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 7, pp. 703–714, July 1991.

[31] E. Bardinet, L. D. Cohen, and N. Ayache, "Fitting of iso-surfaces using superquadrics and free-form deformations," in *Proc. IEEE Workshop Biomedical Image Anal.*, 1994, pp. 184–193.

[32] A. P. Pentland, "Automatic extraction of deformable part models," *Int. J. Computer Vision*, vol. 4, pp. 107–126, 1990.

[33] B. Vemuri, A. Radisavljevic, and C. Leonard, "Multi-resolution stochastic 3D shape models for image segmentation," in *Information Proc. Med. Imaging*, H. H. Barrett and A. F. Gmitro, Eds., pp. 62–76. LNCS 687, Springer-Verlag, Berlin, 1993.

[34] R. B. Schudy, "Harmonic surfaces and parametric image operators: Their use in locating the moving endocardial surface from three-dimensional cardiac ultrasound data," Computer Science Tech. Rpt. 112, University of Rochester, Rochester, New York, Mar. 1981.

[35] C. Brechbühler, G. Gerig, and O. Kübler, "Surface parametrization and shape description," in *Visualization Biomed. Comp. 1992, Proc. SPIE 1808*, R. A. Robb, Ed., 1992, pp. 80–89.

[36] E. Persoon and K. S. Fu, "Shape discrimination using Fourier descriptors," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 3, pp. 388–397, May 1986.

[37] G. P. Tolstov, *Fourier Series*, Prentice-Hall, Englewood Cliffs, 1962.

[38] P. J. Besl, "Geometric modeling and computer vision," *Proc. IEEE*, vol. 76, no. 8, pp. 936–958, 1988.

[39] P. Shi, G. Robinson, and J. Duncan, "Myocardial motion and function assessment using 4D images," in *Visualization Biomed. Comp. 1994, Proc. SPIE 2359*, R. Robb, Ed., 1994, pp. 148–159.

[40] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.

[41] D. B. Cooper, "Maximum likelihood estimation of Markov-process blob boundaries in noisy images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 1, no. 4, pp. 372–384, Oct. 1979.

[42] A. Chakraborty, L. H. Staib, and J. S. Duncan, "Deformable boundary finding influenced by region homogeneity," in *Proc. Comp. Vision Pattern Recog.*, 1994.

[43] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling, *Numerical Recipes*, Cambridge U. Press, Cambridge, 1986.

[44] R. van den Boomgaard, *Mathematical Morphology: Extensions Towards Computer Vision*, Ph.D. thesis, U. Amsterdam, 1992.

[45] R. A. Robb, "High speed three-dimensional x-ray computed tomography: The dynamic spatial reconstructor," *Proc. IEEE*, vol. 71, pp. 308–319, 1983.