Segmentation, Registration, and Measurement of Shape Variation via Image Object Shape

Stephen M. Pizer, Daniel S. Fritsch, Paul A. Yushkevich,
Valen E. Johnson*, Edward L. Chaney
Medical Image Display & Analysis Group
University of North Carolina

*Inst. for Statistics & Decision Science, Duke University

Abstract

A model of object shape by nets of medial and boundary primitives is justified as richly capturing multiple aspects of shape and yet requiring representation space and image analysis work proportional to the number of primitives. Metrics are described that compute an object representation's prior probability of local geometry by reflecting variabilities in the net's node and link parameter values and that compute a likelihood function measuring the degree of match of an image to that object representation. A paradigm for image analysis of deforming such a model to optimize a posterior probability is described, and this paradigm is shown to be usable as a uniform approach for object definition, object-based registration between images of the same or different imaging modalities, and measurement of shape variation of an abnormal anatomical object compared with a normal. Examples of applications of these methods in radiotherapy, surgery, and psychiatry are given.

Introduction

Our intuition tells us that to segment, display, match, or analyze an object in a 2D or 3D medical image, we should look for a region of the image that is consistent with what we know about the object's geometry. We should be able, for example, to find a ventricle-shaped object in an MR image of the brain. This paper gives a means of representing shape that allows the efficient, stable application of this idea to problems of image analysis.

We describe a representation that
1) richly and efficiently captures object geometry, and
2) allows measurement of the shape difference between two instances of an object, in the framework of realistic images, and thereby
3) provides a common framework for the medical image operations of segmentation, object-based registration, measurement of shape variation, and indication of regions of pathological shape change.

The general paradigm for image analysis that we will present here involves two components. The first component is to create a model R comprised of interconnected figures, where a *figure* is either a whole object or the main object part or a protrusion or indentation in another figure. It is this focus on the figure as the basic unit of shape that distinguishes the approach that we will describe in this paper. The second component of the approach is to determine a geometric transformation T that maps R into the most probable variant, given the image data. The measure of probability involves two components:

1) the probability P_{geom}, of observed change in the modeled shape, position, orientation, and size of the object;
2) the probability P_{img}, that the image data would arise from the object geometry. In this paper we decompose P_{geom} into the probability of the position, orientation, and size of the object and the probability P_{shape} of the shape of the figures and in the interfigural spatial variables, and we focus on P_{shape}. Then

1) for a segmentation objective, R models the object over a family of training images, and T(R) yields the segmented object;

2) for a registration objective, R is the object representation in the reference image, T(R) is the object representation in the target image, and T is the desired registration transformation;

3) for a shape variation measurement objective, R is the object representation at its base position, and -log P_{shape} is the measure of change in shape from R to T(R).

Implementing this conceptual framework requires us to specify the representation R of object geometry and to specify the formation of the probabilities P_{geom} and P_{img}. The remainder of this paper begins with a section on representation of object geometry, leading to a multifigural tree of linked, coarsely sampled medial loci augmented by a boundary displacement map on the medially implied boundary. It is followed by a section on measures of the probability of geometric change and a section on measures of the fit of image data to an object representation. Sections follow indicating how these ideas can be used for image analysis, in particular, segmentation, registration, and measurement of shape anomaly or differences. However, since we wish to focus here on the shape representation, we leave details of the algorithms for these objectives to other papers on the individual algorithms [Wilson 1995, Wilson 1996, Fritsch 1995, Fritsch 1997, Liu 1998]. In the discussion of image analysis we give examples from radiotherapy planning, surgery planning, and psychiatric diagnosis, but we leave fuller illustration of these applications to other papers on the individual clinical topics [Fritsch 1996, Clary 1997].

1. Alternatives for Representing Object Shape

We take the common view that the shape of an object includes all the spatial aspects of its layout that are invariant to similarity transforms, made from translation, rotation, and uniform magnification. Thus an object in d-space is fully spatially described by its shape together with its similarity transform parameters: its position, orientation, and scale. We will call the shape together with these similarity transform parameters of an object its *morphography*. But how can we represent the shape in a way that richly captures the aspects that are visually important, allows efficient calculation of shape variation and location of regions of shape variation, and allows stable application to image data, with all their disturbances of intensity noise and blurring, boundary texture, and especially complexity of background?

Traditionally shape has been derived from points with recognizable local geometry called landmarks [Kendall 1989, Bookstein 1991] or from a representation of the boundary by sampled points [Cootes 1993] or by boundary-representing basis function coefficients [Staib 1992, 1996; Székely 1996, Kelemen 1998]. Hybrids of boundary and other curve loci with landmarks, e.g., [Yuille 1992], and representations by sampling the full space [Christensen 1997, Grenander 1994] have also been suggested. The methods with point primitives have assumed that the boundary representations or landmarks could be accurately, homologously, reproducibly, and conveniently extracted in sufficient number, but they have suffered from the fact that all of these assumptions are frequently questionable in real images.

Efficiency is also a concern. Given N shape primitives, if N is very large, as in the number of pixels in a 2D image or the number of voxels in a 3D image, the time for instantiating a shape representation based on an image can be unacceptably large. If N is small, efficiency may not be a concern but the shape representation is frequently not rich enough. If N is in the moderate range of some hundreds, the efficiency depends on whether the required computation requires $O(N)$ or $O(N^2)$ steps. The similarity transform parameters are easily extracted in $O(N)$ time. However, various methods have extracted shape from the relations between the

primitives, of which there are $O\binom{N}{2} = O(N^2)$, or they require multiplication of an O(N)-vector by a square matrix of side O(N), also requiring $O(N^2)$ steps. The minimal bending energy method of Bookstein [1991] falls into this category.

Among O(N) methods, the Procrustes method has modeled shape by variations in the vectors from the point primitives to their center of mass. Interrelationships among points are not modeled, so the measure of shape produced is somewhat weak. The same can be said about methods based only on the local relations between adjacent boundary points, thus capturing only the local curvatures of the boundary.

Another way of producing an O(N) method has been to take only a constant number of linear combinations of the N primitives, as provided by Principal Component Analysis of the primitive point positions or boundary coefficients in training images after normalization for a similarity transform [Cootes 1993; Kelemen 1998]. These approaches yield shape parameters that reflect global information and thus are weak in their ability to capture aspects of shape that are specific to a particular location, protrusion, or indentation.

We suggest that a method for providing O(N)-time shape analysis and the ability to capture both local and global shape variations in a way stable against common image uncertainties and complexities lies in
   a) recognizing that certain boundary points are related in special ways determined by figural properties,
   b) realizing that a shape is made up of related figures,
   c) focusing on the facts that shape is magnification invariant and that the *tolerance* of positions in object representation primitives is a critical factor (see Fig. 5, later), and
   d) making coarse-to-fine analysis an inherent aspect of the shape description.
This leads us to a representation that is a hierarchy of coarse-to-fine collections of nets of medial primitives, augmented by nets of boundary displacement primitives, all with associated probability distributions. We will describe these ideas with examples in two spatial dimensions. We have shown these ideas to work without change in 3D as well, but we leave the 3D aspects to another paper.

2. Figural Shape Representation

2.1. Medial Primitives and Nets of Medial Primitives

Having assumed that positional tolerance is a central aspect of shape [Morse 1996], we focus first on the large-tolerance aspects of shape that allow us to say that the two objects in Fig. 1 are the same figures but differ in the texture, i.e., the small-tolerance detail, of the boundary. We suggest that these large tolerance aspects are those involving cross-figural linking, which we represent via medial primitives. Blum [1967] and Marr, among others, realized early that cross-figural linking was an important aspect of shape, and Psotka [1978], Frome [1972], and Burbeck & Pizer [1995, 1996] elicited psychophysical evidence of human vision's sensitivity to such linking.

While we wish to reverse the common view of medial primitives that they are derived from a boundary to a view in which they *imply* a boundary, with tolerance, for the next few sentences assume that the boundary is known. We focus first on the special relations between boundary points that a simple object, protrusion, or indentation has. The two special relations between boundary points are the adjacency relation traditionally dealt with and the symmetric relation of linking a boundary point across a figure with a figurally opposing point, called its

medial involute (for example in Fig. 2a, each point $\mathbf{A}_i$ and an opposing point $\mathbf{B}_i$ are each other's medial involutes). The involution relationship has traditionally been defined by the fact that the normals to the boundary at the involutes intersect at a so-called *medial point* equidistant from the two involutes. However, we reverse the process, defining a medial primitive as a medial point with two associated boundary-pointing vectors of equal length (Fig. 2b) and taking the medial involution relation between the points at the end of the two boundary-pointing vectors to be a consequence of our primitive (Fig. 2c) rather than a relation defining it. The sections of the boundary so defined are orthogonal to the respective boundary-pointing vectors and are taken to have positional tolerance in the directions of the respective vectors.
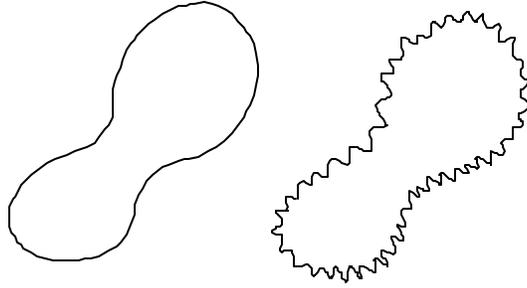

Fig. 1. Two instances of the same figure, one with smooth boundary texture and one with rough boundary texture.
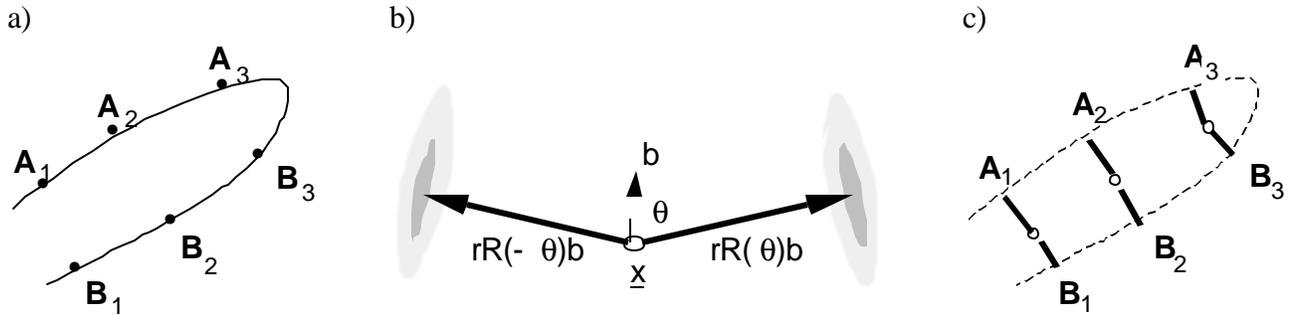

Fig. 2. a) A figure with pairs of medial involutes (**A**,**B**). b) A medial primitive and its implied involutes, with tolerance. c) The three (**A**,**B**) medial involute pairs defined by three medial primitives. The dashed boundary should be thought of as being interpolated through the implied boundary points.

Traditionally, medial information is given by the values of 1) the location x of the intersection of normals to the boundary points and 2) the common distance r from this location to the two corresponding boundary involutes [Blum 1978]. When the boundary points and boundary normals are to be implied by the medial primitive rather than vice-versa, it is necessary to record as well the unit vectors $\vec{n}_1$ and $\vec{n}_2$ pointing from x toward the respective boundary points. For reasons of geometric informativeness to be discussed later, we choose to define a medial primitive by the medial position x, the width r, a unit vector $\vec{b}$ bisecting (averaging) the boundary-pointing vectors, and an angle θ between $\vec{b}$ and the two boundary-pointing vectors (half the angle between the boundary-pointing vectors). θ is called the *object angle*. The implied medial involutes are at $x + rR(q)\vec{b}$ and $x + rR(-q)\vec{b}$, where $R(q)$ is the operator rotating its operand by θ. The boundary-pointing vectors are each taken to imply a boundary fragment normal to its vector, so these fragments are at angle $\pi\text{-}2\theta$ to each other.

We are now ready to more formally define "figure". As illustrated in Fig. 3, a figure is a geometric object implied by a sequence of medial primitives such that each of the two implied boundary point sequences $\{x + rR(q)\vec{b}\}$ and $\{x + rR(-q)\vec{b}\}$ is ordered along the implied boundary in the same order as the primitives. The sequence may be closed by an end primitive that has arrows to $x + rR(q)\vec{b}$, $x + rR(-q)\vec{b}$, and $x + ar\vec{b}$ [Clary 1997]. The sequence of boundary points on one side of a figure can straddle an attached subfigure, i.e., a protrusion or indentation on one side of the figure. Fig. 3 illustrates four figures: a large "sweet potato", a protrusion on the sweet potato, an indentation in the sweet potato, and a nearby "bean".

Medial primitives are associated using two types of links: intrafigural and interfigural. An interfigural link connects a primitive in a subfigure to a primitive in a parent figure. Intrafigural links connect medial primitives along a figure. The spacing of approximately r indicated in Fig. 3 is intended to be typical. In 3D the medial primitives form a net, and we will use "net" to refer to the sequence of primitives in 2D as well.

A net of primitives ends either with a medial primitive at the open end of a protrusion or indentation or with an end primitive. Following Leyton [1992], we distinguish end primitives as a special type of medial primitive. According to Leyton, in analogy to the importance of endstopping of lines in human vision, endstopping of figures corresponding to the distal end of a medial locus is also an important aspect of shape and happens at a maximum curvature point on the boundary where a circle centered at the medial locus end osculates the boundary. Such a point of maximum curvature of the appropriate sense is called a *vertex*.
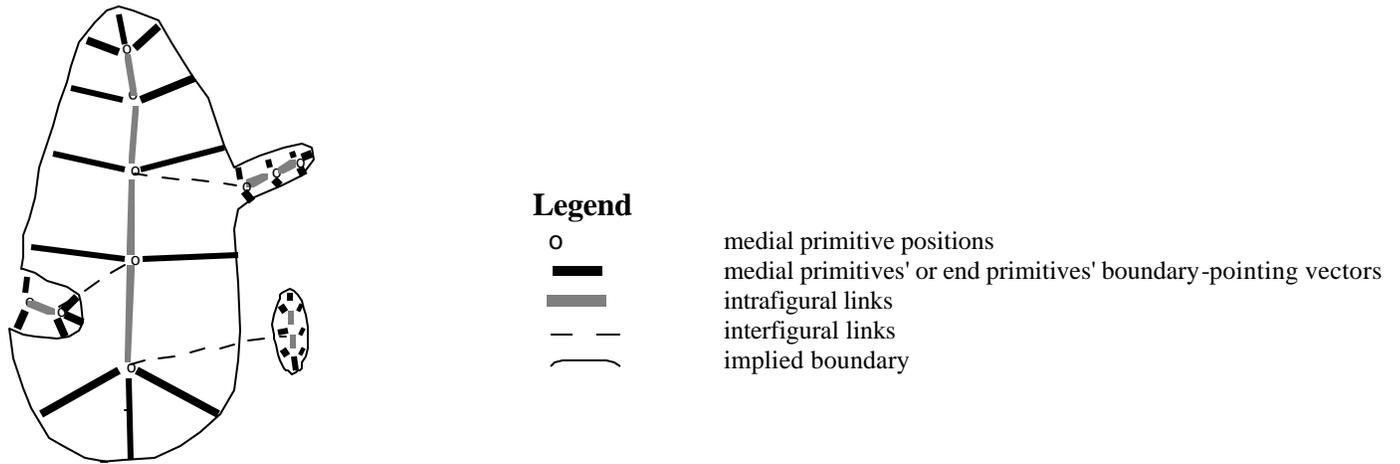

Legend

o	medial primitive positions
▬	medial primitives' or end primitives' boundary-pointing vectors
▬	intrafigural links
— —	interfigural links
⌣	implied boundary

Fig. 3. Linked sequences of medial primitives implying four linked figures.

2.2. Figural Shape

Figural shape is given by the relations among the medial primitives. We first distinguish shape aspects within a figure from those between figures. In our example in Fig. 3, the fact that the main sweet potato figure narrows from bottom to top and curves to the left as it nears the top will fall into the category of intrafigural shape as will the facts that the protrusion on the right is approximately straight and of roughly constant width. The facts that the protrusion is attached somewhat above the middle of the sweet potato, is oriented at around -70 degrees relative to the sweet potato, and has a width about a quarter of the width of the sweet potato will fall into the category of interfigural shape.

While intrafigural shape is fully described by the relations among all the medial primitives forming the figure, our visual capabilities for comprehending shape seem limited to nearby primitives. For this reason and the desire to produce an O(N) description, we will mostly limit the description of shape discussed in this paper to those given by adjacent medial primitives in the net. We will see later, when focusing on coarse-to-fine aspects of the description, that this need not limit us from capturing certain more global figural shape aspects.

In describing interfigural shape, we follow many others in computer vision who have found that an important shape relation is the graph relations between figures and their attached subfigures or neighboring or included figures. We consider one of the figures in this relation the parent and one the child or subfigure, with the result that the whole structure forms a directed acyclic graph, which in many cases is a tree. The structure of this graph will be taken to indicate topology, which in this paper will be treated as deterministic. The consequence is that we will be able to represent shapes such as livers, kidneys, cerebral ventricles, and hippocampi but not be able fully to represent shapes such as blood vessel trees and the cerebrum with its sulci and gyri, since these latter examples vary among individuals in how many branches there are from a particular figure and in what branches from what.

For now we take each relation between a figure and its subfigure to be represented by a single link between a medial point on the parent figure and a proximal medial endpoint of the subfigure. Including these links (dashed in Fig. 3) still leaves us with O(N) links for N medial primitives.

The links, both intrafigural and interfigural, between medial primitives (see Fig. 4) carry the figural shape information. Specifically, the shape information is in the "shape tuple" $\left|\Delta \boldsymbol{x}/r\right|, \Delta r/r, \Delta \vec{b}, \boldsymbol{q}, \Delta \boldsymbol{q}, \hat{\Delta} \boldsymbol{x} - \vec{b}$, where "^" denotes normalization to a unit vector, Δ indicates a difference between the linked primitives, differences between unit vectors denote an angle, and r and θ are measured at the tail of the link. All of these measures are dimensionless. If the link is between adjacent primitives within a figure, $\left|\Delta \boldsymbol{x}/r\right|$ carries information about the local elongation of the figure, $\Delta r/r$ and θ carry information about the local widening or narrowing rate of the figure, $\Delta \vec{b}$ carries information about the local curvature of the figural (medial) axis, and $\Delta \boldsymbol{q}$ carries information about the local curvature of the boundary relative to the axis [Blum & Nagel 1978]. The direction of $\Delta \vec{x}$ relative to $\vec{b}$ of the reference primitive carries information of spatial derivative order higher than two.

If the link is between a primitive in a parent figure and a primitive in a subfigure, $\Delta r/r$ carries information about the proximal width of the subfigure relative to the local width of its parent figure, $\Delta \vec{b}$ carries information about the proximal orientation of the subfigure relative to the local width of its parent figure, and the direction of $\Delta \vec{x}$ relative to $\vec{b}$ of the parent primitive carries information on the location on the parent figure of the subfigure (e.g., in Fig. 3 whether the protrusion is high on the sweet potato, towards the middle, or low).

The local shape related to a particular medial primitive can be taken to be given by the set of shape parameters of all of the links incident on that primitive. This set of dimensionless quantities and vectors is invariant to the similarity transform operations of uniform magnification, translation, and rotation. But
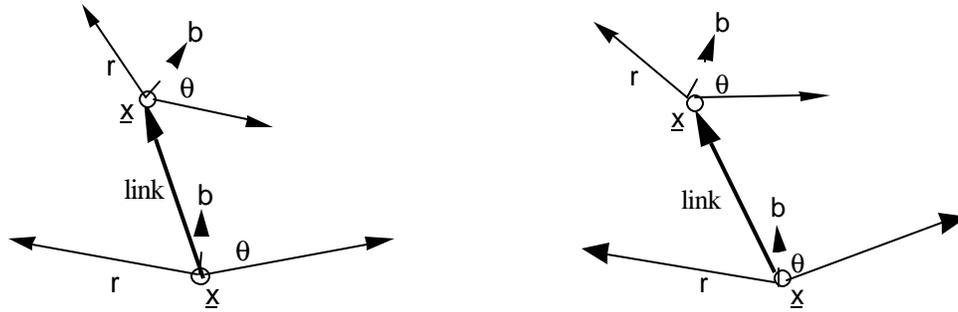
Fig. 4. Two linked medial primitives, before and after figural deformation. Take the lower medial primitive to be the reference primitive, and let the upper one be either an adjacent primitive along the same figure (connected by a grey line in Fig. 3) or a linked subfigural primitive (connected by a dashed line in Fig. 3).

note with emphasis that we have taken the tolerance of the boundary fragments implied by a medial primitive to be a significant aspect of the primitive. The spacing of primitives (the length of the links between primitives) is also an important aspect. In addition, we may follow the common wisdom that boundary radius of curvature is an important aspect, even when it is of the (fuzzy) boundary implied by the medial primitives. For medial width (r), tolerance, link length, and boundary radius of curvature all to be variables associated with the medial primitives used to represent a shape locally, these distance-dimensioned variables must scale together. That is, if we need to describe shape with locality, then as illustrated in Fig. 5a, tolerance, the distribution of link length, and the distribution of 1/boundary curvature (i.e., radius of curvature) must scale locally with the local value of r. The wider the figure is, the larger the tolerance of the figurally implied boundary must be, the smoother the implied boundary must be, and the longer the intrafigural links must be.
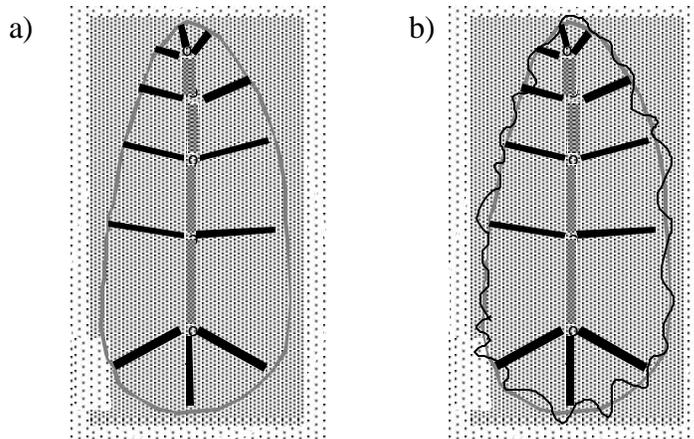

Fig. 5. a) The net of medial primitives shown represents a figure and implies a boundary whose mean is shown by the curve and whose tolerance, as a Gaussian probability distribution in the direction of the normal to the mean, is shown by the two levels of shading. b) Boundary displacements. The boundary of small tolerance, indicated by the narrower curve, is represented by displacements, at a fine sampling, from the mean curve implied by the medial primitives in a direction normal to that mean curve.

Interpolation between adjacent medial primitives by interpolation in $\vec{b}$, q, and *log r* is consistent with these requirements. The resulting continuous locus of medial primitives implies a continuous boundary which is the boundary implied by the net of medial primitives.

In the case of the root figure in the tree of figures, we can think of it having a link to the full image space. That is, the spatial variables describing its link to the image are the global similarity transform variables of the whole object. These are not shape variables but are also important in the object representation.

2.3. Boundary Displacement Shape

Figural shape only implies a boundary with a large tolerance and with restrictions on the range of curvatures it can have. This large tolerance allows a coarse spacing of medial primitives and thus an efficient representation.

However, an object boundary certainly also has shape at a smaller scale. This aspect of shape information can be determined more stably if, instead of following the common practice where the primitives are boundary primitives containing a position and normal that are located only in relation to the image space, the final boundary location is located relative to the mean boundary implied by the medial primitives and is constrained by the tolerance of the boundary implied by the medial primitives. That is, we define the boundary at small scale by displacements relative to and along the normal to the mean boundary implied by the net of medial primitives (Fig. 5b). Thus these displacements can be defined in one dimension, along the normal direction to the medially implied mean boundary. Relations among nearby displacements can reflect local textural properties of the boundary.

Forming a net of boundary displacement primitives requires piecing together the boundary primitives between subfigures and their parent figures. Chen [1998] has shown how to do this in his work on object rendering via the representations discussed in this paper.

On what basis is a particular topology of medial primitives and boundary displacement primitives defined to be the right set to represent the figural shape information in a particular population of objects relative to a particular set of images? We take this to depend on the task at hand and not to be somehow innate to the images. For example, for a purpose in which finger shape is to be studied, a hand might be modeled as a palm figure, with 5 protrusion subfigures representing four fingers and a thumb, and the extensions of the fingernails represented by boundary displacements. For another purpose, the same hand might be modeled as a mitten, with 4 subfigures being the between-finger and finger-thumb indentations. For yet another purpose the fingernail extensions might be modeled as subfigures of the fingers instead of by boundary displacements. For the time being, we take these modeling decisions to be most appropriately done by human intelligence, whereas the values of the primitives ought to be extracted algorithmically according to the values of training images. We leave the effect of these modeling decisions to future work.

2.4. Coarse-to-Fine Representation

The principle that medial primitives represent the object information at one scale, proportional to the inter-medial-primitive distances, and that boundary primitives represent the object information at a smaller scale, proportional to the inter-boundary-primitive distances, and that boundary primitive location is dependent on boundary locations implied by the medial primitives is an instance of general principle of coarse-to-fine representations. For our purposes we will deal with four levels of coarseness (see Fig. 6): a) for the whole object or group of objects, b) for each figure, c) for local intrafigural relations, and d) for boundary displacements. These have interprimitive connections at successively smaller maximal distances (i.e., scale). Such representations can lead simultaneously to efficiency, to handling both long- and short-distance shape

relations, and to stability of image analysis against image disturbances. We leave to a future paper the major discussion of this principle and its implementation and especially the discussion of successively diced nets of medial primitives and of boundary primitives.
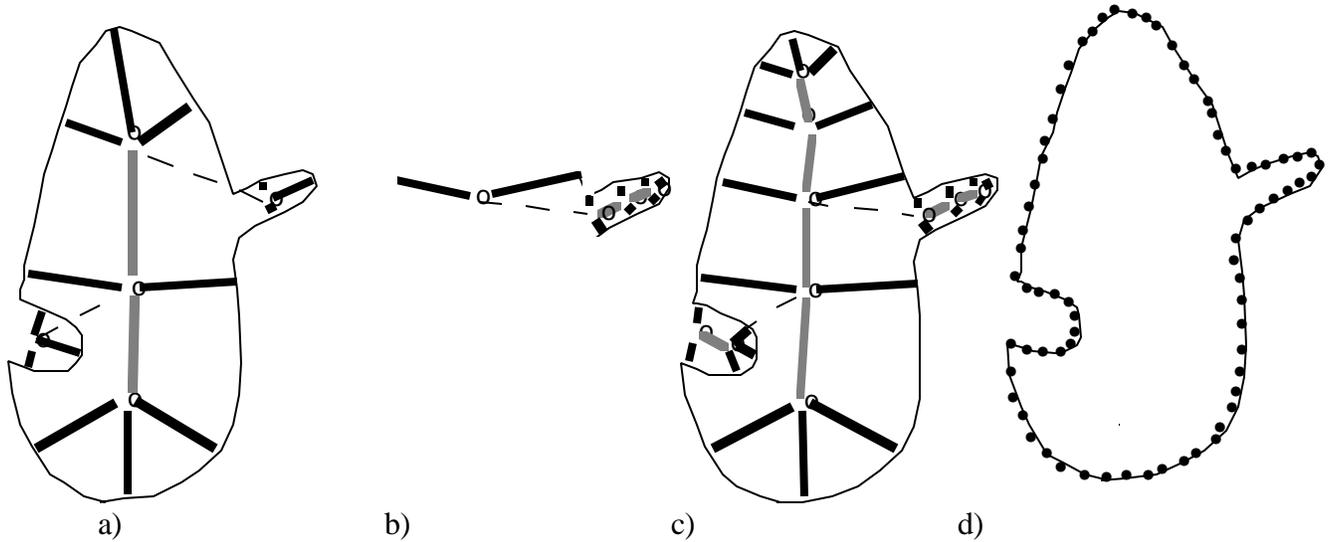

Fig. 6. Coarse-to-fine shape representations: a) Primitives at object level; all can interact with all others. b) Primitives at figural level. Each figure is handled separately, except for its interfigural link to its parent. Within each figure all primitives can interact. c) Individual medial primitives, each interacting only via its intrafigural links to adjacent primitives.
d) Boundary primitives (normals not shown); each can interact only with a few neighbors.

At each stage there are collections of medial primitives that can interact (change in a correlated fashion). Put another way, the shape-tuples formed from the interprimitive links described in section 2.2 can change in a correlated fashion.

At the object stage (Fig. 6a), there are the most coarsely spaced primitives, but all primitives can interact. At the figural stage (Fig. 6b), only primitives within the figure and between the figure and the object (the interfigural link) can interact, and these are at shorter distances than for the object. Thus the primitives are spaced more finely than for the object stage. Moreover, at increasing levels of fineness we treat the parent figure, then its subfigures, then their subfigures, etc. At the medial primitive stage (Fig. 6c), there need be no interprimitive interaction, but there is interaction between the primitive and the primitives at the figural stage. At the boundary primitive stage (Fig. 6d), the spacing is smaller than at the medial primitive stage, and while there can be interaction between nearby boundary primitives, we now include only interaction between each boundary primitive and the adjacent medial primitives. We refine these ideas further in Section 3.

At each stage, the principle of local shape requires that tolerance be proportional to primitive interaction distance. Thus tolerance decreases as one moves through the stages of scale. The primitives at one scale are taken as variations from the values interpolated from the just larger-scale level. The previous level can be thought of as providing a prior for the next smaller level of primitive values. In this way, the coarsest medial nets can represent shape relations at the full length of a figure while finer medial nets represent more local shape relations. All this is done with a representation that still has O(N) links, over all levels of coarseness.

The scheme just laid out may fail when small-scale boundary changes in an object are correlated at a long distance along the figure. For example, the method may not be able to represent the case off a worm's head always twitching when its tail twitches. Representing such information will require either giving up shape locality or finding another means of achieving an O(N) representation.

2.5. *Measurement Aperture Size and Magnification Invariance*

We have seen that different aspects of our shape representations provide boundary information at different tolerances, that is, at different scales. Accordingly, image information can be used to position primitives of an object representation R at different scales by using proportionately sized apertures to measure the property of being on the boundary or being normal to the boundary.

The width of the aperture used to locate a locus in an image is an important variable. In previous writings Koenderink [1984], Lindeberg & ter Haar Romeny [1994], and we [Pizer 1996] have used the word *scale* for this quantity and have gone as far as to say that this is not a parameter but a variable as important as spatial position in characterizing an image. In particular, the aperture width is proportional to the tolerance with which the locus, a boundary in our case, can be determined. Thus, for the object level of representation (see Fig. 6a), the aperture width will be quite large. If measurements are to be made to determine the value of individual medial primitives, the aperture width must be proportional to the r value of the primitive [Pizer 1996] and have significant extent. The aperture width determining boundary displacements must be much smaller.

Burbeck & Pizer [1995] have found that in human vision the aperture width s at which cross-figural linking (medial primitive location) is done is nearly in proportion to figural radial width ($s \approx 0.25r$) and for all but the thinnest of figures is significantly greater than the scale at which boundary location and boundary shape is determined. With such a measurement scale and in figural components with boundaries not too far from parallel, sampling theory would suggest a sampling rate of approximately one medial primitive per cr/3, where c is the elongation of the aperture along the boundary (see Fig. 7 in Section 4). This sampling rate leads to under 16 medial primitives for all but the most elongated 2D figures, providing a considerable efficiency of representation. In the coarse-to-fine regime that we have suggested, even coarser sampling would be used at earlier stages.

We now see that maximum interprimitive interaction distances, tolerance of the implied boundary, and aperture width ought all to be proportional. Therefore, we can broaden our use of the word "scale" to mean any of these.

We close Section 2 with a summary: The representation of the morphography of an object that we are proposing involves a) the similarity transform parameters, b) a linked set of nets of medial primitives, each net possibly in a coarse-to-fine collection, and c) a net of boundary displacement primitives, possibly in a coarse-to-fine collection. The linked nets of medial primitives specify the interfigural relations and the general shape of each figure, and the net of boundary displacements specifies the boundary texture. A linked net of medial primitives, with or without a net of boundary displacements, form a geometric description that imply a boundary with tolerance.

As suggested earlier, this representation can be used for image analysis to extract the most probable instance of a modeled object from an image or to measure the degree of abnormality of this instance. To do this, one needs to be able to measure both how abnormal a represented object is and how consistent it is with a particular image. The next two sections treat these two measurements.

3. Measuring shape anomaly

We take a Bayesian point of view of image analysis, designing methods that extract the most probable object given the image, i.e., that maximize the posterior probability *p(object | image)*, or equivalently *log p(object | image)*. We will refer to the negative of the log of a probability as a "metric". For example, if a probability is a Gaussian, the metric is proportional to a Mahalanobis distance.

Bayes' Theorem implies that maximizing *log p(object | image)* extracts that object representation that minimizes *[-log p(image | object) + -log p(object)]*. The first term, the log likelihood, is a metric of the match between the image data and the object representation. We treat this metric in Section 4 and continue here with *-log p(object)*. This is the -log of the prior probability of this object representation's occurrence in the population of all objects that might appear in the imaged region. It measures how anomalous a shape, pose, and size is relative to some standard.

There are two major issues in this section on measuring shape anomaly. In Section 3.2 we explain how metrics involving medial or boundary primitives can be either statistical or geometric. The statistical metrics measure *-log p(object | image)* based on information extracted from training images. Geometric metrics measure how far the object deviates from a reference object by some combination of spatial distances, size differences, and orientation differences. First, in Section 3.1, we discuss how to find metrics by successive refinement of scale and an associated factoring of the morphographic information into the information at a larger scale and the information at a smaller scale, conditioned on the information at the larger scale.

3.1 Factoring morphographic probabilities coarse-to-fine

The overall morphographic probability participating in the posterior optimization is *-log p(object)*, i.e., *-log p(all primitives at all scales)*. This metric can be simplified by probabilistic factoring by scale level (where levels are numbered by decreasing scale so that level 1 has the largest scale), in a way similar to that described in [Grenander 1994]. There are a number of scale levels, and at each scale level the boundary and a locus of medial primitives is implied at some tolerance, decreasing in order of level. The primitives at each level thus can be considered as displacements from those implied at the previous, larger scale level. Because the information at larger scale largely ignores information at significantly smaller scales, we can use the Markov assumption that the probabilistic dependence of information of displacements at a smaller scale on information at larger scales can be expressed entirely through its dependence on the next larger scale. For example, the boundary displacements are conditioned only on the local medial primitive locations and need not account for the object-level or figural-level information except transitively through the local medial primitive level. While this hierarchical modeling assumption can be expected to be most often valid for objects in medical images, it will fail in situations where there is significant correlation between parameter values at widely different scale levels, e.g., where boundary wrinkling texture is correlated with the position of one of the major figures on the object.

The probabilistic factoring is thus a successive application of the factoring,

p(displacements at scale level k | info. at all larger scales) =

11

p(displacements at scale level k from rep'n interpolated at scale level k-1 | displacements at scale level k-1) ×
p(displacements at scale level k-1 | info. at all larger scales). This can be rewritten
-log p(displacements at scale level k | info. all larger scales) =
-log p(displacements at scale level k from rep'n interpolated at scale level k-1 | tolerances implied by scale level k-1) +
-log p(displacements at scale level k-1 | info. at all larger scales) .

We will consider the successive refinement presented in Section 2.4: object, figure, medial primitive, and boundary primitive. The posterior optimization at each level geometrically transforms the implied boundary and leaves a smaller tolerance than before. The final tolerance is that associated with the boundary displacements.

The first coarse-to-fine stage factors *-log p(all primitives at all scales)* into *-log p(level 1 primitives) +*
-log p(displacements at scale level 2 | level 1 primitives). In the applications reported in Section 5 we have found it possible to restrict the transformation of the level 1 primitives to a similarity transform, so that
-log p(level 1 primitives) = -log p(location, orientation, and size). Similarly, in these applications we have found it possible to restrict the transformation of the figural primitives to a similarity transformation relative to the level 1 primitives, so that *-log p(displacements at scale level 2 | level 1 primitives) =*
-log p(figural translation, rotation, and size relative to the level 1 primitives).
It is possible to include a full-object warp at these two scales, but we leave such a possibility to a later paper. With the formulation we have used, the first stage determines the probability of the similarity transform and the second and following stages determine P_{shape}.

The third stage of the factoring describes the displacement of the individual medial primitives relative to the collection of medial primitives forming a figure. In the applications in Section 5, where the medial primitives representing the figure at scale level 2 are the same as those representing local information at the scale level 3, we have handled the third scale level via a Markov assumption, treating each medial primitive as probabilistically dependent only on the intrafigurally adjacent medial primitives:
-log p(displacement of a medial primitive at scale level 3 | level 2 primitives) =
-log p(displacement of the medial primitive | the intrafigurally adjacent medial primitives). The values of this metric are discussed further in Section 3.2.

The final stage of the factoring treats the boundary displacements along the normal to the boundary interpolated from the medial primitives as dependent only on the tolerance implied by the medial primitives. Chen [1998] has assumed that for all boundary positions i,
boundary displacement$_i$ | linked medial primitives is independent of all other displacements, and he has modeled *p(boundary displacement$_i$ | linked medial primitives)* by a Gaussian with standard deviation equal to the tolerance at implied boundary position i implied by the medial primitives (see Fig. 5). That is, if the standard deviation giving the displacement tolerance at implied boundary displacement i is $\sigma_i \propto$ the figural radius at the medially-implied boundary position,
-log p(boundary displacements | linked medial primitives) =
  $-\boldsymbol{S}_i$ *log p(boundary displacement$_i$ | linked medial primitives) =* $\boldsymbol{0.5}\, \boldsymbol{S}_i\, (displacement_i)^2/\boldsymbol{s}_i{}^2$.
3.2 Statistical and geometric shape metrics for medial primitives

To ease our discussion of metrics for medial primitives, let us adopt the notation $m_i = \left(x_i, r_i, \vec{b}_i, q_i\right)$ for the ith medial primitive, $l_j = \left(\Delta x_j / r_j, \Delta r_j / r_j, \Delta \vec{b}_j, q_j, \Delta q_j\right)$ for the jth link, and $l_j \sim m_i$ for the indication that the jth link involves the ith medial primitive. When the Markov assumptions leading to probabilistic factoring are valid, Gibbs distribution theory allows us to conclude that the shape probability at a medial primitive is given by a local sum of the link metric over the links to the medial primitive: $\|m_i\| = \sum_{l_j \sim m_i} \|l_j\|$. This is the localized shape metric that we have been seeking. Even when combined with the metric components at larger scale, the result has only O(N) terms.

We are working at present on generating statistical (i.e., -log probability) metrics $\|l_j\|$ from families of training images using the principal components approach of Cootes et al. on the incident medial primitive links to a medial primitive. When combined with a similar form of statistical analysis at the larger scales, a global, statistically trained morphographic metric, with explicit local aspects, results. However, the metrics that we have used till now have been geometric instead of statistical. That is, for the local metric, relative to a reference link value l_j^{ref} a link l_j is measured by its geometric difference from the reference, e.g., the difference in its object angle, rather than its degree of unusualness. With this assumption, $\|m_i\| = \sum_{l_j \sim m_i} \|l_j - l_j^{ref}\|_{geom}$.

While we presently recommend that the geometric metric be a weighted sum of metrics on the components, all dimensionless, of the incident $l_j - l_j^{ref}$, some of the image analysis methods that are reported in Section 5 preceded medial primitives including orientation components. They thus used a metric that only involved the position-offset and width-difference components of interprimitive links. This geometric metric was suggested by Wilson [1995]. It uses the size of a vector in scale space, $d_{ss}(\Delta \mathbf{x}, \Delta r, \mathbf{x}, r)$ according to the scale space norm, which takes sizes relative to scale, and the scale space cosine $\cos_{ss}(\mathbf{x}, r; \mathbf{x} + \Delta \mathbf{x}, r + \Delta r)$ between two vectors in scale space [Eberly, 1994].

The shape metric can be used to measure the geometric difference between a link connecting a pair of medial primitives in a particular instance of the shape and the link connecting that pair in a standard version of the shape. It can also be used to measure the difference between two homologous objects, e.g., between a normal organ and the organ of a particular patient or between a patient's organ at one time and that patient's organ at another time. Where the weighted difference is high can be identified to indicate pathology. In addition, the shape metric can be used to measure the difference between the pair in a model and its deformed version in a deformable loci method of segmentation (see Section 5.1). That is, we can produce a formal local Markov random field prior [Wilson 1995], which can be used via the Iterative Conditional Modes (ICM) [Besag 1986] method to optimize the local posterior probability given the linked medial primitives' present values.

4. Measuring accordance of image with linked medial primitives

In this section we focus on the metric *-log p(image | object)*, which measures how consistent an image is with a particular object representation. For the purposes of this section the object representation is assumed to include a set of medial primitives, already placed, oriented, and sized according to values for the similarity transform parameters. Just as with the morphographic probability, the image match probability can be factored by scale if we make a Markov inter-scale-level assumption, resulting in

-log p(image | object) = *-S_k log p(image at scale level k | object displacements at level k and below)*. That is, the image match information needs to be provided at each of the scales at which the object displacements are represented, conditioned on positions established at a larger scale of measurement. When used in posterior optimization (see section 5), the morphographic metric, factored by scale level, and the image match metric, factored by scale level produces and overall metric in which the morphographic and image match terms at a given level are grouped: *object | image metric =*
S_k [-log p(object displacements at level k | object represented at scale k+1) +
-log p(image at scale level k | object displacements at level k and below)]. Each of the terms of the sum are successively optimized, in order coarse to fine. In this section we focus on what happens at any particular level of refinement.

First of all, as discussed in section 2.5, the image measurement at any level must use aperture widths proportional to the inter-primitive distances at that level. The fact that successive scale levels will use derivatives with successively smaller apertures makes reasonable our assumption that the image information at successive scale levels, conditioned on their respective object representations, are independent. Also, the aperture widths at the first, full-object level will be quite large. The use of a large aperture serves to give the first stage of the method a large capture distance. That is, the model may be placed quite far from its final position and still converge to an optimum posterior. For example, if a significant portion of a major figure is within 2 full aperture standard widths away from its final position, convergence to a visually reasonable similarity transform can be expected.

Second, typically the information from the image telling how well it matches with a medial primitive has to do with the behavior of the image intensities near the boundary positions implied by the medial primitive. We assume that this information is probabilistically independent across the primitives, and we make the Markov assumption that the local intensity dependence on all primitives is only on the local primitive. Stated mathematically, *-log p(image | collection of medial primitives) =*
-S_i log p(image at boundary positions implied by primitive i | collection of medial primitives). This metric of the match between an image and a medial primitive is what we have called "medialness" in our previous work.

As discussed in section 2.5, at the stage of match at which the individual medial primitives, relative only to their intrafigural neighbors, are being displaced, the metric of the match between an image and a medial primitive must involve image information measured at a scale proportional to the width parameter of the medial primitive. Moreover, we take it to involve only the image information at the two boundary positions and normals implied by the medial primitive. We assume the image information in these two boundary regions are probabilistically independent. This yields the summary mathematical statement *-log p(image | a medial primitive i or an end primitive i) =*
S_i (-log p((image | 2 or 3 boundary positions and normals i at scale μ r_i)), where the boundary positions, normals, and scale are those implied by the respective medial or end primitive. We call these summands "boundariness" or boundary strength metrics.

As with the shape metrics, there are both statistical and geometric image match metrics. The statistical ones derive from training images and measure the accordance of a particular local distribution of intensities with ones in the homologous positions in the training images [Cootes 1993]. While we are working toward using these, to

date we have used geometric image match metrics, which measure the accordance of the local image intensity function with a template, e.g., a directional Gaussian derivative. Such measures of accordance are standard in image analysis; normally they are measured by treating the template as a weighting function.

By the previous discussion, an image accords with a medial primitive to the degree that its implied boundary points and normals behave in a boundary-like fashion. Using the directional derivative of an aperture function geometric metric, the medial strength of the medial primitive $\boldsymbol{m} = \left(\boldsymbol{x}, r, \vec{b}, \boldsymbol{q}\right)$ is given by the sum of two boundary strength values:

$M(\boldsymbol{m}) = \left(R(\boldsymbol{q})\vec{b}\right) \bullet \nabla A(\boldsymbol{y}, r) \boldsymbol{I}\left(\boldsymbol{x} + rR(\boldsymbol{q})\vec{b} + \boldsymbol{y}\right) + \left(R(-\boldsymbol{q})\vec{b}\right) \bullet \nabla A(\boldsymbol{y}, r) \boldsymbol{I}\left(\boldsymbol{x} + rR(-\boldsymbol{q})\vec{b} + \boldsymbol{y}\right)$, where A is the

aperture function (Fig. 7a). The aperture function should have an elongation suitable to the medial primitive sampling distance. By the theory given in Section 2.5, both the length and the width of the aperture should grow in proportion to r.

If one has only the information in a medial primitive, the along-boundary ridge of each aperture function would be straight and orthogonal to the direction of differentiation, as in Fig. 7. Alternatively, one can use the information in the links from the medial primitive to derive boundary curvature and give this curvature to the aperture function.

The image property whose difference determines the boundary can vary across the boundary points. For example, the polarity of the intensity difference, whether the change at the boundary is in texture or luminance, and whether a luminance change is signaled as an outline (a bar) or a gradual step, and the expected sharpness of the step or bar can all vary from point to point. The training can allow the particular boundary property which is to be measured to be set at each boundary point.

At the endpoint medial primitives $\boldsymbol{m}$, the medial strength $M(\boldsymbol{m})$ adds to the aforementioned formula the property of end strength (Fig. 7b). That is, these end primitives have not only two boundary-pointing vectors that behave just like those in internal medial primitives, but also an additional vector $\vec{b}$ for which the two boundary strength kernels have coalesced into a single operator covering a sector of a circle with radius of curvature equal to the radius of the primitive. There is an additional aperture whose position and orientation is implied by this vector. The spatial aperture size normal to the circle at the implied position of this additional aperture is proportional to the radius of the circle, and the template is differentiated in the $\vec{b}$ direction. Like boundary strength operators, the end strength operator forms a sort of template of intensity variation to which the image data is fit.

For scale levels at which a collection of medial and end primitives are jointly being displaced, according to the probabilistic discussion earlier in this section and the overall agreement of a net of medial primitives $\{\boldsymbol{m}_i\}$ is

$$\text{Image match at medial locus} = \sum_{i=1}^{N} M\left(\boldsymbol{m}_i\right)$$

The items of the sum which are especially small indicate positions where the image information disagrees with the locus.

Similarly, the overall agreement of a set of boundary primitives $\{\boldsymbol{b}_k\}$ derived from displacements from the medially implied boundary positions and normals, is $\sum_{k=1}^{N_b} B(\boldsymbol{b}_k)$, where B can be a statistical or geometric measure of agreement of the image with the model in the normal direction, at a scale smaller than that used in the medial primitive match measurements.
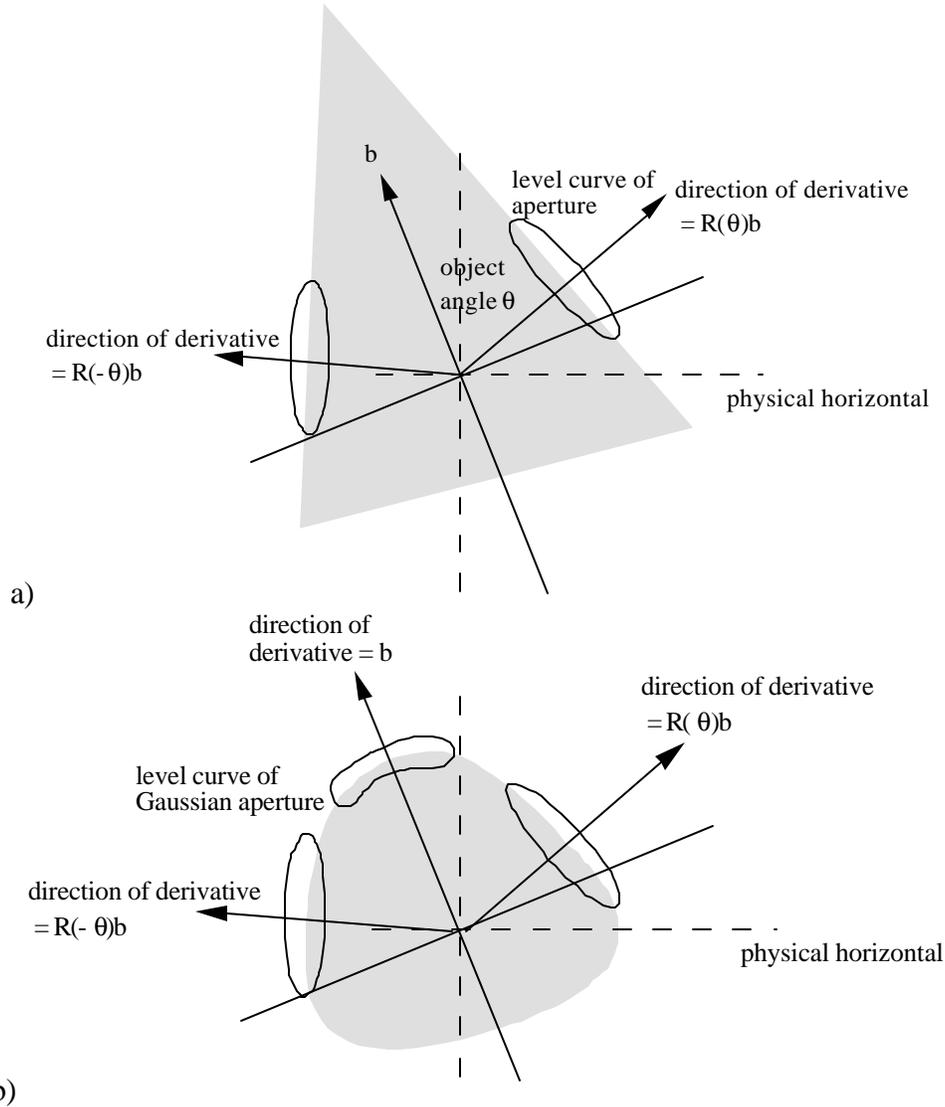

a)

b)

Fig. 7. a) Example of a medial strength kernel. b) Example of an end strength kernel.

5. Shape-based medical image analysis

The examples in the following are intended to give a sense of the possibilities of the image analysis paradigm we have described. Being from the past, they use only geometric image match and geometric shape metrics, and the geometric metrics use only position and width values. Moreover, in some cases the medial strength kernels used are isotropic, summing boundary strength over all object angles rather than just for two or three vector directions.

5.1 Segmentation and recognition

Consider the problem of the definition of anatomic objects, for example, to facilitate 3D radiation treatment planning. Here the objective is to define the boundary of the object that most probably is the organ of interest, so that radiation beams can be positioned with respect to this boundary and dose to these healthy organs can be minimized.

A method of Bayesian deformable loci which generalizes the deformable contours method to use a medial+boundary model has been developed and tested for 2D images [Fritsch 1997]. Given a model, the method involves deforming the boundary and medial loci (moving the points defining these loci) to optimize the Bayesian log posterior formed by summing the log likelihood given by the measure of image match and the log prior giving the measure of morphographic match with the model. This view of segmentation as the deformation of the object representation has led us to call an object represented in this way, Deformable Shape Loci (DSL). The method can be considered not only a method of segmentation, which yields the object boundary, but also a method of recognition that gives the probability that the segmented object is the one modeled and which also provides an indication of pathology, indicating the places where the segmented object is particularly different in shape or morphography from the model.

Methods for first interactively forming a template from a training image and then using a set of training images to determine the weights the model are described in [Wilson 1996] and [Fritsch 1997]. Forming a model, we believe, must involve human interaction, to reflect understanding of which are the object figures that matter. The method for forming a model described in [Fritsch 1997] is based on core extraction (a *core* is a height ridge of medial strength [Pizer 1996]) and a variation of deformable contours, followed by interactive positioning of primitives.

Also described in [Fritsch 1997] is a method for optimizing the log posterior for a particular target image. The method is based on treating the measure of image match described in Section 3 as a log likelihood and treating the measure of morphographic difference from the model described in Section 2 as a log prior. In this work the object morphographic measure, the figural morphographic measure, and the boundary morphographic measure were combined into a single morphographic measure, and the image match measures for these three scale levels were also combined, with the result that the objective function was

$$-\log P = w_b \sum_{i=1}^{N_b} B(\boldsymbol{b}_i) + w_m \sum_{i=1}^{N_m} M(\boldsymbol{m}_i) - w_l \sum_{i=1}^{N_l+1} \left\| \boldsymbol{l}_i - \boldsymbol{l}_i^{ref} \right\|,$$

where the weights w_b, w_m, and w_l give the relative strength of the small-apertured boundary match, the larger apertured medial match used to describe both object-scale and figural scale match, and the image match, and the link N_l+1 is between the parent figure and the image space, i.e., specifies the similarity transformation for the whole object. Ultimately, as discussed in sections 3 and 4, forming the model must also involve a set of training images to determine the variabilities of each term in the objective function. At this early stage, a geometric image match and a geometric morphographic difference were used and the weights of each sum were set to interactively defined constants.

Briefly, the segmentation method uses the model itself is used to start the optimization, and the location, orientation, and magnification parameters of first the full model and then successive figures are first iteratively modified to optimize the posterior. At this stage the part of the prior reflecting intrafigural links (medial-medial, medial-boundary, and boundary-boundary) does not change as the figural shapes are not changed. Then all of

the points on the figural loci are iteratively moved to optimize the posterior, moving medial points and their associated boundary involutes before moving the boundary points themselves. This strategy is particularly effective because the large apertures used for the medial match allow the model to conform to large-scale object changes and, moreover, serve to stabilize the finding of the boundary by the small-aperture boundary points.

All of the previous stages were done with rather coarsely sampled medial and boundary loci. A final stage uses a constrained deformable contours method to fill in boundary points between the coarsely sampled boundary points, which are constrained to their positions at the end of the previous stage.

A series of extractions of the ventricle from 2D MRI slices, using a 2-figure model, has shown the DSL segmentation to be robust against deformation and intensity noise [Fritsch 1997]. A 5-figure model of the ventricle has also been used to automatically segment from 2D MRI images of patients with quite a variety of ventricle shapes [Fritsch 1997]. The model and some results are shown in Figs. 8 and 9, respectively.
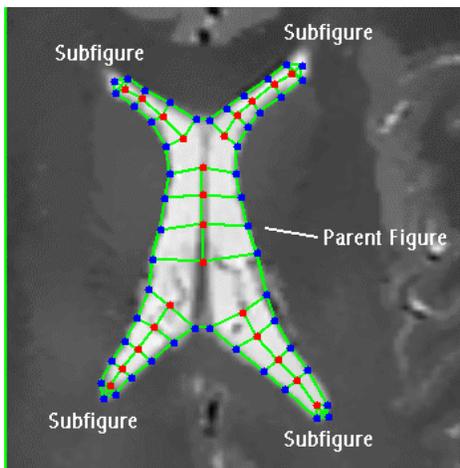

Fig. 8. A five-figure model of a ventricle, with inter-figural links, superimposed on its training image.
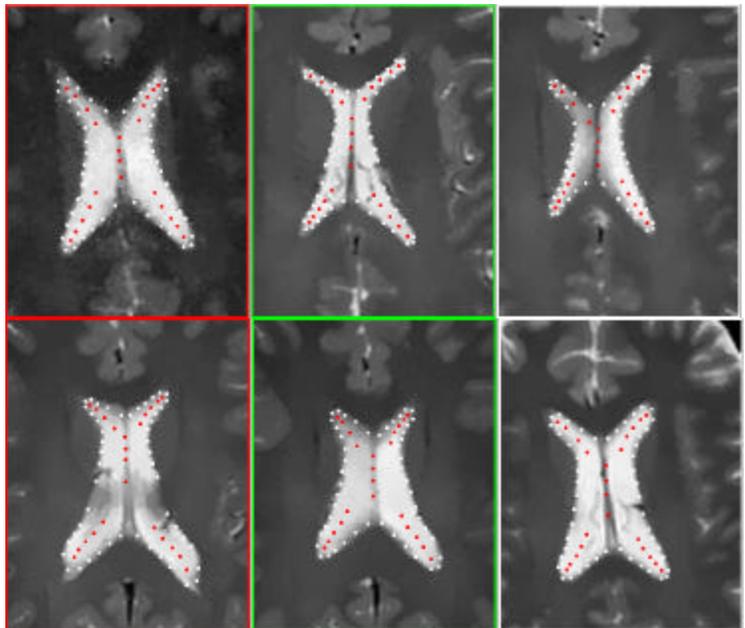

Fig. 9. Final position of the deformed template sites on ventricles taken from a population of images Fig. 9. Final position of the deformed template sites on ventricles taken from a population of images from different patients. The model template used in the optimization is shown in Figure 7.

More recently, this segmentation strategy has been modified to explicitly use the concept of linked medial primitives described in this paper. Here, during the local stage of medial primitive optimization, each medial primitive unit (the medial point and its two associated boundary involutes) move together as a function of position x, orientation $\vec{b}$, object angle θ, and radius r. An example model is shown in Fig. 10.

Shown in Fig. 11 is an example of applying the brainstem model to a target image that was generated via thin-plate spline warping of the model image in Fig. 10. The segmentation algorithm operates as follows: First, the model is initialized by placing it at approximately the right position, orientation, and size in the target image (Fig. 11a). Next, the model is registered with the image by optimizing the image match over the similarity transform parameters (which do not result in shape change penalty) using large-tolerance apertures at the boundary

primitives. Finally, the registered model is warped using an iterative conditional modes (ICM) algorithm where, during an iteration, one after the other each

Fig. 10. Training MR brain image and model of brainstem designed from it.

primitive unit is optimized over its parameters $\left(\boldsymbol{x}, r, \vec{b}, \boldsymbol{q}\right)$. Due to the efficiency of the representation, the optimization program requires only O(N) time for N boundary points, and in our experience frequently fewer than a half dozen iterations of the optimization are typically required. For a typical anatomic structure an iteration calculating new values for the whole set of medial primitives requires about a second at present on a high-end PC.

| a | b | c |

Fig. 11. Segmentation produced from a MR brain images. a) Initial placement of model. b) First-level registration of model with image. c) Final configuration of primitives.

We have also performed segmentations using a hierarchy of objects where the more stably found objects are used to initialize and constrain the positions, sizes, and orientations of subobjects. An example using a CT slice through the abdomen is shown in Fig. 12. Here, only the linearly interpolated boundary primitives are shown. The procedure operates as follows. First, only the large-scale model for the entire abdomen is fit to the image

data via a similarity transformation. This determined transformation is then applied to each of the smaller-scale objects (kidneys, vertebra, and spinal cord). Next, each of the smaller-scale objects is registered via a similarity transformation with the image data but is constrained in translation, orientation, and size change by a link to the large-scale abdomen model. Finally, each of the objects is allowed to deform according to the paradigm outlined at the beginning of this section.
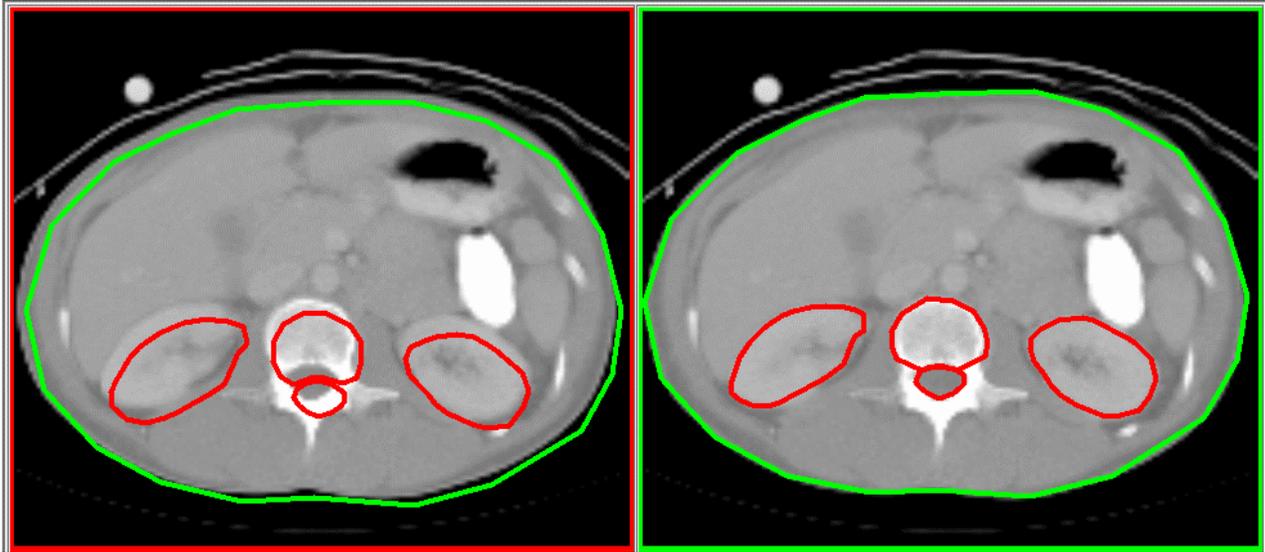

 Fig. 12. Use of an object hierarchy to segment structures in an axial CT slice through the abdomen. (a) Original configuration of object models on an unclassified image. (b) Final template configurations following hierarchical segmentation.

As a final example, Fig. 13 illustrates the ability of the model to interpolate across sections of boundary with little or no image contrast. The shape term dominates the objective function when there is little change in the contrast at the medially implied boundary as the medial primitives are translated.
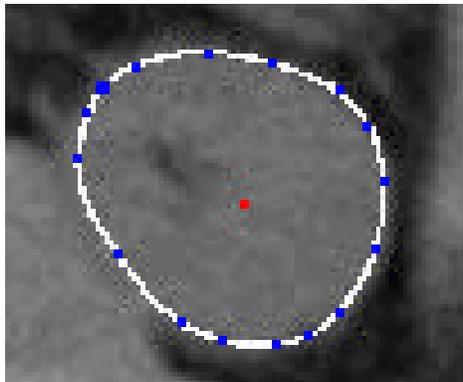

Fig. 13. A model fit to a kidney in a CT slice.

5.2 Registration

In object-based registration the model is extracted from a single reference image of the patient in question, although this model may in turn be the result of a segmentation based on a more highly trained model, as discussed earlier in this paper. After the registration's model extraction, one or more target images may be

registered. For example, to verify the positioning of the treatment beam in radiotherapy with respect to the patient, a portal image obtained during each treatment needs to be registered with a planning image such as a simulator radiograph or a radiograph digitally reconstructed from a CT image.

First assume that the two images to be registered are both in 2D (or both in 3D) and that the registration transformation T does not change shape (i.e., the images may be registered via a similarity transformation). Furthermore, assume that there is no warp between the images to be registered. For this registration objective, the global magnification, translation, and rotation may be chosen as the most probable, given the image data, and thus it might reflect a prior on these transformations. In many situations it may be that knowledge of the prior is not present, so a uniform prior, and thus a maximum likelihood approach ought to be used. In this case the morphographic choice, i.e., the choice of T, is not to be penalized. Thus the objective function, -log P, consists only of the image match terms:

$$w_b \sum_{i=1}^{N_b} B(\boldsymbol{b}_i) + w_m \sum_{i=1}^{N_m} M(\boldsymbol{m}_i).$$

The weights are interactively set, because there is only one reference image and thus no measurement of variability over family of images.

Because of the stability of medial strength measurement, we have found good success in registering 2D portal images, determining a rotation and translation only, by optimizing $\sum_{i=1}^{N_m} M(\boldsymbol{m}_i)$ (using medial strength kernels for which the two normals were in opposite directions and the response was optimized over the normal direction). In fact, Fritsch [1995] shows this registration can be done automatically and produces an accuracy superior to manual registration. For the example in Fig. 14 the registration requires around 10 seconds at present on a high-end PC.
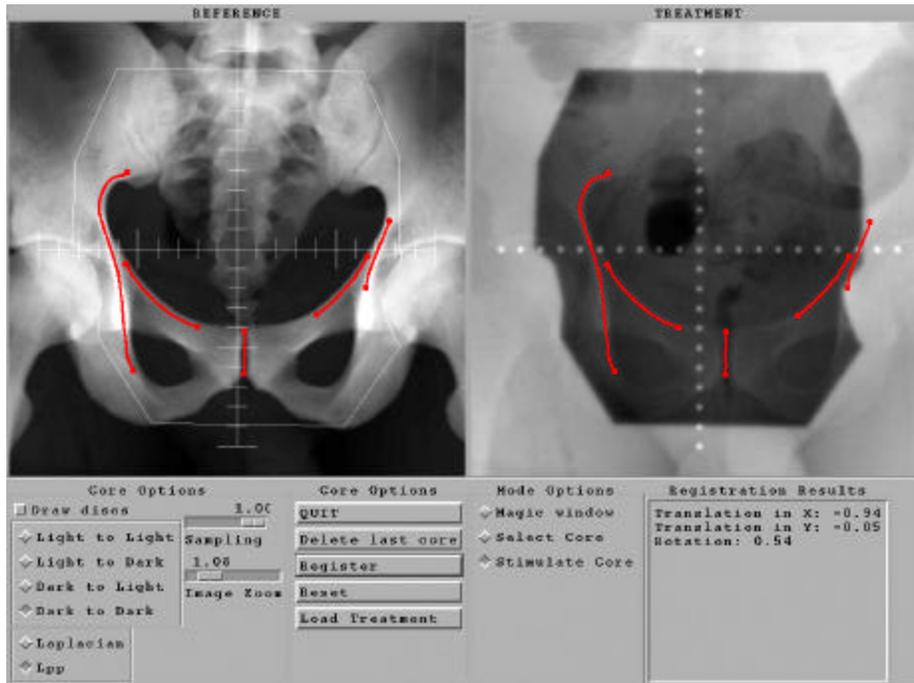
Fig. 14. Registration of 2D radiotherapy portal and simulation images via medial strength in portal image on cores in simulation image. Left: simulation image digitally reconstructed from CT. Right: portal image registered with the simulation image.

In some registrations the geometric transformation allowed may consist not only of a similarity transform but also a warp. In that case, the probability distribution should cover not only the coefficients of the similarity transform, but also the coefficient of the warp terms. Then the situation becomes the same as with segmentation, and all of the discussion of sections 2-4 apply to the registration.

5.3 Measuring shape variation and recognition of pathology

As mentioned at the end of Section 3, the analysis of the sublocus of pathological shape change of an organ can be determined by locating especially large terms in $-\log P_{geom} = \sum_{i=1}^{N_l} \left\| \vec{l}_i - \vec{l}_i^{\,ref} \right\|$.

This expression for $-log\ P_{geom}$ or other expressions for $-log\ P_{geom}$ reflecting a full training set and hierarchical modeling, as discussed in Section 3, can also be used as a measure of shape variation from normal. Thus, for example, in psychiatry it is important to correlate the shape change of various components of the brain: the ventricle, the hippocampus, and putamen to name a few, with schizophrenic disease. Others have used shape models based on boundary primitives [Cootes 1993] or on boundary representation in terms of cyclic basis functions [Székely 1996] for this purpose, with principal component analysis providing the statistical means for turning interactively segmented training images into a measure of shape variation based on Gaussian probability distributions. This measure is then applicable to measure the Mahalanobis distance of an object in a patient image from the mean object shape.

We are applying this same idea to models based on medial primitives. In an early attempt a family of 20 random corpus callosum MR images was created by Aylward as follows. From a corpus callosum in an MR image 5 (x,y,r) medial values approximately equally spaced along the structure were measured. These formed the mean of a 15-variable Gaussian probability distribution. The standard deviations of the distribution were chosen to reflect two independent changes: bulging/narrowing toward the anterior portion and a swing of the tail. Twenty samples from that probability distribution were taken, and each sample, comprising 5 (x,y,r) values, was used as the knots forming splines of x, y, and r, respectively. For each such spline triple, a binary image was formed as the union of the disks of radius r at position x,y, and this image was superimposed on a low contrast midsagittal MRI slice.

We fitted a set of medial primitives to the corpus callosum in each of these 20 simulated images. We started by extracting a core (height ridge in (x,y,r,$\boldsymbol{q}$) of medial strength) from each training image. We then approximated the medial primitive position, radius, and bisector direction at 16 equally spaced positions along the core, fixing the object angle at $\pi/2$. Finally, we optimized each primitive's medial strength by modifying its object angle, also allowing slight variation in position perpendicular to the core, bisector direction, and radius.

The resulting set of medial primitive parameters were analyzed by principal component analysis – this would be the first (full object) stage of a coarse-to-fine shape analysis according to our overall theory. The full coarse-to-fine analysis is part of ongoing research. Fig. 15 shows the first two principal modes of variation of the one-stage analysis. These modes respectively capture the very two independent variations that were inserted to form the training data.
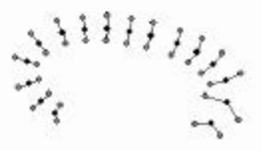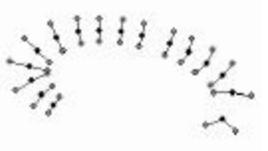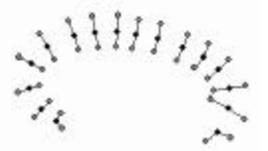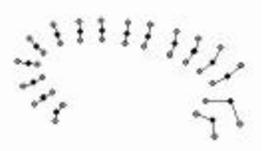
22

	$m\text{-}2\sigma_i$	m	$m\text{+}2\sigma_i$
Eigenmode 1 (i = 1)			
Eigenmode 2 (i = 2)			

Fig 15. Principal shape variations of a corpus callosum from simulated training images. The center column shows the mean medial primitive set, and the left and right columns show respectively +/- two standard deviations of variation of the medial primitives according to that eigenmode.

6. Summary and Discussion

The principal objective of this paper has been to show how to construct a stable, efficiently calculable measure of shape and other geometric object properties and how to use this measure in a uniform method for segmentation and recognition of image objects, object-based registration, and object shape measurement that has a large capture distance. The abilities of this method have been illustrated by examples. A more detailed description of the segmentation programs and a presentation of the segmentation results of these methods can be found in [Fritsch 1997]. For some of the clinical uses described in Section 4 the method has been validated by clinical usefulness or accuracy [Fritsch 1995, Fritsch 1997]. In others, this validation is yet to be done.

The measure of shape that we have constructed is dependent on a variety of reasonable probabilistic assumptions of the hierarchical dependence of information at the collection of scale levels and of the independence and Markov properties of intensity vs. primitive-implied positions and directions. A careful study of the realizability and uniqueness of these intensity states remains to be done. Such hierarchical modeling is not new to shape based image analysis, having been applied to voxel-flow descriptions of shape by Christensen [1997] and Grenander [1994], as well as by McCulloch et al. [Laading 1998] in author Valen Johnson's laboratory. However, our application of hierarchical modeling to the richer medial and boundary primitives appears to add capability and efficiency to the approach, and it brings a deep understanding of the ideas of scale in shape theory to bear. At the same time, it raises issues of how to handle abutment relations between model figures (e.g., adjacent organs).

One area of promise of the representation that we have described is in achieving homology between primitives as the representation deforms. With the rich collection of parameters contained in medial primitives and their relation to boundary primitives, not only the loci of vertices (extrema of boundary curvature; crests and troughs in 3D) but also figural loci such as end points (curves in 3D), pinch points (curves in 3D), and medial axis vertices (articulations?) can be designated. However, it remains to future research to measure the extent to which the present metrics cause correspondences at such loci to be maintained and to design metrics that accomplish this goal. Moreover, because it is not always the case that biological homology corresponds to geometric homology, the import of the promise to achieve geometric homology of primitives is yet to be understood.

Compared to a deformable model based on boundary points alone and only on the relations between adjacent boundary points, our experience is that the DSL segmentation method much more reliably converges to a satisfactory result. A major advantage appears to be in the first stages of the segmentation, in which the capture distance from an initial placement, orientation, and scaling of the model is considerably greater than methods based on boundary points with small apertures. A capture distance of about half the width of the parent figure is typical. As compared to methods using all $O(N^2)$ relations between the N boundary points, the DSL method is very much faster.

This paper's first motivation of DSLs as shape models was that both model formation and segmentation take time O(N). But the method of Cootes & Taylor [1993] requires only time O(N) if a fixed number of eigenvectors of shape change are selected, and in many cases it has rather good capture distance. DSLs have a different advantage over this method. The first is that the training requires only O(N) time, as compared to the minimum of $O(N^2)$ and O(the number of training images) for the Cootes & Taylor method. Second, and more importantly, the training requires fewer training images because the number of variables at each scale stage, and thus the number of major modes of variation at each scale stage is few. Third, whereas the Cootes & Taylor method uses the prior information as to shape to constrain the deformation, DSL segmentation uses the prior in a weighted fashion. As a result, the DSL method has the possibility of converging to a pathological version (one with low prior probability) of the DSL model if the image information is convincing enough. Fourth, the DSL method has the advantage of being able to determine the locality of a shape difference (or medial strength or boundary difference), compared to the training set. The DSL method has the disadvantage that, unlike the boundary based methods, medial primitives cannot handle sections of boundary without also handling their involutional partners. However, all of this comparison is hypothetical; the Cootes & Taylor method has been found successful on a very large number of segmentation problems, whereas the DSL method has been tested on rather few problems.

A difficulty with the expressions for P_{geom} that we have discussed is that interfigural homology is required between a model and an individual. While many organs that retain such homology over normal patients exist, many, such as blood vessel trees and the cerebral cortex, do not. In these the figure-subfigure structure is not deterministic but stochastic. Similar differences in the figure-subfigure structure may occur with pathology. A theory for the geometric difference between figural graphs with different and possibly stochastic structures remains to be developed.

We have generalized to 3D of the DSL representation, the measure of image match, the model creation tools, and the optimization of the posterior, and we have implemented both model building and posterior optimization tools. The model differs only by the boundary and medial loci forming 2-manifolds rather than 1-manifolds, with the consequence that the intra-medial-locus links form a mesh except at endpoints, and by the medial primitive benefiting from having the normal to the medial locus as an additional component. The morphographic difference measure does not change at all, except that the number of dimensions in links, translation, and orientation is one higher than with 2D images. The measure of image match differs only in the boundary strength kernels contributing to both boundary strength and medial strength are one dimension higher than with 2D images. A paper describing this 3D medial primitive based deformable models method awaits our carrying out a satisfactory evaluation of its performance.

Acknowledgments

References

Besag, J.E. (1986). On the statistical analysis of dirty pictures. *J. Royal Statistical Society, Series B*, **48**: 259-302.

Blum, H (1967). A new model of global brain function. *Perspectives in Biology & Medicine*, **10**: 381-407.

Blum, H & RN Nagel (1978). Shape description using weighted symmetric axis features. *Pattern Recognition*, **10**: 167-180.

Bookstein, FL (1991). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press.

Burbeck, C.A., Pizer, SM (1995). Object representation by cores: Identifying and representing primitive spatial regions. *Vision Research*, **35**,13:1917-1930, 1995.

Burbeck, C.A., Pizer, SM, Morse, BS, Ariely, D, Zauberman, G, Rolland, J (1996). Linking object boundaries at scale: a common mechanism for size and shape judgements. *Vision Research.*, **36**, 3:361-372, 1996.

Chen, D (1998). *Volume Rendering Guided by Multiscale Medial Models*. PhD dissertation, Univ. of NC Dept. of Computer Science.

Christensen, G, S Joshi, M Miller (1997). Volumetric transformation of brain anatomy. *IEEE Trans. Med. Imaging* **16**(6): 864-867.

Clary, GJ, SM Pizer, DS Fritsch, JR Perry (1997). Left ventricular wall motion tracking via deformable shape loci. In HU Lemke, MW Vannier, K Inamura, eds., *CAR '97, Computer Assisted Radiology and Surgery,* Elsevier Science, Amsterdam: 271-276.

Cootes, TF, A Hill, CJ Taylor, J Haslam (1993). The use of active shape models for locating structures in medical images. *Information Processing in Medical Imaging*, HH Barrett & AF Gmitro, eds., Lecture Notes in Computer Science **687**: 33-47, Springer-Verlag, Berlin.

Eberly, D (1994). Chapter 14: A differential approach to anisotropic diffusion, Chapter 4, *Geometry-Driven Diffusion in Computer Vision*:, Kluwer Academic Publishers, Dordrecht: 371-392.

Fritsch, DS, EL Chaney, A Boxwala, MJ McAuliffe, S Raghavan, A Thall, JRD Earnhart (1995). Core-based portal image registration for automatic radiotherapy treatment verification. *International Journal of Radiation, Oncology, Biology, Physics*; special issue on Conformal Therapy, **33**(5): 1287-1300.

Fritsch, DS, Yu L, Johnson V, McAuliffe MJ, Pizer SM, and Chaney EL (1996). A probabilistic approach using deformable organ models for automatic definition of normal anatomical structures for 3D treatment planning. *International Journal of Radiation, Oncology, Biology, Physics*, **36** (Supp. 1):187 (Abstract).

Fritsch DS, Pizer SM, Yu L, and Chaney EL (1997). Segmentation of medical image objects using deformable shape loci. *Proc. Information Processing in Medical Imaging* (IPMI97), J Duncan and G Gindi, eds., Springer-Verlag, Lecture Notes in Computer Science 1230:127-140.

Frome, FS (1972). A psychophysical study of shape alignment. Technical Report TR-198, University of Maryland, Computer Science Center.

Grenander, U & MI Miller (1994). Representations of knowledge in complex systems. *J. Royal Statistics Soc.* **56**: 549-603.

Kelemen, A, G Székely, G Gerig (1998). Three-dimensional m4odel-based segmentation. In B Vemuri, ed., *Proc. Workshop on Biomedical Image Analysis*, IEEE Computer Science Press: 4-13.

Kendall, DG (1984). A survey of the statistical theory of shape, *Statistical Science* **4**: 87-120.

Koenderink, JJ (1984). The structure of images. *Bio.. Cybernetics* **50**:363-370.

Laading, J, C McCulloch, V Johnson, D Gilland and R Jasczak (1998). A hierarchical feature based deformation model applied to 4d cardiac SPECT data. Discussion paper, Duke U. Inst for Statistics & Decisions Scis., www.isds.duke.edu, 98-33.

Leyton, M. (1992). *Symmetry, Causality, Mind*, The MIT Press, Cambridge, Massachusetts.

Lindeberg, T, BM ter Haar Romeny (1994). Chapter 1: Linear scale-space I: Basic theory, and Chapter 2: Linear scale-space II: Early visual operations, *Geometry-Driven Diffusion in Computer Vision*:, Kluwer Academic Publishers, Dordrecht: 1-72.

Liu, A., Bullitt, E., Pizer, S. M. (1998). 3D/2D registration via skeletal near-projective invariance in tubular objects. In WM Wells, A Colchester, S Delp eds., *Medical Image Computing and Computer-Assistend Intervention – MICCAI '98.* Lecture Notes in Computer Science **1496**: 953-963, Springer.

Liu, A, E Bullitt, S. M. Pizer (1997). Surgical instrument guidance using synthesized anatomical structures. *CVRMed-MRCAS '97*, J Troccaz, E Grimson, R Mösges, eds., Lecture Notes in Computer Science **1205**: 99-108.

Morse, BS, SM Pizer, DT Puff, C Gu (1996). Zoom-invariant vision of figural shape: effects on cores of image disturbances. *Computer Vision and Image Understanding*, **69**:72-86, 1998.

Pizer, SM, D Eberly, BS Morse, DS Fritsch (1996). Zoom-Invariant Vision of Figural Shape: The Mathematics of Cores. *Computer Vision and Image Understanding*, **69**:55-71, 1998.

Pizer, SM, Fritsch, DS, Low, KC, Furst, JD (1998). 2D & 3D figural models of anatomic objects from medical images. *Mathematical Morphology and Its Applications to Image Processing*, HJAM Heijmans, JBTM Roerdink, eds. (invited paper, Proc. ISMM '98), Kluwer Computational Imaging and Vision Series: 139-150.

Psotka, J (1978). Perceptual processes that may create stick figures and balance. *Journal of Psychology: Human Perception & Performance* **4**: 101-111.

Staib LH. JS Duncan (1992). Boundary Finding with Parametrically Deformable Models, *IEEE Trans PAMI* **14**(11):1061-1075

Staib LH, JS Duncan (1996). Model-based Deformable Surface Finding for Medical Images, *IEEE Trans Med Img*. **15**: 720-731.

Székely, G, A Kelemen, C Brechbühler, G Gerig (1996). Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models. *Medical Image Analysis* **1**(1): 19-34.

Wilson, A, V Johnson, SM Pizer, D Fritsch, L Yu, E Chaney (1996). Towards a Framework for Automated Image Analysis, Proc. 16th Leeds Annual Statistical Workshop: Image Fusion and Shape Variability Techniques. Univ. of Leeds Press, Leeds, England.

Wilson, AG (1995). *Statistical models for shapes and deformations*. Ph.D. dissertation, Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina.

Yuille, AL, PW Hallinan, DS Cohen (1992). Feature extraction from faces using deformable templates. *IJ Comp Vis* **8**(2): 99-111.