Introduction

Stephen Pizer and Kaleem Siddiqi and Paul Yushkevich

Abstract In the late 1960s Blum (1967) first suggested that medial loci, later generalized and called symmetry sets and central sets by mathematicians (Yomdin, 1981; Mather, 1983; Millman, 1980), would provide an effective means of representing objects appearing in 2D images. Soon thereafter Blum suggested the extension of medial loci to objects in 3D images. These object representations have since come to have an important role in the description of shape. In this introductory chapter we explore the forms the medial representation of objects can take, and we give some history of its development and use. To start, Section 1 places the medial representation of objects into the context of alternative representations.

This chapter sets the context for the subsequent material covered in this book, which is organized in three parts. Part I of the book describes in detail the the mathematical properties of the medial representation and the relation between it and the corresponding object boundary. Section 2 of this chapter introduces this material by giving the basic mathematical definition and properties of the medial representation. Section 3 cites evidence for medial representations as models of human vision. Part II of the book presents algorithms that have been proposed for going from a boundary to a medial representation or from a medial representation to a boundary representation, and it discusses their performance. Section 4 of this chapter overviews such methods, including some that are not detailed later in the book. Finally, Part III of the book covers selected applications of the medial representation in image

Stephen Pizer
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: pizer@cs.unc.edu

Kaleem Siddiqi
School of Computer Science & Centre for Intelligent Machines, McGill University, Canada
e-mail: siddiqi@cim.mcgill.ca

Paul Yushkevich
Department of Radiology, University of Pennsylvania, USA
e-mail: pauly2@grasp.upenn.edu

analysis. These include segmentation, shape characterization, recognition, object labeling, and registration. Section 5 of this chapter introduces these applications.

1 Object representations

A variety of alternative means of representing objects or multi-object ensembles in 3D or 2D have appeared. The main alternatives to the medial representation fall into three categories. In *landmark representations* (Fig. 1e) an object or object ensemble is described by an ordered set of geometrically recognizable and salient locations on the object(s). Typically this is a sparse set, due to the difficulty of manually or automatically extracting a large set of landmarks from the 2D or 3D image data. In *boundary representations* (b-reps) the object or objects are represented by a relatively dense set of points sampling its (their) boundary (Fig. 1a), by a mesh of tiles whose vertices form a boundary sampling set but from which normal and curvature information can also be extracted (Fig. 1b), or by an orthogonal function decomposition of the boundary surface, e.g., by Fourier (spherical, in 3D) harmonics (Fig. 1c). In this decomposition the representation is formed by the collection of coefficients of the basis orthogonal functions. In *displacement by voxel representations* the object is represented by an atlas and a displacement vector field (Fig. 1d). The atlas is formed as a template label image giving for each voxel the name of the object class (including background as a class) in which the voxel falls in a template image. The displacement vector field gives a displacement of each voxel and thereby describes a mapping of space and thus of the object labels from the template image to a particular instance of the object or object ensemble. Finally, there is the medial representation (Fig. 1f), which, like the displacement by voxels representation, directly represents the object interior.

Landmark representations have a long and rich history. They have the major strength that, by the very choice of the landmarks, there is a strong correspondence between each specified landmark in one image and that in another image in the population. While their sparseness typically limits them to object descriptions of large spatial scale and the statistics has been largely in terms of positional displacements and not local changes of orientation or magnification, there has been an extensive study of their geometry and their statistics. Prolific and early authors on this subject are Bookstein (1991) and Kendall (1989). A very effective summary of statistics of landmark geometry is provided in Dryden and Mardia (1998). Methods to compute diffeomorphic warps from landmark displacements can be found in Joshi and Miller (2000). Medial representations typically capture many of the landmark correspondences but also allow the study, at many spatial scale levels, of such additional transformations as bending, widening, and elongation.

Boundary representations via tile meshes, (e.g., Delingette (1999)) provide all of the capabilities of boundary representations via points and share the difficulty of analysis at larger scales. In addition, they allow access to boundary normals and cur-

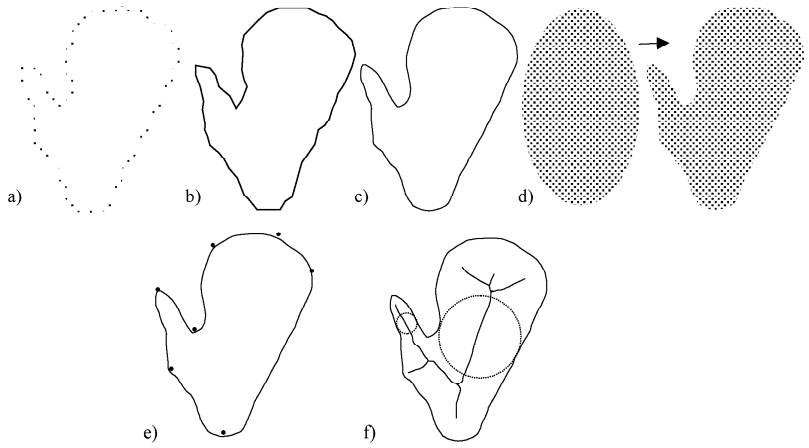


Fig. 1 Object representations of a mitten shape. a) Boundary points. b) Boundary tiles. c) Fourier harmonics. This parametrized representation is by coefficients of sinusoids on the parameter. d) Atlas displacements with labels. Shown are the binary labels (inside vs. outside the object) in a base space and the new locations of the labels after the displacements have been applied. e) Landmarks. f) Medial. Shown is the continuous locus of centers of tangent disks (spheres in 3D) and two of these disks, which in the representation are given at every center point. Notice how the natural landmarks, at boundary vertices, correspond to endpoints of the internal medial axis (shown) or the external medial axis (not shown)

vatures and thus to the analyses of differential geometry. Nevertheless, they remain shell and not interior representations.

Boundary representation via basis functions includes both spline fits and orthogonal function decomposition. These represent an object of a particular topology (e.g., spherical topology) by a smooth continuous mapping from the surface of a standard object of the same topology (in our example, a sphere). This vector mapping function can be decomposed into orthonormal functions (Kelemen et al., 1999). Being smooth functions, all derivatives of the object surface are available, and hence boundary normals and curvatures can be derived analytically. Nevertheless, interior properties are not directly accessible.

The displacement by voxel representation gives easy access both to boundaries and to interiors. It requires methodology to limit the displacement maps to diffeomorphisms (1-to-1 smooth warps); these typically involve velocity fields integrated via differential equations (Grenander and Miller, 1998). The difficulties of this representation derive from the large computer storage space required by the representation, the correspondingly large computations needed to derive and analyze these representations, and the difficulty of determining large scale properties from this small scale representation. However, both displacement by voxel and displacements of boundary representations provide a useful small scale description when combined with a medial representation describing the larger spatial scales.

The medial representation has special strengths in directly capturing various aspects of shape, in giving direct access to both object interiors and object boundaries, and in providing rich geometrical relationships within objects. In the remainder of this chapter we give a careful definition of this representation, and we introduce some of its basic mathematical properties. In addition we introduce some of the basic algorithms for extraction of the medial representation of an object, and we sketch some of the applications of this representation in computer vision and image analysis that will be covered in this book.

2 Medial representations of objects

The medial approach to representing an object (Figs. 1f, 2) is to describe a locus midway between (at the center of a sphere bitangent to) two sections of boundary and to give the distance to the boundary (called the medial radius), yielding the object as the union of overlapping bitangent spheres. One way to think of this representation is a locus of $(\mathbf{p}, r)$, where $\mathbf{p}$ gives the sphere center and r gives the radius of the sphere. In some representations the vectors from the medial point to the two or more corresponding boundary points are included; in others they are derived via first derivatives of the medial locus and of the radius function on that locus. When these vectors are included, the primitive, called a medial atom, is a hub point $\mathbf{p}$ with two equal-length spokes $\mathbf{S}^{\pm 1}$ (Fig.2 b). Blum (1967) described the interior medial locus by restricting the bitangent spheres to those entirely contained in the object's interior. The spokes of these medial atoms do not overlap and sweep out the object's interior.

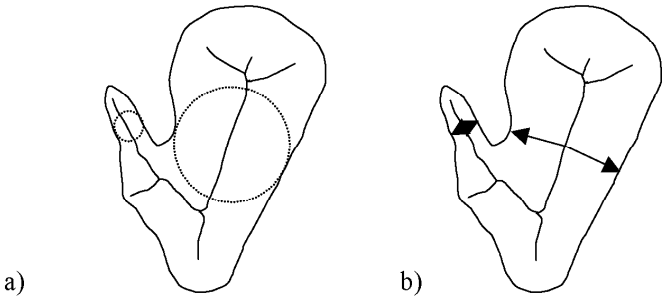


Fig. 2 a) The medial spheres representation: a continuous locus (or collection of continuous loci) of medial spheres (disks in 2D). b) The medial atoms representation, a continuous locus (or collection of continuous loci) of medial atoms, where a medial atom consists of a hub and two spokes.

The Blum medial axis is a transformation of an object boundary with the same topology as the object. Thus the boundary can generate the medial locus of $(\mathbf{p}, r)$ or the medial atoms, but equally the medial locus, in either form, can also generate

the object boundary. In the first direction the transformation is a function, but in the second direction it is one-to-many, since a medial point describes more than one boundary point. In the boundary-to-medial mapping, low level boundary noise can be transformed into large changes in the medial locus. Therefore, the inverse of the boundary-to-medial mapping, which maps in the medial-to-boundary direction, turns certain large changes in the medial locus into small changes in the boundary. We cover methods working in both directions in this book.

One of the strengths of using the medial representation as a primitive is that any unbranching, connected subset of the medial locus generates intrinsic space coordinates for the part of the object interior that the medial atoms in that subset reach. These coordinates can be thought of as follows: coordinate system location in the medial sheet, choice of spoke (left or right) and fraction of the spoke length along that spoke.

Two alternative definitions of the Blum medial locus have appeared. 1) In the so-called "grassfire" or "Eikonal flow" definition the locus is the set of quench points along with their times of formation when a fire burning on equally dense grass within the object is lit at time $t = 0$ at all points on the object boundary. This has been shown to be equivalent to the location of shocks in a partial differential equation of motion at fixed speed in a direction initially normal to the object boundary (Kimia et al., 1995). 2) The "Maxwell set" definition of the medial locus is the set of locations internal to the object with more than one corresponding closest boundary point in the sense of Euclidean distance (Mather, 1983). Each point in this set is augmented with its distance to the boundary.

The Blum medial locus generalizes to the locus defined by all bitangent spheres. This locus is called the *symmetry set* and has been analyzed in some detail in the literature (Bruce et al., 1985; Giblin and Brassett, 1985). As an object deforms, the symmetry set can undergo a complex deformation. Likewise, as discussed in Chapter 3, as the medial locus and the associated spokes deform, the corresponding boundary can undergo a complex deformation. Other medial loci than the Blum medial locus and the symmetry set have also been derived from the set of spheres bitangent to the object. Asada and Brady (1983) suggested the use of the locus of middles of the chord connecting the two points of bitangency. This representation is also referred to as the *midpoint locus* and its properties in relation to that of the symmetry set are studied in Giblin and Brassett (1985). It is the basis for the skeletonization technique developed in Zhu (1999). This locus turns out to be of special interest when specialized to the family of tubes, i.e., objects with circular cross-sections, whose skeletons turn out to be curves in 3-space. In this case the chord center becomes the center of a disk orthogonal to the skeletal space curve.

Leyton (1992) described another locus derived from the set of spheres bitangent to the object boundary. This locus is defined by connecting the two points of bitangency along the shortest-distance geodesic path on the bitangent sphere and taking the associated point to be the center of this geodesic path. Leyton called this locus the Process Induced Symmetric Axis, or PISA. In his theory more complex objects are formed by protrusion or indentation processes acting on the boundary of simpler ones (Leyton, 1988, 1989). A formal justification for this view is a symmetry-

curvature duality principle whereby the endpoints of this construction are related to curvature extrema on the object's boundary (Leyton, 1987; Yuille and Leyton, 1990).

Both the locus of chordal centers and the locus of geodesic centers require more information than that required by the Blum medial axis, namely just a point and a distance. Whereas the midpoint locus has been shown to be less singular than the symmetry set (Giblin and Brassett, 1985), reconstruction of the object from it is far from clear. Since neither of these medial loci have found broad acceptance or application, we will not cover them further in this book.

Medial loci based on spheres change topologically as an object deforms. This behavior was motivation to seek a medial axis that is stable under affine deformation. Such an axis, based on a consideration of affine area enclosed by a region was invented by Betelu et al. (2000). This construct is also interesting because it, unlike the Blum locus, is insensitive to noise. However, it is unintuitive and is very new, and as a result it has so far received little study or use. Therefore, we will not cover it in this book.

The major disadvantage of medial representations based on bitangent spheres is the fact that the branching structure of the medial locus is very sensitive to the small scale geometric aspects of the object. In Fig. 2 the lower joint of the thumb produces its own longish branch, and though this may be desirable in this case, an even smaller pimple due to noise would also produce a similar branch. Blum noticed this deficiency and coined the term ligature (meaning glue) to describe such portions of the medial axis that describe only small portions of the boundary such as regions of high negative curvature. Such ideas have lead to modifications of the medial axis to identify salient portions of it (August et al., 1999; Katz, 2002) but can also be overcome in other ways, including the marrying of the medial representation to the displacement by voxel representation or by displacement of the boundary representation at small scale.

The medial representation has a variety of strengths and as a result we have written this book. In particular:

1. It is an interior representation of the object and thus is subject to both geometric and mechanical operations applicable on the object's interior, such as bending, widening, elongation, and warping.
2. It provides rich geometric information, giving simultaneously locational, orientational, and metric (size) description in any locality of the interior and near exterior of an object.
3. It provides a basis for description at multiple spatial scales and thus provides efficiency of computation and efficiency in the number of population samples needed to estimate object geometry probabilities.
4. It allows one to distinguish object deformations into along-object deviations, namely elongations and bendings, and across-object deviations, namely bulgings and attachment of protrusions or indentations.
5. Its branches at the larger scales divide objects in a way that makes automatic object recognition effective.

6. It provides descriptions of objects and their geometric transformations that are intuitive to nonmathematical users.
7. It generates object-relative coordinate systems for object interiors and their near neighborhoods that provide useful correspondences across instances of an object.
8. It provides a means for describing the locational, orientational, and size relations between one object and a neighboring region of another object within a complex of objects (see Chapter 8).

It is often the case that the interior of an object is described as a binary image, e.g., a set of pixels in $\mathbb{R}^2$ or a set of voxels in $\mathbb{R}^3$. A frequent algorithmic objective for researchers is to transform this binary image, or equivalently the curve or surface that forms its boundary, to its medial axis. In this book we refer to this operation as the *Medial Axis Transform* or the MAT. If the output of the MAT is discretely represented, i.e., as a set of pixels or voxels, we will refer to the operation as the Discrete MAT.

In the remainder of this section we define medial loci and some of their properties in detail, discuss their history, sketch the basic algorithms for their extraction from boundaries and overview their applications in computer vision. This material is organized as follows. Section 2.1 precisely defines the *Blum medial locus*. Section 2.2 describes the structural geometry of Blum medial loci and their use for decomposing objects into simple figures. Section 2.3 explains the local geometric properties of Blum medial loci. Finally, Section 2.4 defines medial atoms and describes their use in skeletonization via deformable models. This section also provides an introduction to m-reps which are used as an object representation, with medial atoms as the building blocks.

2.1 The Definition of the Medial Locus

Medial loci enjoy wide use in computer vision and image analysis, as well as in other fields of computer science such as graphics, computer aided design, and human-computer interfaces (Bloomenthal and Shoemake, 1991; Sherstyuk, 1999; Storti et al., 1997; Blanding et al., 1999; Igarashi et al., 1999). The modern interest in medial loci originates with the work of Blum, who defined the medial locus of a two-dimensional object, studied its geometric properties, and noted its usefulness for describing the shape of objects (Blum, 1967; Blum and Nagel, 1978). While the definition itself is deceptively simple, the thorough understanding of the properties of the medial locus requires rigorous mathematical treatment. Only recently have a number of mathematically rigorous studies of the medial loci of higher-dimensional objects been published (Damon, 2003; Giblin and Kimia, 2002, 2004). These studies and their extensions are covered in Chapter 2 and Chapter 3.

We begin by defining the medial locus using a formulation that slightly extends Blum's original definition. Later in this section we mention some alternative definitions of the medial locus. The basic element of Blum's definition is the *maximal inscribed ball*.

Definition 0.1. Let S be a connected closed set in $\mathbb{R}^n$. A closed ball $B \subset \mathbb{R}^n$ is called a *maximal inscribed ball* in S if $B \subset S$ and there does not exist another ball $B' \neq B$ such that $B \subset B' \subset S$.

Next let us give a formal definition of the term *object* to avoid any ambiguity that the use of this word may bring.

Definition 0.2. A set in $\mathbb{R}^n$ is called an n-dimensional *object* if it is homeomorphic to the n-dimensional closed ball.

Let Ω denote an n-dimensional object and let $\partial\Omega$ denote its boundary. We are now ready to define the medial locus of Ω.

Definition 0.3. The *internal medial locus* of Ω is the set of centers and radii of all the maximal inscribed balls in Ω.

Definition 0.4. The *external medial locus* of Ω is the set of centers and radii of all the maximal inscribed balls in the closure of $\mathbb{R}^n/\Omega$.

Definition 0.5. The *Blum medial locus* of Ω is the union of its internal and external medial loci.

Definition 0.6. The tuple $\{\mathbf{p}, r\}$ that belongs to the Blum medial locus of an object Ω is called a *medial point* of Ω.

The medial locus is thus a subset of the space $\mathbb{R}^n \times [0, +\infty]$. We will sometimes use the term medial locus to refer just to the set of the centers of the maximal inscribed balls, forgetting about their radii. The terms *skeleton*, *medial axis*, and *symmetric axis* have been used by other authors to describe both the internal medial locus and the entire medial locus as a whole, with or without the inclusion of the radial component.

It turns out that the medial locus consists of a countable number of manifolds whose codimension in the space $\mathbb{R}^n \times (0, +\infty)$ is no less than 2. Hence, the medial locus of a two-dimensional object consists of a number of curves and isolated points, and the medial locus of a three-dimensional object consists of surface patches, curves, and isolated points. The manifolds composing the internal medial locus lie inside the object and are bounded, while the manifolds in the external medial locus lie outside of the object and extend to infinity. Fig. 3a. shows the internal and external medial loci of a two-dimensional object.

Moreover, it turns out that the inscribed disks whose centers and radii compose the medial locus of an object generically[1] are bitangent to the object's boundary. In fact, the medial locus is a subset of a more general geometric construct called the *symmetry set*, defined as the closure of the locus of centers and radii of all balls bitangent to the boundary of an object (Giblin and Brassett, 1985). The balls that generate the symmetry set are not restricted to lie either inside or outside of the

[1] The word *generic* is being used in a precise mathematical sense, as defined in the glossary.

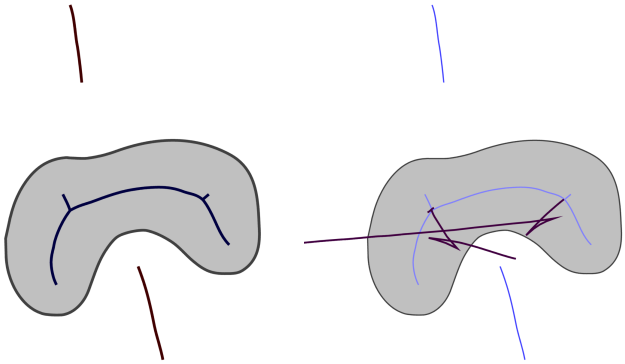


Fig. 3 a. The internal and external medial loci of an object. b. The symmetry set of the same object, which contains the internal and external medial loci (shown in blue) as well as some cusped structures.

boundary of the object. Hence, in addition to the medial locus, the symmetry set of an object contains connective structures such as the cusps shown on Fig. 3b.

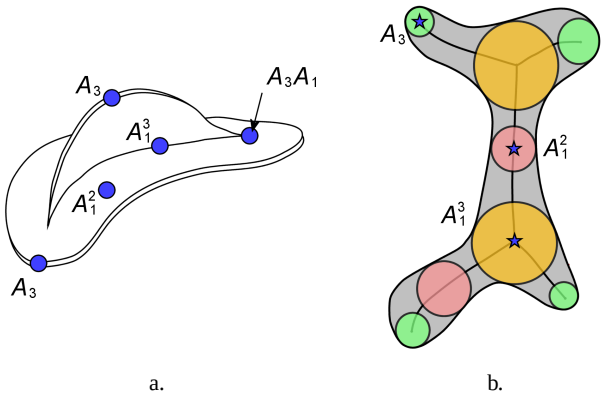
The medial locus can also be defined analytically using the following *grassfire analogy*. The object is imagined to be a patch of grass whose boundary is set on fire instantaneously. As the grass burns away, the fire fronts propagate from the boundary. This propagation can be described by the following differential equation:

$$\frac{\partial \mathscr{C}(t,p)}{\partial t} = -\alpha \mathbf{n}(p), \tag{1}$$

where $\mathscr{C}(t,p)$ denotes the fire front at time t, parameterized by p, $\mathbf{n}(p))$ is the unit outward normal to the fire front, and α is a constant, positive for inward propagation and negative for outward propagation. As the propagation progresses, segments of fire fronts that originate from disjoint parts of the boundary begin to meet and quench themselves at points that are called *shocks*. The medial locus is defined as the set of all the shocks, along with associated values of time t at which each shock is formed. This analytic definition of the medial locus is equivalent to the geometric definition given previously; a proof was given by Calabi (1965a) and Calabi (1965b); Calabi and Hartnett (1968). Kimia et al. (1995) combine the grassfire flow with an additional additive term based on the Euclidean curvature of the evolving front to yield a reaction-diffusion space for shape analysis.

2.2 Structural Geometry of Medial Loci

A rigorous description of the structural composition and local geometric properties of Blum medial loci of three-dimensional objects is given in Chapter 2 and Chapter



a. b.

Fig. 4 a. Different classes of points that compose the medial locus of a three-dimensional object, as categorized by Giblin and Kimia. b. Three possible ways in which maximal inscribed disks can be tangent to the boundary of a two-dimensional object.

3 (see also Giblin and Kimia (2000)). The classification of types of points on the medial locus was also given in (Yomdin, 1981) and in (Mather, 1983). Their description classifies medial points based on the multiplicity and order of contact that occurs between the boundary of an object and the maximal inscribed ball centered at a medial point.

Each medial point $P = \{\mathbf{p}, r\}$ in the object Ω is assigned a label of form A_k^m. The superscript m indicates the number of distinct points at which a ball of radius r centered at $\mathbf{p}$ has contact with the boundary $\partial\Omega$. The subscript k indicates the order of contact between the ball and the boundary. The order of tangent contact is a number that indicates how tightly a ball B is fitted to a surface S at a point of contact P.

The following theorem specifies all the possible types of contact that can generically occur between the boundary of a three-dimensional object and the maximal inscribed balls that form its medial locus. The theorem also specifies how medial points with different associated type of contact are organized to form surfaces and curves in the medial locus.

Theorem 0.1 (Giblin and Kimia). *The internal medial locus of a three-dimensional object Ω generically consists of*

1. *sheets (manifolds with boundary) of A_1^2 medial points;*
2. *curves of A_1^3 points, along which these sheets join, three at a time;*
3. *curves of A_3 points, which bound the free (unconnected) edges of the sheets and for which the corresponding boundary points fall on a crest;*
4. *points of type A_1^4, which occur when four A_1^3 curves meet;*
5. *points of type A_1A_3 (i.e., A_1 contact and A_3 contact at a distinct pair of points) which occur when an A_3 curve meets an A_1^3 curve.*

Proof. See (Giblin and Kimia, 2000) for a rigorous proof.

In two dimensions, a similar classification of medial points is possible. The internal medial locus of a two-dimensional object generically consists of (i) curves of bitangent A_1^2 points, (ii) points of type A_1^3 at which these curves meet, three at a time, and (iii) points of type A_3 which form the free ends of the curves. The three classes of contact are illustrated in Fig. 4a. In two dimensions, A_3 contact means that the inscribed disk and the boundary osculate at a local maximum of boundary curvature.

The geometric properties of the external medial locus are similar to those of the internal locus, with the exception that the sheets and curves are no longer completely bounded and may stretch out to infinity. Less effort has been devoted in the literature to the study of external medial loci.

Theorem 0.1 states that each surface composing the internal medial locus of an object joins another two such surfaces or terminates at a point of type A_3^1, which correspond to ridges of curvature on the boundary surface. Similarly, curve segments composing the internal medial loci of two-dimensional objects either connect with pairs of other curve segments or terminate at points corresponding to positive maxima of boundary curvature. Hence, the number of such ridges or maximal points limits the number of surfaces and curves in the internal medial locus. It can be shown by induction that the number of curve segments composing the internal medial locus of an object whose boundary has M positive maxima of curvature may not exceed $2M-3$.

We will use the term *stratum* to refer collectively to curves in medial loci of two-dimensional objects and to surfaces in medial loci of three dimensional objects. The composition of medial loci into interconnected strata makes it possible to decompose geometrically complex objects into simple components called *figures* (Fig. 5). Roughly speaking, a figure is the part of an object that corresponds to a particular stratum in the medial locus. A particularly simple mathematical definition of a figure is the following:

Definition 0.7. The union of closed balls whose centers and radii form a single stratum in the medial locus of an object is called a figure with joint[2].

Figures generated by strata belonging to the internal medial locus of an object are bounded, and the union of all such figures is the object itself. The intersection of a pair of figures with joints is non-empty if the generating strata of the two figures are connected. This non-empty intersection is called the *joint*, and it comprises of balls of triple boundary contact.

The internal medial locus of a figure has only a single stratum and figures can be said to be geometrically simple and easier to study than whole objects. Fig. 6 shows some examples of two-dimensional figures with joints.

The relationship between the structure of symmetry sets and the extrema of boundary curvature of two-dimensional objects are central to Leyton's theory of symmetry (Leyton, 1987). For planar objects, Leyton's curvature-symmetry duality theorem states that

[2] As distinguished from *figure*, which will be discussed later in the context of m-reps.

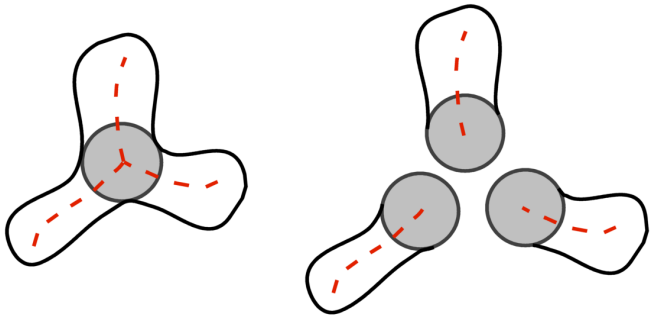


Fig. 5 Decomposition of a planar object into figures with joints. Each curve in the medial locus corresponds to a single figure.

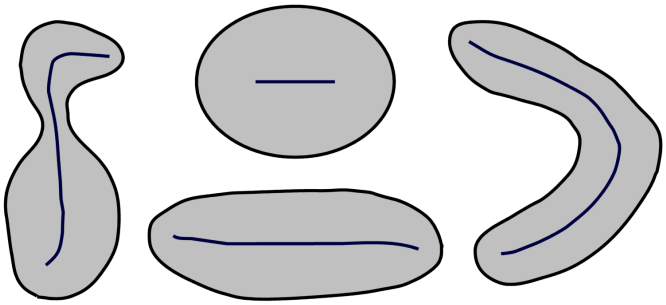


Fig. 6 All four of these objects fall into the category of *figures with joints* according to Def. 0.7, even though none of them have an actual "joint". Notice that the figure on the right has more than two positive maxima of curvature.

> Any section of curve, that has one and only one curvature extremum, has one and only one symmetry axis. This axis is forced to terminate at the extremum itself.

The extension of this theorem to three dimensions is given by Yuille and Leyton (1990).

Leyton's theory states that the curves composing the symmetry set of an object represent the history of events that have formed the object. According to Leyton's postulates, "memory is always in the form of asymmetry," meaning that asymmetry makes it possible to recover information about the formation of an object, while "symmetry is always the absence of memory." He suggests that the extrema of curvature are the places where the boundary has been pushed in from the outside or pushed out from the inside, indicating a growth or deformation process. The medial curves that terminate at these extrema are in a sense arrows in the direction of the

push. Hence, the symmetry set is a diagram of protrusion and indentation operations that have been applied to an object (Leyton, 1992).

2.3 Local Geometry of Medial Loci

Prior to describing the local geometry of Blum medial loci, let us introduce a useful notation for referring to the points of contact between a ball placed at a medial point and the boundary of an object.

Definition 0.8. If $P = \{\mathbf{p}, r\}$ is a medial point of an object Ω, then the set of points of contact between a ball of radius r centered at $\mathbf{p}$ and $\partial\Omega$ is called the *boundary pre-image* of P.

In other words, a medial point labeled A_k^m has a boundary pre-image that contains m points. Since for most of the medial points $m = 2$, the following definition is quite useful.

Definition 0.9. If points A and B form the boundary pre-image of a medial point P, then A is called a *medial involute* of B and vice versa.

Said in another way, medial involutes are pairs of points on the boundary of an object that are symmetric with respect to the medial locus. It is possible for a point to have multiple medial involutes, for example one with respect to the interior medial locus and one with respect to the exterior medial locus.

The major part of Blum's work on the internal medial loci of two-dimensional objects is devoted to the study of the geometric relationships between medial points and their boundary pre-images (Blum, 1967; Blum and Nagel, 1978). Blum showed that the points in the boundary pre-image can be expressed in terms of the position and radius of the medial point and from their derivatives with respect to movement along the medial locus.

For the purposes of studying local geometry of internal two-dimensional medial loci it suffices to focus on medial points that lie on the interior of the curves composing the medial locus and thus have two-point boundary pre-images. The geometric properties of the free endpoints and connecting endpoints of medial curves can be derived as the limit cases of the interior point properties.

In addition to using the position $\mathbf{p}$ and the radius r to characterize each point on the medial locus, Blum uses two first order properties. The first property is the slope of the medial curve at the medial point, which can be expressed as a unit-length tangent vector $\mathbf{U}^0$. The second is called the *object angle* and is given by

$$\theta = \arccos\left(-\frac{dr}{ds}\right), \tag{2}$$

where s is the arc length along the medial curve. The object angle is indicative of the narrowing rate of the object with respect to movement along the medial curve.

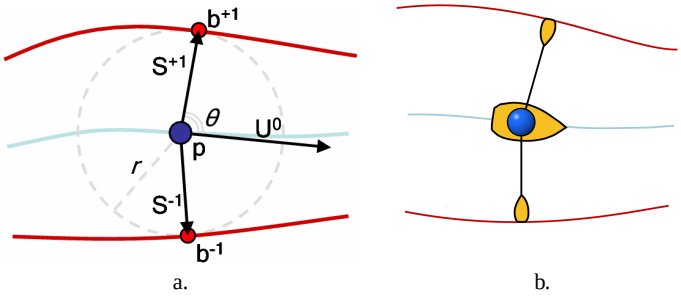


Fig. 7 Local medial geometry. a. Local geometric properties of a medial point and its boundary pre-image. b. The rowboat analogy for medial points: the oars (spokes) are $\mathbf{S}^{\pm 1} = r\mathbf{U}^{\pm 1}$.

When the object angle is equal to $\pi/2$, the radius has a critical point, and as one moves along the medial locus, the object retains its local width.

A medial point can be characterized by $\mathbf{p}, r, \mathbf{U}^0$, and θ, or preferably, by $\mathbf{p}, r, \mathbf{U}^{+1}, \mathbf{U}^{-1}$, where $\mathbf{U}^{+1}$ and $\mathbf{U}^{-1}$ are unit-length vectors orthogonal to the boundary of the object at $\mathbf{b}^{\pm 1} = \mathbf{p} + r\mathbf{U}^{\pm 1}$.

Fig. 7a describes the local geometry of a medial point and associated boundary pre-image points $\mathbf{b}^{-1}$ and $\mathbf{b}^{+1}$. The angle formed by the points $\mathbf{b}^{-1}$, $\mathbf{p}$, and $\mathbf{b}^{+1}$ is bisected by the vector $\mathbf{U}^0$, the unit tangent vector of the medial curve at $\mathbf{p}$. The angle between $\mathbf{U}^0$ and the vectors $\mathbf{b}^{-1} - \mathbf{p}$ and $\mathbf{b}^{+1} - \mathbf{p}$ is the object angle θ.

The quantities $\mathbf{p}, \mathbf{b}^{\pm 1}, r, \mathbf{U}^{\pm 1}, \mathbf{U}^0$, and θ appear frequently in this book. To better remember these quantities, consider an analogy between a medial point and a one-person rowboat, illustrated in Fig 7b. The position of the rower in the boat corresponds to $\mathbf{p}$, and the length of the oars corresponds to r. The vector $\mathbf{U}^0$ represents the direction in which the boat is moving and θ is the angle that each oar makes with $\mathbf{U}^0$. The points $\mathbf{b}^{-1}$ and $\mathbf{b}^{+1}$ correspond to the tips of the oars, and the directions of the oars are given by the vectors $\mathbf{U}^{+1}$ and $\mathbf{U}^{-1}$. The movement of a point along the medial locus is analogous to the rowboat navigating down the middle of a stream, with the rower adjusting his oars in such a way that their tips always just touch the banks of the stream (of course, the oars are made of a stretchable material, and as the boat moves, their length changes). A similar analogy to a wheel, corresponding to the bitangent disk, is made in m-rep literature, and the term *spoke* is used instead of the term *oar*. In this book we have adopted the term *spoke* for this vector between the medial locus and the boundary.

The values of $\mathbf{p}$, r, and their derivatives can be used to qualitatively describe the local bending and thickness of an object, as first shown by Blum and Nagel (1978). The measurements $\mathbf{p}$ and $\mathbf{U}^0$ along with the curvature of the medial curve describe the local shape of the medial locus, and subsequently describe how a figure bends at $\mathbf{p}$. A figure that has a line for its medial curve is symmetrical under reflection across that line. The measurement r describes how thick the figure is locally, while $\cos\theta$ describes how quickly the object is narrowing with respect to movement along the medial curve. A figure with a constant value of r has the shape of a worm.

Free ends of medial curves, where the maximal inscribed disk and the boundary osculate and the boundary pre-image contains a single point, are a limiting case of the bitangent disk situation. As our imaginary rowboat approaches such a point, its oars, i.e., the spokes, come closer and closer together until they collapse infinitely quickly at the end-point, forming a single vector in the direction $\mathbf{U}^0$. The object angle θ, which is equal to half of the angle between the spokes, is zero at such endpoints.

The geometry of medial loci of three-dimensional objects is harder to visualize and express than the planar medial geometry. A number of researchers have studied the differential geometry of three dimensional medial loci (Nackman, 1982; Vermeer, 1994; Gelston and Dutta, 1995; Hoffmann and Vermeer, 1996; Teixeira, 1998). However, the basic relations are not too different from the two-dimensional case. Instead of bitangent circles, we have bitangent spheres, with the spokes still forming normal vectors to the object boundary and their vector difference forming a normal $\mathbf{N}$ to the medial locus. Generically the medial locus forms two-dimensional manifolds and there are still two spokes at all but branch points and endpoints. The edges of the medial locus form space curves, as do the branches. Branch curves typically end; the normal situation is of a branching sheet attached to a parent sheet to form a fin.

In non-generic situations, sections of the sheet can degenerate to a space curve, in which case the object is a tube, i.e., has circular cross section and the medial atom becomes an entire cone of spokes with the $\mathbf{U}^0$ -vector as their axis. The relation between the $\mathbf{U}^0$ vector, the object angle θ and derivatives of r on the medial sheet becomes

$$\nabla r = -\mathbf{U}^0 cos(\theta),$$

where the gradient is intrinsic to the medial sheet.

The local reconstruction of the boundary from the medial axis or the symmetry set is discussed in Chapter 2 (Sections 6 and 7) following which the relationship between medial points of three-dimensional objects and the boundary pre-images of these points is presented in detail in Chapter 3. Many of the results reported in Chapter 3 are based on Damon's work on *skeletal structures* (Damon, 2003). This work follows the generative approach to medial geometry, as opposed to the previously described approaches that derive the medial locus from the boundary description of a given object. In the generative approach, the medial locus is defined first, and the object and its boundary are generated by an outward flow from it. Since the medial locus can be defined to contain a fixed number of figures with a specified topology, this approach to medial geometry is often better suited to problems of object modeling and shape description than the approaches deriving the medial axis from the object boundary. In the latter class of methods the derived medial loci can vary greatly and can thus be difficult to compare or analyze. The generative approach is the cornerstone of m-rep methodology, which will be introduced in the following section.

Damon's skeletal set is a stratified set[3] of arbitrary dimension, on which a multi-valued vector field **S**, called the *radial vector field*, is defined. At most points in the skeletal set a pair of radial vectors is defined; these vectors point in the different directions relative to the tangent space of the skeletal set. At edges of skeletal manifolds that are not shared by more than one manifold, the radial vectors come together to form a single vector that lies in the tangent space of the manifold. At shared edges, more than two radial vectors are defined. The endpoints of the radial vectors form a locus that is called the boundary of the object described by the skeletal set. Damon describes a number of constraints that must be satisfied by the skeletal set and the radial vector field in order for the boundary to be continuous and differentiable. These constraints are expressed in terms of the *radial shape operator* S_{rad} and a skeletal *edge shape operator* S_E, which measure how the radial vectors bend with respect to the skeletal set. These operators not only describe the local properties of the radial vector field but can also be used to express the local differential geometry of the boundary.

Another way to view the skeletal set for an object is that the skeletal locus is a fully cyclic piece of plastic wrap fit on both sides of the sheets forming the medial counterpart and that each point on the plastic wrap has a single spoke emanating from it, with none of the spokes crossing. Thus, at points that are neither branch points nor end points, there are two sheets of the plastic wrap touching each other and thus two spokes emanating, one from each sheet. These spokes need not be of equal length. At branch points there are three pieces of plastic wrap touching, and at end points the plastic wrap doubles back on itself and there is a single spoke, as illustrated in Fig. 8. Damon has shown that a natural measure on the skeletal sheet is the product of ordinary Riemannian distance on the sheet and the sine of the angle between the spoke vector and the tangent plane of the sheet. In Chapter 3 it is shown that this measure and a skeletal generalization of the shape operator can be used to pull back integrals over the object interior or its boundary onto the skeletal sheet. The Blum medial locus can be constructed as a special case of the general skeletal set by requiring that the radial vectors at each point of the skeletal set be symmetric with respect to its tangent space.

2.4 Medial Atoms and M-Reps

In this section we show how medial atoms can be used as building blocks for a particular type of object representation called m-reps. Via this scheme a collection of medial atoms can be used to come up with an approximate representation of a graph of figures with fixed topology. Such a view can provide advantages when the task is to draw comparisons across a population of similar structures such as that obtained by drawing several instances from a particular class of objects.

[3] As described in the glossary, a stratified set consists of interconnected *manifolds* of different codimensions.

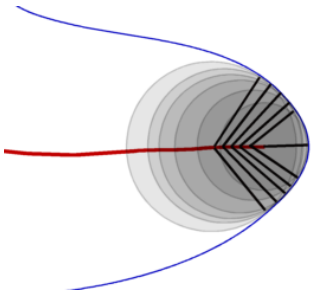


Fig. 8 Medial geometry of end atoms. The continuous relationship between a point on a medial curve and the points of contact between a disk inscribed at the point and the boundary of the object asymptotes at the end point of the curve: equal steps along the medial curve result in increasing steps along the boundary.

2.4.1 Medial Atoms

A medial atom is a modeling primitive that represents a place on the medial locus of an object. A medial atom describes such a place up to a specified differential order and with a specified level of tolerance. The continuous medial locus of an object can be discretely sampled into a set of medial atoms. However, medial atoms should be thought of as entities that exist independently of any medial locus, as demonstrated by the following definition.

Definition 0.10. An *n-dimensional medial atom of order 0* ($n = 2,3$) is a tuple $\mathbf{m} = \{\mathbf{p}, r\}$ that satisfies

$$\mathbf{p} \in \mathbb{R}^n,\ r \in \mathbb{R}^+. \tag{3}$$

Geometrically, a medial atom of order 0 is simply interpreted as a ball. Such medial atoms essentially correspond to maximal inscribed balls whose centers and radii form the medial loci of objects. Given a structured collection of medial atoms of order 0 sampled from the medial locus of an object, it is possible to approximately reconstruct the object's boundary by "shrink-wrapping" a sheet around the balls defined by the atoms, as shown in Fig. 9.

The use of the word "order" in the above definition refers to the fact that medial atoms can be used to approximate medial loci up to a given order. An order 0 atom describes zeroth order medial properties, which are position and radius. An order 1 atom, which is defined below, describes the derivatives of position and radius.

The shortcoming of medial atoms of order 0 and the shrink-wrap boundaries derived from them lies in the fact that while both the medial locus and the boundary can be approximately recovered from a set of order 0 medial atoms, the local symmetry relationships between pairs of medial involutes can not be directly expressed. However, medial atoms of higher order can be used to capture these symmetry relationships and are hence more useful for object representation.

Definition 0.11. An *n-dimensional medial atom of order 1* is a tuple $m = \{\mathbf{p}, r, \mathbf{U}^{+1}, \mathbf{U}^{-1}\}$ that satisfies

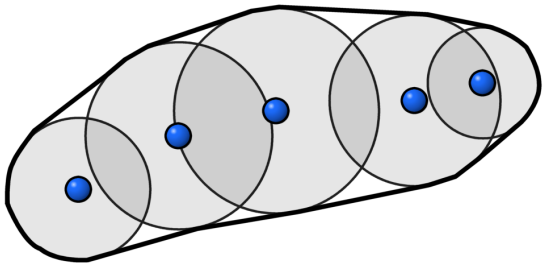


Fig. 9 Boundary of an object reconstructed by shrink-wrapping a collection of order 0 medial atoms.

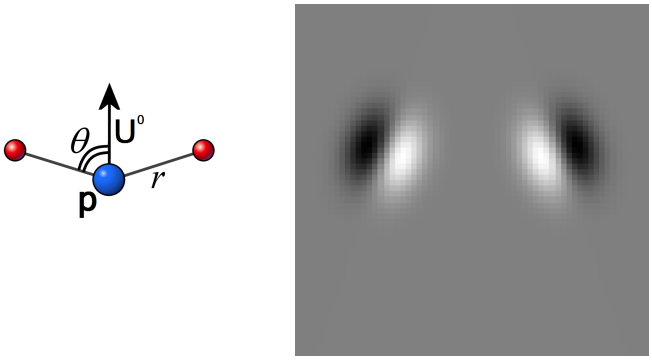
$$\mathbf{p} \in \mathbb{R}^n,\ r \in \mathbb{R}^+,\ \mathbf{U}^{+1} \in S^{n-1},\ \mathbf{U}^{-1} \in S^{n-1}, \tag{4}$$

where S^n is the unit n-dimensional sphere.

The additional components of the medial atom of order 1 are the two spoke orientations, $\mathbf{U}^{+1}, \mathbf{U}^{-1}$. Recall that these are precisely the same first order quantities that were used, after multiplication by the spoke length r to compute $\mathbf{S}^{\pm 1}$, to relate medial points to their boundary pre-images (Fig. 7).

The rowboat analogy used to describe the first order geometry at a point on the medial locus can be used equally well to describe medial atoms of order 1. The atom's position and radius correspond to the position of the rower and the length of the oars (spokes); the orientation corresponds to the direction of the boat's bow, and the object angle to the angle between the bow and the oars (spokes). Unlike their order 0 cousins, which one could visualize as an inflated rubber tube, medial atoms of order 1 explicitly define a pairs of points on the boundary of the object that they describe. These points are, of course, the tips of the spokes and are given by the following definition.

Definition 0.12. The tuples $\{\mathbf{b}^{-1}, \mathbf{U}^{-1}\}$ and $\{\mathbf{b}^{+1}, \mathbf{U}^{+1}\}$ given by $\mathbf{b}^{\pm 1} = \mathbf{p} + r\mathbf{U}^{\pm 1}$ are called the *implied boundary nodes* of the medial atom $m = \{\mathbf{p}, r, \mathbf{U}^{+1}\mathbf{U}^{-1}\}$. Here $\mathbf{U}^{\pm 1} = (cos(\theta), \pm\sin(\theta))^T$ in 2D and $(cos(\theta), \pm\sin(\theta), 0)^T$ in 3D in the coordinate system whose first basis vector is a unit bisector $\mathbf{U}^0$ of $\mathbf{U}^{+1}$ and $\mathbf{U}^{-1}$ and whose second basis vector is a unit vector in the direction of $\mathbf{U}^{+1} - \mathbf{U}^{-1}$, where θ is the object angle, i.e., half the angle between $\mathbf{U}^{+1}$ and $\mathbf{U}^{-1}$.



a. b.

Fig. 10 Medial atoms used for core tracking. a) A medial atom defined by position **p**, radius r, orientation determined by $\mathbf{U}^0$ and object angle θ. b) A Gaussian derivative filter associated with the atom, as used for core tracking.

Medial atoms of order 1 correspond to medial points of type A_1^2, whose boundary pre-images contain two points.[4]. Medial points of other types can be represented either using special medial atoms, as discussed later in Sec. 2.4.2, or as limit cases of A_1^2 medial points. In this section we deal only with order 1 medial atoms because these atoms contain all the information necessary to reconstruct the boundary pre-images of medial points and the symmetry relationships between pairs of medial involutes.

The agreement between a medial atom and an image is measured by a *medialness function*, the domain of which is the set of parameters defining a medial atom. Given a tuple of parameter values, the medialness function measures how well the position and orientation of the boundary nodes of a medial atom defined by these parameters match edge-like structures in the image.

As illustrated in Fig. 10, the medialness function most widely used in the core tracking literature is computed using the image intensity gradient at boundary node locations. Given an image I and an order 1 medial atom $\underline{\mathbf{m}} = \{\mathbf{p}, r, \mathbf{U}^{+1}, \mathbf{U}^{-1}\}$, this function is defined as

$$M(\underline{\mathbf{m}}) = \nabla_\sigma I(\mathbf{b}^{-1}) \cdot \mathbf{U}^{-1} + \nabla_\sigma I(\mathbf{b}^{+1}) \cdot \mathbf{U}^{+1} \ , \tag{5}$$

where $\nabla_\sigma I(\mathbf{b})$ is the image gradient computed at the point **b** by convolution with the gradient of the isotropic Gaussian kernel with aperture σ :

$$\nabla_\sigma I(\mathbf{b}) = \int \nabla G_\sigma(\mathbf{b} - \mathbf{z}) I(\mathbf{z}) \, \mathrm{d}\mathbf{z} \ . \tag{6}$$

[4] The A_i^j taxonomy of medial points, as used by Giblin and Kimia (2000), was discussed in Section 2.1.

The aperture σ is proportional to the radius of the medial atom. This proportionality makes the medialness function invariant to the magnification of the structures in the image.

2.4.2 M-Reps: A Medial Object Representation

First-order medial atoms serve as the building blocks for m-reps, which have the following distinguishing properties:

- The medial locus of an object is represented explicitly.
- A smooth approximate representation of the object's boundary, with tolerance, is implied by the medial locus representation.
- An accurate description of the boundary is given by a smooth fine-scale deformation of the implied boundary.

The explicit specification of an object's medial locus by m-reps makes it possible to compare similar objects in terms of symmetries. In contrast, medial loci yielded by applying a skeletonization method to similar objects can pose challenges for comparison because their branching topology may differ, as in the case of nearly circular objects.

M-reps come in different flavors. Discrete m-reps, due to Pizer et al. (1999), represent the medial locus using a structured sparse set of medial atoms (Fig. 11 shows a 2D example, and Fig. 12 shows a 3D example). Continuous m-reps, defined in Yushkevich et al. (2003), represent the medial locus as approximation curves or surfaces defined by a set of $\{\mathbf{p}, r\}$ control points. This section describes discrete m-reps.

M-reps specify the figural composition of an object explicitly. An m-rep representing a complex object contains multiple components which we shall refer to as *figures*. Each figure is a sheet of medial atoms. These figures are organized into a hierarchy of parent-child relationships, with parents representing the substantial parts of the object, such as the palm of the hand, and children representing protrusions and indentations, such as the fingers. The figural graph of an m-rep resembles the composition of a geometric object into figures with joints but does so only at a conceptual level. The manner in which an m-rep is organized into figures is guided by structural, conceptual, and populational properties, rather than by a desire to precisely mimic the medial branching topology of the objects being represented. The difference between the figural composition of m-reps and the branching topology of corresponding objects is illustrated in Fig. 11.

Various means for representing m-reps in a computer are discussed in Chapter 8.

In direct contrast to skeletonization methods, m-reps derive the boundary description of an object from its medial description. Since the composition of an m-rep into figures and atoms is explicitly imposed, one can describe different objects using m-reps with the same figural composition. This makes it possible to then compare such objects based on their medial properties. In particular, one only needs to

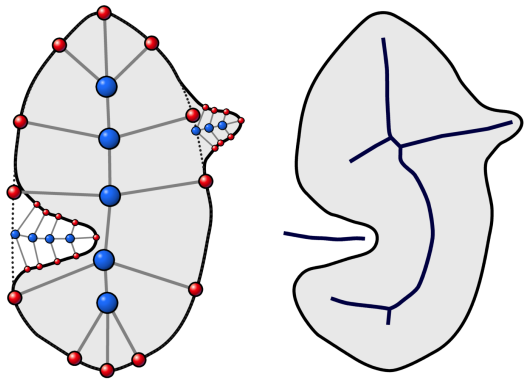


Fig. 11 (Left). Representation of an object using a discrete m-rep. The m-rep is organized into a hierarchy of figures based on structural properties of the object. At the root of the hierarchy lies the main figure whose implied boundary is indicated by the dotted curve. The children of the main figure are the protrusion and indentation figures. (Right). Continuous medial locus of the same object. Branches in the medial locus are determined by the geometry of the boundary and include branches that do not contribute to the structural description of the object.

compare the values of the medial atom parameters, assuming that medial atoms represent corresponding locations in the objects. This type of analysis of course makes

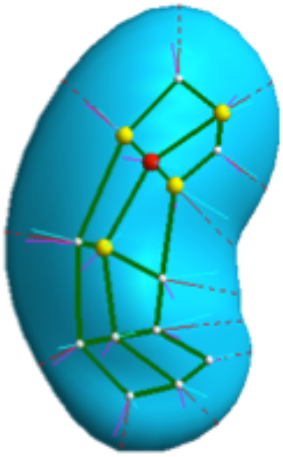


Fig. 12 A three-dimensional discrete m-rep figure of a kidney organized as a 3×5 quadrilateral mesh of medial atoms. The atoms with a larger light hub are the neighbors of the atom with a darker hub. The atom with a darker hub and the two atoms below it are interior medial atoms; the rest are end atoms.

an implicit assumption that the instances being compared are well described by the imposed m-rep fit. The alternative is to attempt to use the more standard skeletonization approaches to obtain a representation, but then one has to compare medial structures with different branching topologies. There is a growing literature on the use of graph theoretic methods for matching skeletal structures; some of this material is covered in Chapter 10. However, these techniques are more complex than the comparison of m-reps. For example, the use of m-reps with a common figural composition makes it possible to estimate probability distributions on the parameters of medial atoms and hence makes medial based shape characterization across a population of similar objects possible.

3 Psychophysical and Neurophysiological Evidence for Medial Loci

The medial representations described in Section 2 have not only found many applications by engineers but have also been found to be relevant components of human vision models. For example, there is literature which suggests that figural decomposition of objects using medial loci might correspond to the cognitive processing performed by the human brain. Rock and Linnett (1993) argue that figures are captured preattentively by the human visual system. Other studies have shown that figures which arise from locations of extrema of negative curvature on the boundary are often associated with the visual decomposition of objects (Hoffman and Richards, 1984; Biederman, 1987; Braunstein et al., 1989; Siddiqi et al., 1996; Singh and Hoffman, 1997). The junctions of figures can also have special importance, matching Biederman's work on the junctions of visual parts (Biederman, 1987). Whereas we shall not cover this literature in great detail in this book, in this section we review some of the findings that point to a role for medial loci in shape perception. We shall also cover the literature that provides neurophysiological evidence for medial axes.

Among the first reported psychophysical data is that of Frome (1972), who examined the role of medial axes in predicting human performance in shape alignment tasks. In these experiments subjects were required to position an ellipse so as to be parallel to a reference line. It was found that for this task, the acuity with which the stimulus was placed could be explained by the length of the medial axis within the ellipse, i.e., the straight line connecting the centers of curvatures corresponding to the end points of the major axis of the ellipse.

Later, Psotka (1978) examined the role of medial loci in representing the perceived form of more complex outlines. Subjects were given outline forms (a square, a circle, a humanoid form and various rectangles) and were asked to draw a dot within each outline in the first place that came to mind. The superimposed dots for each outline were found to coincide well, for the most part, with the locations predicted by Blum's grassfire flow as opposed to locations suggested by other field theories of form perception that had been proposed, such as McCulloch's size constancy proposal (McCulloch, 1965).

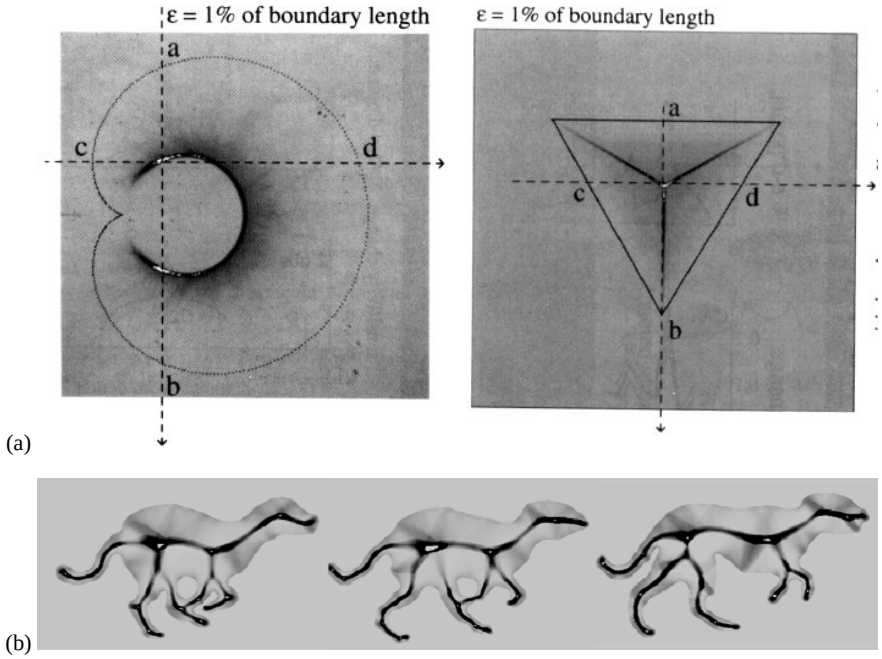


Fig. 13 (Adapted from (Kovács et al., 1998, Fig. 2, 8, 9 & 10), with permission from Ilona Kovács.) The function D_ε for various objects. Dark shading corresponds to increasing values of D_ε, and the "white spot" denotes its maximum. (a) D_ε for a cardioid on the left and a triangle on the right; cross-sections through maximum loci are indicated as dotted lines a-b and c-d. (b) D_ε for a few frames in a sequence depicting the movement of an animal.

Adopting a different experimental procedure, the effect of closure on figure-ground segmentation along with a possible role for medial loci was examined by Kovács and Julesz (1993, 1994) and Kovács et al. (1998). In the experiments reported in Kovács and Julesz (1993) subjects viewed a display of Gabor patches (GPs) aligned along a sampled curve presented in a background of randomly oriented Gabor patches playing the role of distractors. Using a two alternative forced choice paradigm, subjects were required to decide whether the display included an open curve or a closed one. The percentage of correct responses were recorded as a function of the separation distance between successive GPs. A significant advantage was found for the correct detection of configurations of closed (roughly circular) targets. In a second series of experiments subjects were required to detect a target GP of varying contrast placed either inside or outside a circular arrangement of GPs in a field of distractors. It was found that the contrast threshold at which the target could be detected was decreased by a factor of 2 when it was located at the center of the circle, as opposed to the periphery, suggesting a special role for a medial location.

This second finding of increased contrast sensitivity at the center was later examined more carefully in Kovács and Julesz (1994). For an elliptical configuration

of GPs it was found that the peak locations of increased contrast sensitivity were in fact predicted by a type of medial model. Specifically, these locations coincided with the local maxima of a D_ε distance function, representing at each location the percentage of boundary locations over the entire outline that were equidistant within a tolerance of ε. Examples of the D_ε function for more complex forms were presented in Kovács et al. (1998) (see Fig. 13) along with the proposal that its maxima could play an important role in form perception tasks such as the processing of motion (see also Chapter 11, Section 11).

Whereas the above notion of an ε parameter is fixed as a global quantity, a different notion of scale provides the key motivation for medially subserved perceptual models based on *cores* (Burbeck and Pizer, 1995), which are discussed in Section 4.4. Underlying this model is the hypothesis that the scale at which the visual system integrates local boundary information is proportional to object width at medial loci. This view is corroborated by the findings of Burbeck et al. (1996), where the core model was shown to explain human performance in bisection tasks on shapes. In these experiments elongated stimuli were created by placing two sinusoidal waves in phase, side by side, and then filling in the region in between. An example of such a "wiggle" stimulus is shown in Fig. 14. The amount that it is perceived to bend depends on the frequency and amplitude of the sinusoids. For any given stimulus a subject was asked to judge whether a probe dot, placed between two sinusoidal peaks, appeared to the left or right of the object's center. By varying the position of the probe dot the perceived center was chosen to be the point about which a subject was equally likely to choose left or right in the task. The perceived central modulation was then defined as the horizontal distance between the centers for a successive left and right peak. The experiments revealed that for a fixed width the central modulation increased with increasing amplitude but decreased with increasing frequency. Furthermore, the modulation effects were greater for a narrow object, in a manner that was adequately explained by the linking of object boundaries at a scale determined by object width, as predicted by the core model.

These wiggle stimuli were revisited by Siddiqi et al. (2001), who showed that in fact the perceived centers in the study of Burbeck et al. (1996) were located precisely on the Blum medial axis, at locations that coincided with local maxima of the radius function. In several experiments using similar stimuli, but with varying degrees of translation between the sinusoidal boundaries, properties of medial loci were shown to account for human performance in shape discrimination tasks using a visual search paradigm.

Taken together, the above body of work provides a wealth of support for the role of medial loci in shape perception. Unfortunately, far less research has been carried out to provide neurophysiological support for medial axes. The one exception is the work of Lee (1995) and Lee et al. (1998) where neurons were isolated in the primary visual cortices of awake rhesus monkeys and their response to a set of texture images was examined. In the first study (Lee, 1995) the input images consisted of either a linear boundary with two regions of contrasting texture, or a rectangular strip or a square on a background of contrasting texture. In each case the texture was comprised of scattered bars in a vertical or horizontal orientation. The findings

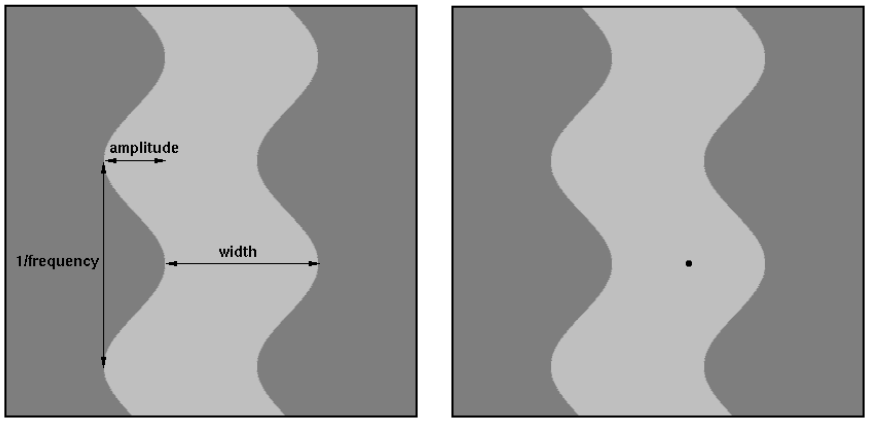


Fig. 14 (Adapted from Siddiqi et al. (2001)). LEFT: The geometry of a "wiggle" stimulus used in (Burbeck et al., 1996). RIGHT: The task is for the subject to judge whether the probe dot is to the left or to the right of the object's center.

revealed a subset of neurons that had a peak response when their receptive fields were centered at the texture boundaries. Some of these neurons also had a sharp response when centered at the center of the rectangular strip or square, i.e., at locations predicted by the Blum medial axis. The subsequent more comprehensive results reported in Lee et al. (1998) revealed that the neurons with interior response peaks appeared to focus in the vicinity of the center of mass of compact shapes (squares and diamonds) but along the entire medial axis for elongated shapes (rectangles). This finding is consistent with the special status attributed to a local maximum of the radius function, e.g., in the context of the wiggle bisection experiments and the contrast sensitivity enhancement experiments discussed earlier. However, the neurophysiological data in support of medial axes has not yet been corroborated in the literature by other researchers.

4 Extracting Medial Loci of Objects

The human vision models described in Section 3 suggest that medial information is somehow extracted from the contrast available at boundaries or from the boundaries themselves. How do we get a computer to extract medial loci?

The computer vision literature describes a large number of *skeletonization* methods, which extract medial loci of objects starting from some boundary representation. In most practical applications the object and its boundary are represented discretely, for example as a set of pixels of the same intensity in a characteristic image or as a mesh of points. Skeletonization is made difficult by the inherent sensitivity of the medial locus to the fine details of the boundary representation. Given

two different discrete representations of the same object, the true Blum medial loci of the two representations can have a different *medial branching topology*, i.e., a different number of figures and a different connectivity graph between the figures.

Hence, the challenge of skeletonization is not to find the precise medial locus of an imprecisely specified boundary, but rather it is to compute an approximation of the medial locus that is consistent with respect to different discrete representations of the same object. Moreover, a good skeletonization method should yield similarly structured medial loci for objects that are similar objects and for versions of the same object that have undergone similarity transformations or magnification.

This section introduces the notions of distance transforms, the Hessian, thinning and pruning, which are common ingredients to many skeletonization algorithms. It then overviews the various approaches to extracting medial loci in the literature. The approach of shocks of the grassfire flow, interpreted as singularities of the Euclidean distance function, is developed in more detail in Chapter 4. Methods based on digital distance transforms are described in Chapter 5, and those based on Voronoi techniques are detailed in Chapters 6 and 7. The boundary evolution and Voronoi methods are also compared to approaches based on height ridges of medialness (cores) in the overview paper by Pizer et al. (2003b). A special property of these three approaches is that they provide a scale parameter that makes it possible to tune the accuracy with which the result matches the precise medial locus of the input boundary. The loci computed at larger values of the scale parameter generalize better to different discrete representations of the input object as well as to other similar objects.

4.1 Distance Transforms, the Hessian, Thinning and Pruning

A *distance transform* of a boundary is obtained by assigning to each location in space its distance to the closest point on the boundary. Thus, points which lie on the boundary are assigned a value of zero. The distance transforms in use in the skeletonization literature typically adopt a notion of Euclidean or approximate Euclidean distance. It is also common to distinguish locations which lie in the interior of a boundary from those that lie in the exterior by a sign change, leading to the notion of a signed distance transform. Distance transforms turn out to be very useful representations since their level curves or surfaces represent the locus of positions reached by successive iterations of a grassfire flow. Furthermore, their singularities coincide with the Blum medial axis. Thus, numerous approaches to skeletonization use a distance transform to simulate the grassfire flow along with techniques to locate its singularities. These approaches benefit from the fact the distance transform values at each point on the medial locus give the radius function.

The *Hessian* matrix of a smooth approximation to a distance function provides eigenvalues that can be used to characterize the type of local symmetry of its level sets: in 3D the slab, the tube, and the sphere. While the tube is non-generic and the sphere doubly non-generic, the Hessian eigenvalues give graded results such as

"a slab but almost tubular." With $\lambda_1, \lambda_2, \lambda_3$ the eigenvalues of the Hessian and with $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$, we have

- 0-D ridges (nearly spherical) when $\lambda_1 \approx \lambda_2 \approx \lambda_3 \gg 0$,
- 1-D ridges (nearly tubular) when $\lambda_1 \approx 0, \lambda_2 \approx \lambda_3 \gg 0$ and
- 2-D ridges (very much a slab) when $\lambda_1 \approx \lambda_2 \approx 0, \lambda_3 \gg 0$.

The same analysis can also be applied to pseudo-distance functions, such as the smooth edge strength *v* in Section 4.3.

There exist a number of techniques in the literature to compute the medial axis by successively peeling layers of the distance transform, e.g., by using morphological *thinning*. Binary mathematical morphology (Serra, 1982; Matheron, 1988; Jonker and Vossepoel, 1995) is based on the operations of erosion and dilation of an object via a structuring element. The general idea is to carry out a process of erosion by a ball to the binary image iteratively, thinning the object until it is one pixel thick. These methods take advantage of the fact that successive erosion by this small structuring element is equivalent to erosion by a larger disk or ball and that the discrete approximation to the larger disk improves when a small element is successively applied. As the successive erosions by the small structuring element are applied, the methods check for pixel or voxel removals that change the topology of the skeleton from that of the original object. These voxels are marked as skeleton voxels. This class of algorithms actually has a long history in the pattern analysis literature. The advantage of such methods is that they can be computationally very efficient. However, the challenge faced by such methods derives from the fact that the resolution of both the original object and the skeleton is given by an underlying rectangular lattice: a voxel is either on the skeleton, or it is not. The resulting medial loci can be sensitive to the rasterization of the object and may have difficulty in discerning the local geometry of the medial locus near branch points, particularly in 3D. Moreover, the resulting medial loci can be sensitive to the rotation and magnification of the object prior to the imaging process.

As discussed earlier, an inevitable difficulty faced by skeletonization algorithms is their inherent sensitivity to small perturbations of the boundary of an object. Thus many methods result in an initial coarse computation of the skeleton, followed by a second stage of *pruning* in which components which are thought to represent insignificant boundary details are removed. The types of measures used for pruning are driven in part by the manner in which the original boundary is represented and the computational techniques that are used. For example, when the boundary is defined by a polygon or a mesh and Voronoi techniques are used to compute the medial locus, pruning functions are designed to measure the area or length contribution of a Voronoi edge or face to determine whether or not it should be kept (Chapters 6 and 7). On the other hand, when computations are carried out on a discrete lattice using digital distance transforms, pruning measures are designed to take into account invariance and/or stability properties of the reconstruction with or without a branch or sheet, as in Chapter 5. In both settings it is generally accepted that a pruning method should have the following properties:

1. It should preserve topology (homotopy type).

2. It should be continuous, i.e., small differences in the significance measure should result in small changes to the computed skeleton.
3. The significance measure should be local on the medial locus.

4.2 Skeletons via Shocks of Boundary Evolution

One class of approaches aims to simulate Blum's grassfire flow, as described by Eq. 1, and to then detect the locus of quench points. These approaches are distinct to the approaches of morphological erosion, in that the evolving curve is modeled as a partial differential equation. In practice this raises questions about how the flow should be numerically discretized and the singularities of the evolution detected, both of which are nontrivial issues.

Representative methods in this class are the techniques of Leymarie and Levine (1992); Tari et al. (1997); Siddiqi et al. (2002). In the first method the boundary of a 2D object is represented as an active contour, which is partitioned at locations of positive curvature maxima corresponding to the medial axis endpoints. The various segments of the active contour then propagate inwards driven by a potential function modeled by the negative gradient of the Euclidean distance function to the boundary. The active contours slow down in the vicinity of the medial axis, where the magnitude of the numerical gradient is small.

The skeletonization method in Siddiqi et al. (1999, 2002); Dimitrov et al. (2003) is based on a novel characterization of singularities of the grassfire flow, which in turn lends itself to robust and efficient numerical implementations. The key insight is that in the limit as the area (2D) or volume (3D) within which average outward flux is computed shrinks to zero, the average outward flux of the gradient of the Euclidean distance function has different limiting behaviors at non-medial and medial points. This allows for a uniform treatment in 2D and 3D, along with associated skeletonization algorithms. These techniques, along with related methods based on the gradient of the Euclidean distance function, are discussed in Chapter 4. It turns out that for the case of shrinking circular neighborhoods the limiting values of the average outward flux reveal the object angle and hence allow for the explicit recovery of their boundary pre-image (Dimitrov et al., 2003; Dimitrov, 2003). An advantage of this approach is that the analysis extends to higher dimensions.

4.3 Greyscale Skeletons

It is well known that the grassfire flow is equivalent to a formulation where the time of arrival surface associated with the level curves of the flow satisfy an eikonal equation. A generalization of these results to the case of greyscale images is described in (Tari et al., 1997; Shah, 1996). The method is based on a linear differential equation that is developed from a model introduced by Modica and Mortola (1977) for

approximating the characteristic function $\chi_{\partial\Omega}$ of the boundary $\partial\Omega$ of an object Ω. Given a boundary approximation $\chi_{\partial\Omega}$, possibly computed by thresholding a simple edge strength operator and thus not necessarily closed, it generates a scalar function v on space (image) giving modified edge strength, whose troughs yield a medial locus.

The modification is designed to produce an edge strength image that is smooth, i.e., has small gradient magnitude, and falls from a value 1 at $\partial\Omega$ towards 0 with distance from $\partial\Omega$. As such it minimizes the functional

$$E_\sigma(v) = \int_\Omega \left[\sigma \|\nabla v\|^2 + \frac{(v-\chi_{\partial\Omega})^2}{\sigma} \right] = \int_\Omega \left[\sigma \|\nabla v\|^2 + \frac{v^2}{\sigma} \right] \tag{7}$$

with the constraint $v = 1$ along $\partial\Omega$. The minimizer of E_σ satisfies the elliptic differential equation

$$\nabla^2 v = \frac{v}{\sigma^2} \tag{8}$$

with the constraint serving as a boundary condition. The parameter σ plays the role of a nominal smoothing radius. When σ is small compared to the local width of the shape and the local radius of curvature of $\partial\Omega$, the level curves of v capture the smoothing of $\partial\Omega$ by a curve evolution model with a combination of a constant motion (grassfire) and curvature motion term, as used in Kimia et al. (1995). In particular, as shown by Mumford and Shah (1989), when σ is small,

$$v(x,y) = \sigma \left(1 + \frac{\sigma\kappa(x,y)}{2} \right) \frac{\partial v}{\partial \eta}(x,y) + O(\sigma^3) \tag{9}$$

where η is the direction of the gradient of v and $\kappa(x,y) = v_{\xi\xi} / \|\nabla v\|$ is the curvature of the level curve of v passing through the point (x,y) with $v_{\xi\xi}$ the second derivative of v in the direction ξ tangent to the level curve.

The skeleton is taken to be the troughs of v. These troughs are computed by finding the extrema of $|\nabla v|$ along the level curves of v. As indicated in equation 9, near the shape boundary the level sets of v mimic a curve (or surface) evolution process with a speed consisting of a constant component and a component proportional to curvature.

In mathematical terms the above extrema are the set of positions where

$$\frac{d\|\nabla v\|}{ds} = 0,$$

with s the arc-length along level curves of v. These are the points where the level curves are parallel and the gradient lines of v have inflection points. Geometrically this means that the gradient vector ∇v is an eigenvector of the Hessian of v at points on the axes of local symmetry. This method is developed in a series of interesting papers (Tari et al., 1997; Tari and Shah, 1998).

The above ideas have been applied to both the computation of medial loci of boundaries implicitly defined in greyscale images, and their segmentation into pro-

trusions and indentations in (Shah, 2001, 2005). The latter is accomplished by grouping points which share the same sign of the second derivative $\frac{d^2\|\nabla v\|}{ds^2}$. Furthermore, extensions of these ideas to higher dimensions have been explored (Tari and Shah, 2000; Shah, 2005b). The development of robust algorithms to compute these medial loci remains an area of investigation, and these computational methods are similar in spirit to the class of methods based on core tracking to be discussed next. A practical difficulty faced here is that the medial loci so obtained are disconnected. Nonetheless, it is possible to use the direction of ∇v to determine a hierarchical interpretation as a graph of components. The 2D case is detailed in Shah (2001).

4.4 Core Tracking

The greyscale skeletons of Shah and Tari can be seen as a special case of cores, with the function v, or more precisely its reciprocal, serving as a measure of medialness. Cores generalize the medialness to functions not just of position but also medial radius and spoke directions. *Cores* are medialness height ridges, i.e., maxima of dimension $p-k$ in a space of dimension p equal to that of medial atoms, i.e., 5 in 2D and 8 in 3D.

These medial loci are computed by tracking (Pizer et al., 1998; Furst and Pizer, 1998; Morse et al., 1998; Eberly et al., 1994; Fritsch et al., 1995a). These particular medial loci do not have the same branching properties as Blum medial loci, so special branch-searching and seeding strategies are needed (Fridman, 2004). Their strength is that they are derived directly from greyscale image data without the need of first extracting boundaries and that, as shown in the references just cited, they are quite insensitive to image noise.

Core tracking literature has made two different choices for defining the subspaces in which maxima are found, resulting in two classes of cores: *maximal convexity cores* and *optimal parameter cores*. In maximal convexity cores the subspaces are defined by the $p-1$ directions of greatest second derivative of M, computed at each point as the unit eigenvectors of the Hessian matrix (Eberly, 1996). In optimal parameter cores, the atoms are required to attain a local maximum in radius, orientation, and object angle, as well as in the direction orthogonal to the Euclidean tangent space of the core, which is defined by the optimal orientation (Furst and Pizer, 1998; Fridman et al., 2004).

Core tracking methods work by following a core from a starting point in the parameter space. A user specifies a location, size, orientation and object angle of an initial medial atom and the algorithm searches for a ridge point in the vicinity of this initialization. The algorithm then tracks the core, taking small steps in the parameter space until some termination condition is met. Core tracking has been implemented using marching cubes methodology (Furst and Pizer, 1998; Lorensen and Cline, 1987) and using predictor-corrector methods (Fritsch et al., 1995b).

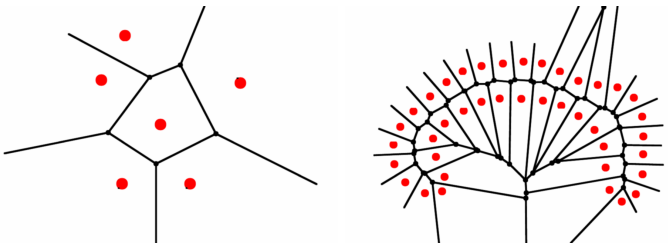


Fig. 15 Examples of Voronoi diagrams. a. Voronoi Diagram of six points. b. Voronoi Diagram of points sampled from the boundary of the corpus callosum. The skeleton of the object is just the internal portion of the Voronoi Diagram.

4.5 Skeletons from Digital Distance Transforms

The pattern analysis literature is replete with a class of methods which use digital distance transforms but are distinct from the methods of morphological erosion. These methods attempt to locate height ridges on a discrete lattice, using techniques from discrete geometry and topology. This is done either by iterative removal of pixels or voxels, or by marking centrally located elements, or a combination of these two ideas. The state of the art algorithms in this class are the subject of Chapter 5. These methods offer the advantage that they are computationally efficient and that formal guarantees on the quality of the results can be provided, including reversibility and homotopy preservation.

4.6 Voronoi Skeletons

Voronoi skeletons (Ogniewicz, 1993; Székely, 1996; Amenta et al., 2001) are computed by calculating the *Voronoi diagram* of a set of points sampled from the boundary of an object. Fig. 15a shows a Voronoi diagram of a set of six points on a plane. The diagram consists of *Voronoi regions*, which are sets of points located closer to a particular generating point than to any other generating point. The line segments in the diagram are called *Voronoi edges*; they separate Voronoi regions and are loci of points that are equidistant from a a pair of generating points. The points where Voronoi edges meet are equidistant from three or more generating points.

When the generating points of a Voronoi diagram are sampled from the boundary of an object, as shown in Fig. 15b, the similarity between Voronoi edges and the curves composing the medial locus becomes apparent. A circle of appropriate radius centered at a point on a Voronoi edge contains two generating points, i.e., has two points of contact with the boundary. A circle centered at an intersection of two Voronoi edges contains three generating points, i.e., has three points of contact with the boundary, as do disks centered at intersections of curves in the medial locus.

The Voronoi diagram is also related to the grassfire analogy: if some points on the boundary are set on fire (as opposed to the whole boundary), the places where fire fronts meet and quench themselves are the edges in the Voronoi diagram of these points.

The Voronoi diagram of a set of boundary points contains edges that extend outside of the object, possibly to infinity. The discrete approximation of the boundary obtained by connecting the generating boundary points with line segments cuts the Voronoi diagram into internal and externals parts. The internal part is called the *Voronoi skeleton*. The Voronoi skeleton generated by a discrete representation of an object's boundary is an approximation of that object's internal medial locus. Schmitt provides a proof that as the number of generating boundary points increases, the Voronoi skeleton converges in the limit to the continuous medial locus, with the exception of the edges generated by neighboring pairs of boundary points (Schmitt, 1989).

The Voronoi skeletons, such as the one shown in Fig. 15b, contain many branches, some of which are spurious and sensitive to the slightest changes to the generating boundary points. For instance, a Voronoi skeleton computed from the set of pixels forming the boundary of an object in an image can change significantly if the object in the image is rotated. In order to make Voronoi skeletons more robust to small boundary changes, researchers have proposed to isolate parts of the Voronoi skeletons that are most stable and significant. A number of measures of significance for edges and groups of edges in the Voronoi skeleton have been introduced in the literature (Ogniewicz and Kübler, 1995; Székely, 1996). The significance measures make it possible to establish trunk-branch relationships between connected edges in the Voronoi skeleton, and thus to establish a hierarchy of figures and sub-figures. The edges that fall far from the root of this hierarchy and have small significance values do not contribute to the descriptive ability of the Voronoi skeleton and are trimmed. Pruning on the basis of significance introduces a component of scale into Voronoi skeletonization. By adjusting the threshold level of significance at which edges are discarded from the skeleton, it is possible to generate a coarse-to-fine spectrum of skeletons. These ideas are explored in some detail in Chapter 6. Chapter 7 discusses theoretical results on the quality of approximation of the medial locus obtained by Voronoi diagram methods in 3D.

Both local and global measurements of significance have been proposed to organize and prune Voronoi skeletons. Local measurements assign a significance score to each edge in the Voronoi skeleton using a heuristic, such as the distance along the boundary between the pair of generating points to which the edge is equidistant (Ogniewicz, 1993). Global methods, on the other hand, compute the significance of an edge or a group of connected edges by measuring its impact on the appearance of the whole object, for example measuring the effect that removing the edge or edges from the skeleton would have on the shape of the boundary (Näf, 1996; Styner, 2001; Katz and Pizer, 2003).

The construction of Voronoi skeletons of three dimensional objects, while analogous to the two-dimensional construction, is much more difficult to implement. One difficulty arises during the traversal of Voronoi edges. The connectivity of Voronoi

edges in two dimensions organizes them into a tree structure that can be traversed from the trunk to the leaves; in three dimensions the Voronoi edges may form a graph that contains cycles and is more difficult to traverse. Methods for organizing three-dimensional Voronoi skeletons into figures have been developed by Näf (1996); Attali et al. (1997); Styner (2001).

4.7 Skeletonization by Deformable Medial Models

The skeletonization methods described so far are *deterministic* in the sense that the medial locus is defined entirely by the boundary of the object. These methods do not provide a way of incorporating prior knowledge about the shape of the medial locus into the extraction process. This limitation leads to the methods frequently producing spurious branches. This limitation also makes it difficult to use skeletonization methods to produce features for shape characterization. A solution is to use deformable methods to fit a medial model of a fixed branching topology to the object.

A discrete medial model can be constructed from a single medial manifold in one of two ways, by sampling or by parametrization. Both of these ways are described in Chapter 8, but here we restrict our discussion to an m-rep as a quadrangular grid of medial atoms sampling the continuous medial locus of atoms. The corresponding continuous locus can be interpolated from the grid, and each atom in this interpolated medial locus has an intrinsic medial coordinate described earlier. The end atoms, i.e., boundary atoms in such a grid represent the crest where the atoms on one side of the continuous sheet meet the atoms on the other side of that sheet. Branching protrusion figures or indentation figures from the figure defined by this main grid can also be represented by such a quadrangular grid of medial atoms, with certain of its end atoms, called *hinge atoms* designated as forming the base of the subfigure and sitting on the medially implied boundary of the host figure. The connections between the subfigure and its host figure are given by the medial coordinates in the host of the subfigure's hinge atoms.

Deformation of such a medially defined object consists of transforming each of its medial atoms in such a way that the implied object remains unfolded and each atom avoids unnecessarily changing its relation to its neighboring atoms in the grid and also taking care that each hinge atom connecting a subfigure to its host figure can transform in ways consistent with sliding, or rotating the subfigure on its host figure's medially implied surface. All of this, including the coordinate system provided by medial atoms that allows the expressions of the relationships between medial atoms, is described in Chapter 8.

Given a characteristic image or a greyscale image, the medialness of each of the interpolated medial atoms from such a deformable set of grids of atoms can be measured, and the deformation which optimizes the integral of this medialness can be chosen as the medial representation of the object. However, this representation is limited to objects with the same branching topology as that of the starting model.

Chapter 8 shows how such models can be formed from training populations of the object. It also shows how probabilities on the deformations of these models can be derived from the training populations and how these probabilities can be used as priors when calculating optimal deformations of the model into a new target. These ideas put deformable m-reps into the realm of Grenander's *pattern theory* (Mumford, 1996; Grenander, 1976, 1978, 1981).

5 Applications of medial loci in computer vision

Due to their effectiveness in representing object shape, medial loci have found many applications in computer vision and image analysis. Since they can be used to represent object interiors via local rotational and magnification transformations as well as displacements, and since they provide an object-centered correspondence of positions between object instances, they provide an effective basis for probabilistic descriptions of populations of objects. In turn, these descriptions can be used for testing hypotheses as to shape differences between object classes, classification of new objects, statistical sampling, and other applications of probability theory, as discussed in Chapter 9.

Another useful application of medial models is the segmentation of objects and multi-object complexes via deformable models. Deformable model segmentation is equivalent to image registration, so the very same approaches can be used to register two images of the same objects. Such segmentations/registrations can be done by optimization of a posterior probability over a shape space of relatively low dimension, making use of the aforementioned probabilistic descriptions as prior probabilities and as the means of limiting the shape space dimension. This application is detailed in Chapter 9.

Segmentation via following a medial locus from image data, without needing an explicit prior probability, has also met with success. While many of the methods apply to both tubular and slab-shaped objects, they have been particularly successful in the case of trees of tubes whose branching structure is variable and thus not specifiable as a prior on a model with fixed branching topology. Aylward and Bullitt (2002) and Fridman et al. (2004) have produced core-based methods for this purpose. Lorigo et al. (2001) and Descoteaux et al. (2004) have produced tubular tree extraction methods leveraging the fact that skeletal curves are entities of codimension greater than 1 in three dimensions. The method of Descoteaux et al. is flux-based and follows the more general form described in (Vasilevskiy and Siddiqi, 2002). These applications are not detailed further in this book.

Because they represent objects as a graph of figures, with geometric descriptions of both the figures and their connections, medial representations are very well suited to object recognition and labeling of object parts. Several graph theoretic methods have been developed for such purposes and Chapter 10 covers this application with a focus on the problem of 3D model retrieval. These methods can also be applied to the analysis and comparison of anatomical structures as viewed in medical im-

ages, including vessel trees, bronchial trees (Tschirren et al., 2005), etc. Centerlines, which may be viewed as the limiting case of a medial manifold shrinking to a 3D curve, have also proven useful for the visualization of structures in virtual endoscopy (Deschamps and Cohen, 2001; Bouix et al., 2004). These latter applications are not covered in this book.

Yet another application that has been developed for medial axes leverages the object-relative coordinates that they provide for applying physical simulations to objects and collections of objects. For example, one can do mechanical simulations based on medial models (Crouch et al., 2004), i.e., with spatial meshings based on medial coordinates. These partial differential equations can be applied to medial models directly in the object-based coordinates that the models provide, or even on the curvilinear high-dimensional spaces that describe medially represented objects (see Chapter 8). However, since there are only a few such applications thus far in the literature, they are not covered in this book.

We conclude the book with Chapter 11, which provides an overview of applications of medial representations in different fields at scales ranging from the very large to the very small.

References

Amenta, N., S. Choi, and R. Kolluri: 2001, 'The Power Crust, Unions of Balls, and the Medial Axis Transform'. *Computational Geometry: Theory and Applications* **19**(2-3), 127–153. [31, 364]

Asada, H. and M. Brady: 1983, 'The Curvature Primal Sketch'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8**, 2–14. [5]

Attali, D., G. Sanniti di Baja, and E. Thiel: 1997, 'Skeleton simplification through nonsignificant branch removal'. *Image Processing and Communications* **3**(3-4), 63–73. [33]

August, J., K. Siddiqi, and S. W. Zucker: 1999, 'Ligature Instabilities in the Perceptual Organization of Shape'. *Computer Vision and Image Understanding* **76**(3), 231–243. [6]

Aylward, S. R. and E. Bullitt: 2002, 'Initialization, Noise, Singularities, and Scale in Height Ridge Traversal for Tubular Object Centerline Extraction'. *IEEE Transactions on Medical Imaging* **21**(2), 61–75. [34]

Betelu, S., G. Sapiro, A. Tannenbaum, and P. Giblin: 2000, 'Noise Resistant Affine Skeletons of Planar Curves'. In: *Proceedings of the European Conference on Computer Vision*. Dublin, Ireland, pp. 742–754. [6]

Biederman, I.: 1987, 'Recognition–by–components: A theory of human image understanding'. *Psychological Review* **94**(2), 115–147. [22]

Blanding, R., C. Brooking, M. Ganter, and D. Storti: 1999, 'A Skeletal-Based Solid Editor'. In: W. F. Bronsvoort and D. C. Anderson (eds.): *Proceedings of the Fifth Symposium on Solid Modeling and Applications (SSMA-99)*. New York, pp. 141–150, ACM Press. [7]

Bloomenthal, J. and K. Shoemake: 1991, 'Convolution surfaces'. *Computer Graphics (SIGGRAPH '91 Proceedings)* **25**(4), 251–256. [7]

Blum, H.: 1967, *A transformation for extracting new descriptors of shape*. MIT Press. [1, 4, 7, 13]

Blum, H. and R. Nagel: 1978, 'Shape Description Using Weighted Symmetric Axis Features'. *Pattern Recognition* **10**(3), 167–180. [7, 13, 14, 99]

Bookstein, F. L.: 1991, *Morphometric Tools for Landmark Data*. Cambridge University Press. [2, 320]

Bouix, S., K. Siddiqi, and A. R. Tannenbaum: 2004, 'Flux Driven Automatic Centerline Extraction'. *Medical Image Analysis*. [35]

Braunstein, M. L., D. D. Hoffman, and A. Saidpour: 1989, 'Parts of Visual Objects: An Experimental Test of the Minima Rule'. *Perception* **18**, 817–826. [22]

Bruce, J. W., P. J. Giblin, and C. G. Gibson: 1985, 'Symmetry Sets'. *Proceedings of the Royal Society of Edinburgh* **101(A)**, 163–186. [5]

Burbeck, C. A. and S. M. Pizer: 1995, 'Object Representation by Cores: Identifying and Representing Primitive Spatial Regions'. *Vision Research* **35**(13), 1917–1930. [24]

Burbeck, C. A., S. M. Pizer, B. S. Morse, D. Ariely, G. S. Zauberman, and J. Rolland: 1996, 'Linking Object Boundaries at Scale: A Common Mechanism for Size and Shape Judgements'. *Vision Research* **36**(3), 361–372. [24, 25]

Calabi, L.: 1965a, 'On the shape of plane figures'. Technical Report PML No 60429 SR-1, Parke Mathematical Laboratories. [9]

Calabi, L.: 1965b, 'A study of the skeleton of plane figures'. Technical Report PML No 60429 SR-2, Parke Mathematical Laboratories. [9]

Calabi, L. and W. Hartnett: 1968, 'Shape recognition, prairie fires, convex deficiencies and skeletons.'. *American Mathematical Monthly* **75**, 335–342. [9]

Crouch, J., S. M. Pizer, E. L. Chaney, and M. Zaider: 2004, 'Medial Techniques To Automate Finite Element Analysis of Prostate Deformation'. *IEEE Transactions on Medical Imaging* **to appear**. [35]

Damon, J.: 2003, 'Smoothness and Geometry of Boundaries Associated to Skeletal Structures I: Sufficient Conditions for Smoothness'. *Annales Inst. Fourier* **53**, 1941–1985. [7, 15]

Delingette, H.: 1999, 'General Object Reconstruction Based On Simplex Meshes'. *International Journal of Computer Vision* **32**(2), 111–146. [2, 305]

Deschamps, T. and L. D. Cohen: 2001, 'Fast Extraction of Minimal Paths in 3D Images and Applications to Virtual Endoscopy'. *Medical Image Analysis* **5**(4), 281–299. [35]

Descoteaux, M., L. Collins, and K. Siddiqi: 2004, 'Geometric Flows for Segmenting Vasculature in MRI: Theory and Validation'. In: *International Conference On Medical Image Computing and Computer Assisted Intervention*, Vol. LNCS 3217(1). pp. 500–507. [34]

Dimitrov, P.: 2003, 'Flux Invariants for Shape'. Master's thesis, McGill University, Montréal, QC, Canada. [28]

Dimitrov, P., J. N. Damon, and K. Siddiqi: 2003, 'Flux Invariants for Shape'. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Madison, WI. [28, 127, 153, 154, 162, 169, 353]

Dryden, I. L. and K. Mardia: 1998, *Statistical Shape Analysis*. Chichester: Wiley. [2]

Eberly, D.: 1996, *Ridges in Image and Data Analysis*, Computational Imaging and Vision Series. Dordrecht, NL: Kluwer Academic Publishers. [30, 63, 106, 107]

Eberly, D., R. Gardner, B. Morse, S. Pizer, and C. Scharlach: 1994, 'Ridges for Image Analysis'. *Journal of Mathematical Imaging and Vision* **4**, 351–371. [30]

Fridman, Y.: 2004, 'Modeling Tubuar Structures Using Cores'. Ph.D. thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC. [30]

Fridman, Y., S. M. Pizer, S. Aylward, and E. Bullitt: 2004, 'Extracting Branching Tubular Object Geometry via Cores'. *Medical Image Analysis* **to appear**. [30, 34]

Fritsch, D., E. Chaney, A. Boxwala, M. McAuliffe, S. Raghavan, A. Thall, and E. J.R.D.: 1995a, '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. [30]

Fritsch, D., D. Eberly, S. Pizer, and M. McAuliffe: 1995b, 'Stimulated Cores and their Applications in Medical Imaging'. In: *Information Processing in Medical Imaging 1995 (IPMI'95)*. pp. 365–368. [30]

Frome, F. S.: 1972, 'A Psychophysical Study of Shape Alignment'. Technical Report TR-198, University of Maryland, Computer Science Center. [22]

Furst, J. D. and S. M. Pizer: 1998, 'Marching Optimal Parameter Ridges: An Algorithm to Extract Shape Loci in 3D Images'. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 780–787. [30]

Gelston, S. M. and D. Dutta: 1995, 'Boundary surface recovery from skeleton curves and surfaces'. *Computer Aided Geometric Design* **12**(1), 27–51. ISSN 0167-8396. [15]

Giblin, P. J. and S. A. Brassett: 1985, 'Local symmetry of plane curves'. *Amer. Math. Monthly* **92**, 689–707. [5, 6, 8]

Giblin, P. J. and B. B. Kimia: 2000, 'A formal Classification of 3D Medial Axis Points and Their Local Geometry'. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 1. pp. 566–573. [10, 19]

Giblin, P. J. and B. B. Kimia: 2002, 'Transitions of the 3D Medial Axis under a One-Parameter Family of Deformations'. In: *Proceedings of the European Conference on Computer Vision*, Vol. 2351 of *Lecture Notes in Computer Science*. pp. 718–734, Springer. [7]

Giblin, P. J. and B. B. Kimia: 2004, 'A formal Classification of 3D Medial Axis Points and Their Local Geometry'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(2), 238–251. [7, 352, 354]

Grenander, U.: 1976, *Pattern Synthesis: Lectures in Pattern Theory, vol. I*. New York: Springer-Verlag. [34]

Grenander, U.: 1978, *Pattern Synthesis: Lectures in Pattern Theory, vol. II*. New York: Springer-Verlag. [34]

Grenander, U.: 1981, *Regular Structures: Lectures in Pattern Theory, vol. III*. New York: Springer-Verlag. [34]

Grenander, U. and M. I. Miller: 1998, 'Computational Anatomy: An Emerging Discipline'. *Quarterly of Applied Mathematics* **56**, 617–694. [3]

Hoffman, D. D. and W. A. Richards: 1984, 'Parts of recognition'. *Cognition* **18**, 65–96. [22]

Hoffmann, C. M. and P. J. Vermeer: 1996, 'Validity Determination for MAT Surface Representation'. In: G. Mullineux (ed.): *Proceedings of the 6th IMA Conference on the Mathematics of Surfaces (IMA-94)*, Vol. VI of *Mathematics of Surfaces*. Oxford, pp. 249–266, Clarendon Press. [15]

Igarashi, T., S. Matsuoka, and H. Tanaka: 1999, 'Teddy: A Sketching Interface for 3D Freeform Design'. *Proceedings of SIGGRAPH 99* pp. 409–416. ISBN 0-20148-560-5. Held in Los Angeles, California. [7]

Jonker, P. and A. Vossepoel: 1995, 'On skeletonization algorithms for 2, 3 .. N dimensional images'. In: D. Dori and A. Bruckstein (eds.): *Shape, Structure and Pattern Recognition*. Singapore, pp. 71–80, World Scientific. [27]

Joshi, S. and M. I. Miller: 2000, 'Landmark Matching Via Large Deformation Diffeomorphisms'. *IEEE Transactions On Image Processiing* **9**, 1357–1370. [2]

Katz, R.: 2002, 'Form Metrics for Interactive Rendering via Figural Models of Perception'. Ph.D. thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC. [6]

Katz, R. and S. M. Pizer: 2003, 'Untangling the Blum Medial Axis Transform'. *International Journal of Computer Vision* **55**(3), 139–153. [32]

Kelemen, A., G. Székely, and G. Gerig: 1999, 'Elastic Model-Based Segmentation of 3D Neuroradiological Data Sets'. *IEEE Transactions on Medical Imaging* **18**(10), 828–839. [3]

Kendall, D. G.: 1989, 'A Survey of the Statistical Theory of Shape'. *Statistical Science* **4**, 87–120. [2]

Kimia, B. B., A. Tannenbaum, and S. W. Zucker: 1995, 'Shape, Shocks, and Deformations I: The Components of Two-Dimensional Shape and the Reaction-Diffusion Space'. *International Journal of Computer Vision* **15**, 189–224. [5, 9, 29]

Kovács, I., A. Feher, and B. Julesz: 1998, 'Medial-point description of shape: A representation for action coding and its psychophysical correlates'. *Vision Research* **38**, 2323–2333. [23, 24]

Kovács, I. and B. Julesz: 1993, 'A Closed Curve is Much More Than an Incomplete One: Effect of Closure in Figure-Ground Segmentation'. *Proceedings of the National Academy of Science of the USA* **90**(16), 7495–7497. [23]

Kovács, I. and B. Julesz: 1994, 'Perceptual Sensitivity Maps within Globally Defined Visual Shapes'. *Nature* **370**, 644–646. [23]

Lee, T. S.: 1995, 'Neurophysiological Evidence for Image Segmentation and Medial Axis Computation in Primate V1'. In: J. Bower (ed.): *Fourth Annual Computational Neuroscience Meeting*. pp. 373–378, Academic Press. [24]

Lee, T. S., D. Mumford, R. Romero, and V. A. F. Lamme: 1998, 'The Role of the Primary Visual Cortex in Higher Level Vision'. *Vision Research* **38**, 2429–2454. [24, 25]

Leymarie, F. and M. D. Levine: 1992, 'Simulating the Grassfire Transform Using an Active Contour Model'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(1), 56–75. [28, 144, 145]

Leyton, M.: 1987, 'Symmetry-Curvature Duality'. *Computer Vision, Graphics, and Image Processing* **37**(3), 327–341. [6, 11]

Leyton, M.: 1988, 'A Process Grammar For Shape'. *Artificial Intelligence* **34**, 213–247. [5]

Leyton, M.: 1989, 'Inferring Causal History From Shape'. *Cognitive Science* **13**, 357–387. [5]

Leyton, M.: 1992, *Symmetry, Causality, Mind*. MIT press. [5, 13]

Lorensen, W. E. and H. E. Cline: 1987, 'Marching cubes: A high resolution 3D surface construction algorithm'. *Computer Graphics (SIGGRAPH '87 Proceedings)* **21**, 163–169. [30]

Lorigo, L. M., O. D. Faugeras, E. L. Grimson, R. Keriven, R. Kikinis, A. Nabavi, and C.-F. Westin: 2001, 'CURVES: Curve evolution for vessel segmentation'. *Medical Image Analysis* **5**, 195–206. [34]

Mather, J. N.: 1983, 'Distance From a Submanifold in Euclidean Space'. *Proceedings of Symposia in Pure Mathematics* **40**(2), 199–216. [1, 5, 10, 85]

Matheron, G.: 1988, 'Examples of Topological Properties of Skeletons'. In: J. Serra (ed.): *Image analysis and Mathematical Morphology*, Vol. 2. Academic Press, pp. 217–238. [27]

McCulloch, W. S.: 1965, *Embodiments of Mind*. MIT Press. [22]

Millman, D.: 1980, 'The Central Function of the boundary of a domain and its differential properties'. *Journal of Geometry* **14**, 182–202. [1]

Modica, L. and S. Mortola: 1977, 'Il limite nella Γ-convergence di una familiglia di funzionali ellittici'. *Boll. Un. Mat. Ital.* **14-A**, 526–529. [28]

Morse, B., S. Pizer, D. Puff, and C. Gu: 1998, 'Zoom-Invariant Vision of Figural Shape: Effects on Cores of Image Disturbances'. *Computer Vision and Image Understanding* **69**, 72–86. [30]

Mumford, D.: 1996, 'Pattern theory: A unifying perspective'. In: D. C. Knill and W. Richards (eds.): *Perception as Bayesian Inference*. Cambridge Universiry Press, pp. 25–62. [34]

Mumford, D. and J. Shah: 1989, 'Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems'. *Comm. Pure Appl. Math.* **42**(5), 577–684. [29]

Nackman, L. R.: 1982, 'Three-Dimensional Shape Description Using the Symetric Axis Transform'. Ph. d. thesis, University of North Carolina at Chapel Hill. Department of Computer Science. [15, 99]

Näf, M.: 1996, 'Voronoi Skeletons: a semicontinous implementation of the 'Symmetric Axis Transform' in 3D space'. Ph.D. thesis, ETH Zürich , Commmunication Technology Institue, Image Analysis Group IKT/BIWI. [32, 33]

Ogniewicz, R. L.: 1993, *Discrete Voronoi Skeletons*. Hartung-Gorre. [31, 32]

Ogniewicz, R. L. and O. Kübler: 1995, 'Hierarchic Voronoi skeletons'. *Pattern Recognition* **28**(3), 343–359. [32]

Pizer, S., D. Eberly, B. Morse, and D. Fritsch: 1998, 'Zoom-Invariant Figural Shape: The Mathematics of Cores'. *Computer Vision and Image Understanding (CVIU '98)* **69**, 55–71. [30, 106]

Pizer, S. M., D. S. Fritsch, P. Yushkevich, V. Johnson, and E. Chaney: 1999, 'Segmentation, Registration and Measurement of Shape Variation via Image Object Shape'. *IEEE Transactions on Medical Imaging* **18**, 851–865. [20]

Pizer, S. M., K. Siddiqi, G. Székeley, J. N. Damon, and S. W. Zucker: 2003, 'Multiscale Medial Axes and Their Properties'. *International Journal of Computer Vision* **55**(3), 155–179. [26, 127]

Psotka, J.: 1978, 'Perceptual Processes That May Create Stick Figures and BalancE'. *Journal of Experimental Psychology: Human Perception and Peformance* **4**(1), 101–111. [22]

Rock, I. and C. Linnett: 1993, 'Is a perceived shape based on its retinal image?'. *Perception* **22**(1), 61–76. [22]

Schmitt, M.: 1989, 'Some examples of algorithms analysis in computational geometry by means of mathematical morphology techniques'. In: J. Boissonnat and J. Laumond (eds.): *Geometry and Robotics*, Vol. LNCS 391. Springer-Verlag, pp. 225–246. [32]

Serra, J.: 1982, *Image Analysis and Mathematical Morphology*. Academic Press. [27]

Shah, J.: 1996, 'A Common Framework for Curve Evolution, Segmentation and Anisotropic Diffusion'. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 136–142. [28]

Shah, J.: 2001, 'Segmentation of Shapes'. In: *Workshop on Scale-Space and Morphology*. [30]

Shah, J.: 2005a, 'Grayscale Skeletons and Segmentation of Shapes'. *Computer Vision and Image Understanding* **99**(1), 96–109. [30, 164, 165, 166, 167]

Shah, J.: 2005b, 'Skeletons of 3D Shapes'. In: *The Fifth International Conference on Scale-Space and PDE Methods in Computer Vision*. [30]

Sherstyuk, A.: 1999, 'Shape design using convolution surfaces'. In: *Proceedings of Shape Modeling International '99*. [7]

Siddiqi, K., S. Bouix, A. Tannenbaum, and S. W. Zucker: 1999, 'The Hamilton-Jacobi Skeleton'. In: *Proceedings of the IEEE International Conference on Computer Vision*. Kerkyra, Greece, pp. 828–834. [28, 154, 370]

Siddiqi, K., S. Bouix, A. Tannenbaum, and S. W. Zucker: 2002, 'Hamilton-Jacobi Skeletons'. *International Journal of Computer Vision* **48**(3), 215–231. [28, 84, 89, 153, 158, 351, 352, 353]

Siddiqi, K., B. B. Kimia, A. R. Tannenbaum, and S. W. Zucker: 2001, 'On the Psychophysics of the Shape Triangle'. *Vision Research* **41**, 1153–1178. [24, 25]

Siddiqi, K., B. B. Kimia, and K. J. Tresness: 1996, 'Parts of Visual Form: Psychophysical Aspects'. *Perception* **25**(4), 399–424. [22]

Singh, M. and D. D. Hoffman: 1997, 'Salience of Visual Parts'. *Cognition* **63**, 29–78. [22]

Storti, D. W., G. M. Turkiyyah, M. A. Ganter, C. T. Lim, and D. M. Stat: 1997, 'Skeleton-based modeling operations on solids'. In: *SMA '97: Proceedings of the Fourth Symposium on Solid Modeling and Applications*. pp. 141–154. held May 14-16, 1997 in Atlanta, Georgia. [7]

Styner, M.: 2001, 'Combined Boundary-Medial Shape Description of Variable Biological Objects'. Ph.D. thesis, University of North Carolina at Chapel Hill, Chapel Hill, NC. [32, 33]

Székely, G.: 1996, 'Shape Characterization by Local Symmetries'. Habilitationsschrift, Institut fur Kommunikationstechnik, Fachgruppe Bildwissenschaft, ETH Zürich. [31, 32]

Tari, S. and J. Shah: 1998, 'Local Symmetries of Shapes in Arbitrary Dimension'. In: *Proceedings of the IEEE International Conference on Computer Vision*. Bombay, India. [29]

Tari, S. and J. Shah: 2000, 'Nested Local Symmetry Set'. *Computer Vision and Image Understanding* **79**(2), 267–280. [30]

Tari, Z. S. G., J. Shah, and H. Pien: 1997, 'Extraction of Shape Skeletons from Grayscale Images'. *Computer Vision and Image Understanding* **66**, 133–146. [28, 29]

Teixeira, R. C.: 1998, 'Curvature Motions, Medial Axes and Distance Transforms'. Ph.D. thesis, Harvard University, Cambridge, Massachusetts. [15]

Tschirren, J., G. McLennan, K. Palyagyi, E. Hoffman, and M. Sonka: 2005, 'Matching and Anatomic Labeling of Human Airway Tree'. *IEEE Transactions on Medical Imaging* **24**(12), 1540–1547. [35]

Vasilevskiy, A. and K. Siddiqi: 2002, 'Flux Maximizing Geometric Flows'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(12). [34]

Vermeer, P. J.: 1994, 'Medial Axis Transform to Boundary Representation Conversion'. Ph. d. thesis, Purdue University. [15]

Yomdin, Y.: 1981, 'On the Local Structure of a Generic Central Set'. *Compositio Mathematica* **43**(2), 225–238. [1, 10]

Yuille, A. and M. Leyton: 1990, '3d Symmetry-Curvature Duality Theorems'. *CVGIP* **52**, 124–140. [6, 12]

Yushkevich, P., P. T. Fletcher, S. Joshi, A. Thall, and S. M. Pizer: 2003, 'Continuous Medial Representations for Geometric Object Modeling in 2D and 3D'. *Image and Vision Computing* **21**(1), 17–28. [20, 292]

Zhu, S. C.: 1999, 'Stochastic Jump-Diffusion Process for Computing Medial Axes in Markov Random Fields'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(11). [5]

Synthesis, Deformation, and Statistics of 3D Objects via M-reps

Stephen Pizer and Qiong Han and Sarang Joshi and P. Thomas Fletcher and Paul A. Yushkevich and Andrew Thall

Abstract The m-rep, a representation of the interior of one or more objects, from which boundaries can be synthesized, is described in detail. An m-rep consists of sheets of medial atoms; both sampled and parametrized representations of these sheets are described. Means of forming objects made from a main sheet (*figure*) and attached protrusion or indentation subfigures are described, as are multiscale hierarchies of object complexes, objects, figures, atoms, and voxels. The object-relative coordinate system provided by m-reps is presented. To allow the estimation of probabilities on populations of m-reps, the m-rep can be understood as an element in a feature space that takes the mathematical form of a symmetric space. Doing this provides the ability to estimate probabilities by a generalization of principal component analysis to these curved spaces.

Stephen Pizer
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: pizer@cs.unc.edu

Qiong Han
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: han@cs.unc.edu

Sarang Joshi
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: sjoshi@cs.unc.edu

P. Thomas Fletcher
Department of Computer Science, University of Utah, USA
e-mail: fletcher@sci.utah.edu

Paul A. Yushkevich
Department of Radiology, University of Pennsylvania, USA
e-mail: pauly2@grasp.upenn.edu

Andrew Thall
Department of Computer Science, Allegheny College, USA
e-mail: athall@allegheny.edu

1 Introduction

Chapters 4-7 have taken the point of view that objects begin from a boundary representation and a medial representation is derived from the boundary. An alternative view is that an object is a member of a population of instances of the object and that a fixed topology of the medial locus can be derived from this population, as well as a probability distribution on the geometry of that medial locus. Thus every medial instance can be seen as a deformation of the mean of this probability distribution, and every boundary can be seen as synthesized from that medial instance. The medial representation called the *m-rep*, discussed in Chapter 1, Section 2.4, enables the formation of these fixed topology medial structures and the estimation of their probabilities from training samples. This chapter covers the details of m-reps, the view of an m-rep as a point on the mathematical entity called a *symmetric space*, and the ideas of probability distributions on symmetric spaces. Moreover, m-reps are described in two forms: grids of order 2 medial atoms called *discrete m-reps* and splines of order 1 medial atoms called *continuous mreps*.

2 M-reps, Medial Atoms, and Figures

We seek a means for representing an object, such as a liver or a car, and ensembles of objects, such as the whole abdomen or a street scene. Following the philosophy of Grenander's *Pattern Theory* (Grenander (1996)) that an object's very shape is described by its deformations into its various instances in a population and also realizing that objects deform mechanically in time, we need a representation in which the associated deformations are rich, natural, and efficiently implemented.

Let us focus on single objects first. What constitutes an object? The intuitive view is that an object is not simply a shell but rather consists of interior material that can be locally transformed by elongation, bending, twisting, swelling or contraction, and displacement. We follow this useful view and choose to synthesize objects from such a description of the object interior.

The medial representation, originally promulgated as a locus of bitangent spheres, is well suited to providing this description because its sphere primitive is locally maximally interior to the object. However, it fails to integrate to the interior. This aim is achieved by a small modification (Fig. 1) - replacing the sphere by a hub formed by its center and the two equal-length spokes to the points of sphere tangency. As explained in Chapter 1, we call this object-interior-component primitive a *medial atom*, and we use the term *m-rep* to refer to a locus of medial atoms that sweep out an object interior.

The locus of medial atoms can be a manifold with boundary; an object or object part represented by such a manifold is called a *figure*. As illustrated in Fig. 2, in 2D the manifold can be a curve, in which case we call the figure a *bar*; in a limiting case the curve degenerates to a point, and the bar degenerates to a disk. In 3D the manifold can be two-dimensional, in which case we call the figure a *slab* and its

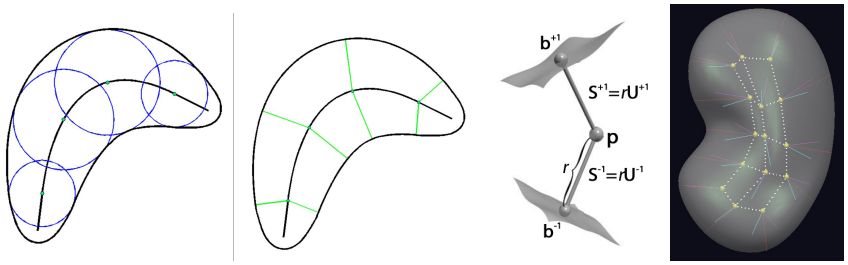


Fig. 1 a) Sample tangent spheres representing an object in 2D. b) Medial atoms representing the same object. c) A medial atom in 3D. The hub is at position $\mathbf{p}$; the two spokes, of length r, are named $\mathbf{S}^{+1}$ and $\mathbf{S}^{-1}$, respectively, the unit vectors in those direction are $\mathbf{U}^{+1}$ and $\mathbf{U}^{-1}$, respectively, and the spoke ends are at positions $\mathbf{b}^{+1}$ and $\mathbf{b}^{-1}$, respectively; finally, fractional distance from the hub to the spoke end is given by τ. d) A 3D kidney with a grid of samples of its medial atoms. This representation is called a *sampled m-rep* or a *discrete m-rep*. The third spoke on the atoms at the edge of the grid was discussed in Chapter 1 and is detailed in Section 8 of this chapter.

locus of hubs the *medial sheet*. In 3D the manifold can also be one-dimensional, in which case we call the locus of the hubs the *medial axis*, the atom must be interpreted as representing all spokes obtained by spoke rotation about the atom spokes' bisector, and the figure is called a *tube*. In the limiting 3D case the axis degenerates to a point, and the figure degenerates to a sphere.

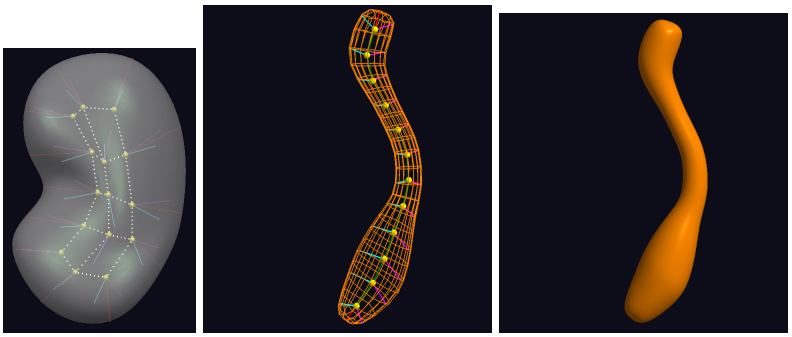
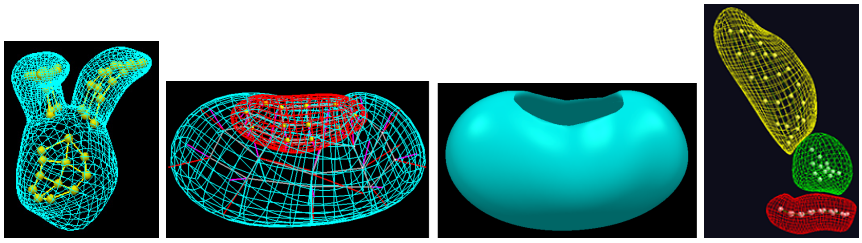


Fig. 2 A slab, in which the medial locus forms a curved 2D manifold of medial atoms. Sample atoms forming a grid are shown, and the balls illustrating their hubs are samples of the medial sheet. b) A tube, in which the medial locus forms a curve of medial atoms. In b1 sample atoms forming a chain are shown; the balls illustrating their hubs are samples of the medial axis. In b2 the boundary implied by the m-rep is shown.

A collection of figures that are attached among themselves form a multifigure object. Each figure can form a protrusion added to the collection (Fig. 3a), or a figure



a. b. c.

Fig. 3 Multifigure objects and multi-object complexes. a) A multi-figure male prostate with two seminal vesicles. b) The kidney minus the renal pelvis. c) The multiple objects making up a male pelvis: bladder, prostate, and rectum.

can form an indentation that is subtracted from the collection (Fig. 3b). Indentations can even pass entirely through the host figure, forming a hole. The form of attachment of subfigures to host figures is described in Section 8. Finally, the collection of figures can form multiple inter-related objects (Fig. 3c) that are non-intersecting but possibly abutting.

3 Object-relative coordinates

A special strength of the medial atom is that it carries a natural local coordinate system, complete with origin (its hub), coordinate frame (see Fig. 4), and metric (its spoke length), though strictly, medial atoms whose spokes are exactly opposed have one dimension of ambiguity in their coordinate frame; see below. Other medial atoms can be expressed in terms of a reference atom as

1. a translation of the reference atom's hub, in the metric of the reference atom (3 parameters). The origin of the atom's coordinate system is thus its hub location;
2. a magnification or demagnification of the spokes in common (1 parameter). The spoke length thus provides the distance metric of the atom's coordinate system and thereby makes local shape description of the figure invariant to local magnification;
3. rotations of the two spokes' directions, or alternatively a 3D rotation of the frame fitted to the spokes together with a 2D rotation of the spokes towards or away from each other in their common plane (both 4 parameters). The latter view breaks down in the singular situation when the spokes are back-to-back, i.e., one is a rotation by π of the other, which occurs at all critical points of the spoke width function. Nevertheless, it exposes the *object angle*, namely half of the angle between the spokes, and the bisector of the spokes. This bisector and the vector orthogonal to it in the plane of the spokes along with the vector orthogonal to this plane forms a natural medially fitted frame. The second of these, which is in the

direction of the difference of the spokes, can be shown to be normal to the medial sheet. A difficulty with this frame is that as one of spokes rotates continuously from nearly back to back with the other spoke, through back to back (object angle $\pi/2$), and beyond, the sense of the bisector changes discontinuously due to the fact that back-to-back spokes is a singular situation ($\nabla r = 0$, and the bisector is not unique).

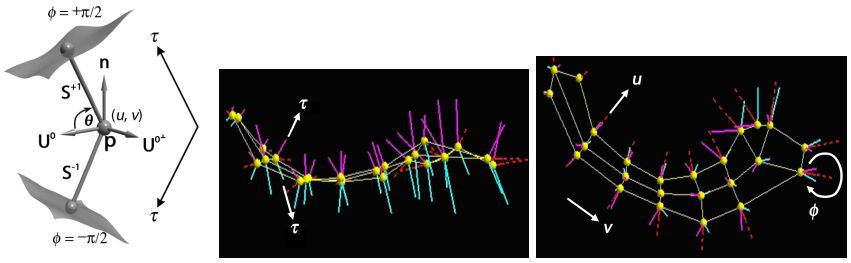


Fig. 4 Left: A medial atom's figural coordinates (u, v, ϕ, τ) and its atom frame ($\mathbf{U}^0$, $\mathbf{U}^{0\perp}$, **n**), where **n** is the normal to the medial sheet. Right and middle: Two views of the figural coordinates for a sheet of medial atoms (shown sampled).

By this approach any medial atom can be written as an 8-parameter (9 for end atoms; see Section 8) transformation of the primal medial atom, whose hub is at the origin, whose spoke length is unity, and whose spokes are both along the cardinal x-axis. These transformations, at various locations on the medial locus, provide an ability that is absent in many other object representations, namely, to provide combinations of local translations, twisting, bending, and swelling/contraction of interior material.

Writing each medial atom in a figure in terms of its immediate neighbors yields an object-relative coordinate system. In this coordinate system, whose coordinates we will write as **u**, the object is seen as a spoke-length-proportional dilation from the medial sheet or axis, according to the theory of Damon (Chapter 3). That is, one of the coordinates in **u**, which we call τ, is the fraction of the distance from the medial sheet or axis to the implied boundary at the spoke ends. τ -1 has the useful property that its sign distinguishes the inside from the outside of the figure. The spoke-length distance metric applied along the medial sheet provides two coordinates that can be called u and v. Applied to a tubular axis, it provides a single coordinate u.

An additional coordinate is needed to distinguish the two sides of the medial sheet and to take one around the crest from one side to the other. In the case of the tube this coordinate ϕ takes one around the tube by varying from -π to π. In the case of the slab this coordinate has a constant value of $\pi/2$ on one side of the object and of $\pi/2$ on the other side and changes smoothly between these while passing around the crest. In the Blum formulation ϕ changes discontinuously between -π and π at the crest. However, letting ϕ change continuously between $-\pi$ and π and be zero at the crest is more consistent with the tube representation and is friendlier to computer

representation because it stabilizes the end definition provided by image data (see Section 8). This decision, however, leaves open the definition of the point at which the end atom is placed, truncating either u or v such that the surface begins to be parametrized with ϕ and the non-truncated of u or v. Since the object angle must begin to move quickly towards 0 as the crest is approached, one possibility is to let ϕ begin to transition when the object angle magnitude falls below some threshold, e.g., $\pi/4$ (Fig. 4).

4 Figures, Subfigures, and Multi-Object Ensembles

When we divide an object into its figures, each of which forms a slab or a tube, we must describe the connection between the figures. Most commonly we think of one figure as a subfigure of a host figure, attached to the host along a connected locus (Fig. 5). However, it is possible for a figure to be attached to itself, or it can be attached to one or more other figures at a disconnected locus of points – consider the handle of a mug. We now treat the means by which the subfigure is attached to its host at any one of the one or more connected loci of points of attachment.

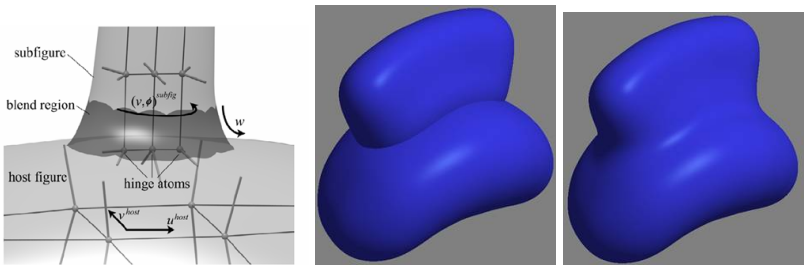


Fig. 5 Host figure and subfigure. Left: representation. Middle: as two separate figures. Right: blended.

One view is to form an attachment of a medial locus branch to the host's medial locus - this is the view that Blum and many mathematicians have taken. In that view, the patch metric on the medial locus described by Damon in Chapter 3 is typically quite small between the branch point and near the host figure boundary, as those medial atoms are responsible for little interior material. In the region of the branching the medial surface has a corner that has complex geometry, especially at the two ends of the corner.

Another view is to describe the subfigure fully by itself and to form a blend region to attach the subfigure to its host. Of course, this requires a description of the blend region and the means of smoothly attaching it both to the host figure and to the subfigure. But this view has the advantage that the subtractive (indentation) subfigure connects in a completely equivalent way as the additive (protrusion) subfigure.

In either case the part of the subfigure near the host figure (we call that part the *hinge*) needs to be understood in the coordinate system of the nearby medial atoms of the host figure. This representation allows the subfigure to be translated, rotated and scaled in the figural coordinates of the host figure (Han et al. (2004)). More precisely, the subfigure atoms near the connection to the host figure translate, rotate, and magnify their coordinates relative to the nearby host atoms while maintaining their coordinates in terms of the intrafigural neighboring atoms, and the remainder of the atoms maintain their coordinates in terms of the intrafigural neighboring atoms, all the while keeping the figural shape as close to constant as possible. The means of maintaining figural shape while deforming is a subject of the next section.

Multiple objects may be described in a similar fashion. That is, each object has its own m-rep, but in addition, to describe the relations among objects the atoms in one object that are near a second object also need to be understood in relation to the nearby atoms of the second object. This arrangement should be held mutually. With such an arrangement interpenetration among objects can be avoided by seeing whether the spoke ends of one object have a negative value of τ in the figural coordinate system of the second object.

5 Synthesis of Objects & Multi-Object Ensembles by Multiscale Figural Description

One point made in the first paragraph has not received adequate emphasis till now: the medial representation is being recommended as the basic means of representing objects, i.e., the representation from which other aspects of the object, such as its boundary and locations in its interior, are *synthesized*. To convey this point better, we jump off from the relation between the figures that make up an object and the object as a whole. Just as one cannot understand the leaves of a tree and the tree as a crown on a trunk at the same spatial scale (Koenderink (1990)), one needs to understand figures and the object as being at separate spatial scales. The figures can be understood as determining with finer tolerance information that is already conveyed by the object as a whole, at a larger tolerance. Moreover, the figural sheets can be allowed to have only two coordinates of parametrization precisely because small deviations from smoothness that the Blum medial axis might require in the sheet's boundary, leading to arbitrarily many subfigures, can be handled at a smaller scale level.

Continuing to a larger scale level than figures, the objects forming a multi-object ensemble determine with finer tolerance information that is already conveyed by the object ensemble as a whole, at a larger tolerance. Similarly, figural sections, each an interior region of the figure corresponding to a neighborhood of medial atoms (Fig. 6), determine with finer tolerance information that is already conveyed by the figure as a whole, at a larger tolerance. This process can be continued, down to the scale of the voxel, where individual voxels may be very locally displaced, rotated, or scaled to refine the tolerance given by the medial representation. In this sense,

an m-rep is a representation that is medial at the large and moderate spatial scales corresponding to objects, figures, and figural interior sections but even more local at the smaller spatial scales.

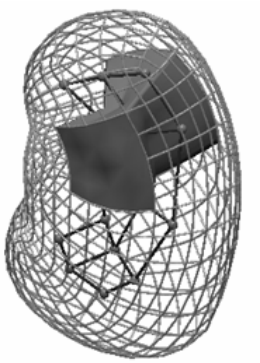


Fig. 6 Figural interior section corresponding to the neighborhood of a medial atom

Why do we represent objects in this large-scale-to-small fashion? It is because it is not useful to synthesize or deform an object in much more than O(N) time, where N is the number of primitives at the smallest scale, since an object or object ensemble will typically have thousands of primitives at that scale. Working at the smallest scale only and yet reflecting relationships of arbitrary degrees of locality would yield an $O(N^2)$ algorithm, which is unacceptably slow. The large-scale-to-small approach is designed so that calculation of the value of a primitive at one scale need only account for that primitive and its relation to neighboring primitives separated by its own scale. The hierarchical application of this principle provides O(N) algorithms for synthesis and deformation.

The synthesis of an object then occurs from large scale to small (Fig. 7). At the largest scale, that of the object ensemble, a base model is geometrically positioned, rotated, scaled and grossly warped, as described by just a dozen or fewer variables. To do this requires knowledge of the gross variations among the family of object ensembles being modeled, information that can be garnered from statistics on samples of the ensemble. How to represent the associated probability densities and estimate its parameters is discussed in Section 7.

For all of the remaining levels of scale, the next smallest scale is one at which the components of the previous scale are each refined. These successively smaller scales will be, at the least, the multi-object ensemble (if there are multiple objects being modeled), the objects, the figures of which they are made up, interior sections of those figures, and smaller subsections of those figures. More finely spaced scale levels are also possible. For example, one might choose to place larger figures in one scale level and smaller ones in the next smaller scale level.

Deformations of an object ensemble occur from large scale to small, with the smaller scale deformations being *residuals* from the larger scale ones, i.e., being

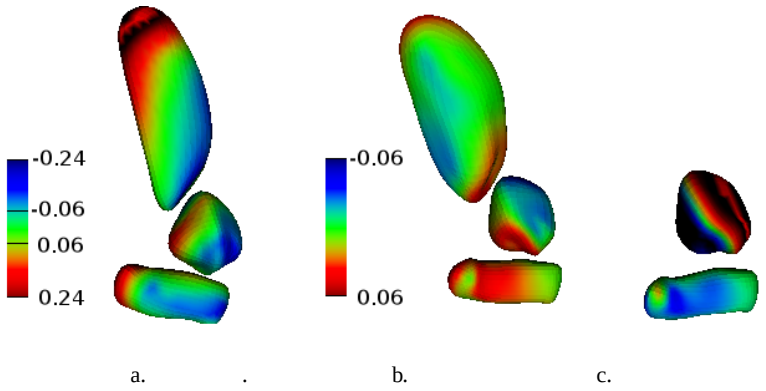


Fig. 7 The multiscale deformation of an object ensemble made from (top to bottom) bladder, prostate and rectum during a segmentation. The color bars next to panel a and panels b-c, respectively, show the amount and direction of the boundary displacement corresponding to each color. Positive values show displacements towards the exterior of the object, and negative values show displacements towards the interior of the object. The scale for panel a has larger values of displacement than that for the other panels. In each case the displacement is shown on the result from the next larger scale level, with scale levels processed from large to small. a. Ensemble deformation, displayed on initial m-rep. b. Deformation of the bladder produced at the bladder object stage, and sympathetic deformations of the prostate and rectum, displayed on ensemble stage result. c. Deformation of the prostate produced at the prostate object stage, and sympathetic deformations of the rectum, displayed on the prostate and rectum in the bladder stage result.

applied on the result of the large scale deformations. Thus, for example, the total deformation applied to a figure making up one object consists of the effect on it of a deformation global to the object ensemble, combined with a deformation global to the object of which it is a part, combined by the residual deformation of the figure itself.

When dealing with the residual deformation of an entity at one scale level, one needs to be aware that smooth deformations by arbitrary transformations on the atoms forming an m-rep figure may produce the illegality of one Euclidean point having two different figural coordinates. This may occur by the two local conditions of the medial sheet kinking and then, in further deformation, self-intersecting, or of the boundary kinking and then, in further deformation folding. Or it may occur by non-local self-penetration. The mathematics of Damon given in Chapter 3 have given local conditions on the medial locus that will allow one to prevent the local conditions, but non-local self-intersection may require a more expensive search.

Until now we have acted as if the topology of the m-rep, i.e., what are the figures and what is a subfigure of what host figure, is a given. But where should the model topology come from in the first place? Two possibilities present themselves.

First, based on knowledge of the application area, the user can understand what the parts are. The first example is of human anatomy: the lobes of the liver come from that discipline, and the smaller lobe can be seen as attached to the larger one.

A second example is of the automobile: a manufacturer understands that the bumper should be a part separately modeled from the car body.

Second, the m-rep topology can come from the statistics of instances of the object, each geometrically analyzed into medial components. Styner (Styner et al. (2003)) has suggested how the Blum medial analysis of instances from their boundaries can lead to a stable set of figures and figure/subfigure relationships even though each individual medial analysis is rather unstable. Essentially, in each instance portions of medial axis are grouped into sheets based on continuity of the medial sheet and of r values, sheets corresponding to an appropriately small fraction of the volume are deleted, and the branchings of the remainder are identified. Then sheets and branchings held in common across the cases are found. Indeed, some populations may need models of multiple branching structures to encompass the whole population, but for anatomic objects it is impressive how frequently a single branching structure will do.

6 M-reps as Symmetric Spaces

M-reps are designed to be deformed. One major use of deformation is in a statistical study of a population, in which an m-rep model is deformed into each instance in the population and then a probability density is derived from the collection of deformations. Another major use of deformation is in mechanical simulations on 3-space that includes one or more objects. Therefore the mathematical relationships among deformations on m-reps is important to understand. In the following we will see that the set of deformations forms what is called a *symmetric space*. A background to the mathematics of symmetric spaces and a description of m-rep deformations as a symmetric space can be found in (Fletcher et al., 2005). Also included in that chapter is a more thorough discussion of statistical analysis of m-rep objects than that included in Section 7.

Let us begin with a single medial atom. Let us consider all transformations on a medial atom, i.e., all combinations of translations of the hub, magnifications or demagnifications of the common spoke length, and rotations of the spokes. We shall see that this set of all transformations on the atom forms what mathematicians call a *symmetric space* made up of a Cartesian product of *Lie groups* and *quotients of Lie groups*, and the medial atom is a single point on this space. In the following we explain these terms and ideas and generalize them to full m-reps, preparing their application to statistics of m-reps, to interpolations of the transformations between m-reps, and to interpolations in space between samples of sampled m-reps (Fig. 1d).

Each medial atom can be understood as a transformation of the primitive medial atom, which, given a base (x,y,z) coordinate system, we will take to be the atom whose hub is at the origin, both of whose spokes point in the positive x direction, and whose spokes have unit length. The medial atom **m** then is understood as the translation of the hub from the origin to the hub position of **m**, combined with the magnification (multiplication) of the unit spoke length by the radial length of the

spokes in **m**, combined with the rotations of each of the spokes to their respective latitudes and longitudes on the sphere. The first two of these transformations are algebraic group operations, with the operations of vector displacement and multiplication respectively. That is, 1) we can compose two vector displacements or two spoke length multiplications, 2) for each of these operations there is an identity transformation, and 3) there is a unique inverse to each transformation. This group property, together with the fact that the set of transformations and the composition and inverse operations are smooth, are the definition of a Lie group.

The spoke rotational transformations do not form a group, among other reasons for lack of a unique inverse. However, each spoke rotation can be thought of as rotation of a sphere that includes no rotation about the spoke axis. Mathematicians call this *the quotient* $SO(3)/SO(2)$ of the Lie group of sphere rotations (3D rotations called $SO(3)$) and the Lie group of rotations about the spoke axis (2D rotations called $SO(2)$). This quotient of Lie groups has adequate smoothness due to the smoothness of the contributing Lie groups.

The result of a medial atom being in a set that is a Cartesian product of Lie groups and quotients of Lie groups is that the set of medial atoms can be understood as an 8-dimensional smooth manifold (9-dimensional for end atoms), albeit one that is curved (see Fig. 8). On this manifold, we can define a distance function and thus geodesic paths. For m-reps, the associated distance-squared function for an appropriately small m-rep transformation might be formed from the sum over the atoms of the sums of the squared displacements of the two spoke ends associated with hub translation, spoke length magnification/demagnification, and swings of the spokes, respectively. Alternatively, the squared displacements from these four atom components might be normalized according to the figural volume changes they produce.

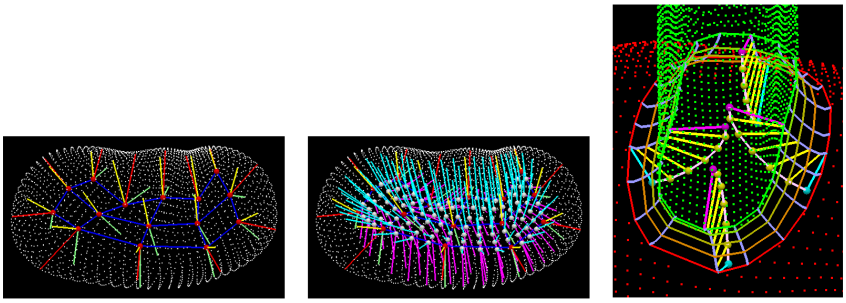
Fig. 8 a) A high-dimensional curved manifold of m-reps. On this manifold a point is a whole m-rep. The geodesic path shown between two such points (m-reps) gives a shortest route of transformations between the m-reps, a sequence shown for a hippocampus in part b). This same diagram (with the surface being of lower dimension) can be taken to illustrate the manifold of medial atoms. In that interpretation a point is a single medial atom. c) The tangent plane to a symmetric space and the Exp mapping of a vector in the linear tangent space to a geodesic path on the symmetric space.

A manifold is called a *symmetric space* if it satisfies the property that at each point there is distance-preserving diffeomorphism (bijective, smooth warp whose inverse is also smooth) on the manifold such that all geodesics through that point are reversed by that diffeomorphism. For the linear space of translations, the diffeomorphism is negation of the translation transferred to the origin. For the multiplicative space, the idea works by applying the linear theory to the logarithm of the magnification. For the rotations of axes on the sphere, the reversing map is rotation about that axis by π.

Cartesian products of symmetric spaces are also symmetric spaces. For example, a medial atom is a Cartesian product of the hub translation, spoke length magnification, and two spoke rotation symmetric spaces. Now consider a sampled m-rep, i.e., a collection of medial atoms to each of which the transformation, relative to its own center, can be applied. This discrete approximation to an m-rep is a Cartesian product of the symmetric spaces corresponding to each atom separately and is thus itself a symmetric space. The set of discrete m-reps of a particular structure made up of n_i internal atoms and n_e end atoms can thus be understood as a curved manifold of dimension $8\,n_i + 9\,n_e$. A point on that surface corresponds to a particular discrete m-rep (see Fig. 8).

Why have we gone to the trouble of thinking of a discrete m-rep as a curved, smooth manifold formed by a symmetric space together with a distance function? It is because for such manifolds geodesic calculations are closed-form algebraic operations implemented by two operations between a point $\mathbf{p}$ on the manifold and a point on the linear tangent space at that point. The mapping between the tangent hyperplane and the curved surface is called $exp_{\mathbf{p}}$, and the inverse mapping, between the curved surface and the tangent hyperplane, is called $log_{\mathbf{p}}$. As illustrated in Fig. 8, these mappings allow the difference between two m-reps to be computed; and as illustrated in Fig. 9. they allow the geodesic to be sampled to form an animation of the deformation between two m-reps. And they allow positions along that geodesic, measured by the fraction of the distance traveled between the endpoints, to provide an interpolation between two m-reps. While they also allow interpolation among two or more medial atoms, interpolation within a fixed m-rep is more appropriately done (Han et al., 2005) using interpolation of the medial sheet from the atom hub positions and the sheet normals given by each atom's spoke differences and interpolation of the two respective spoke swings using the mathematics given in Chapter 3. Strictly speaking, neither form of interpolation produces medial atoms in that the atoms' spokes are not orthogonal to the surface over the atoms' ends, but rather they are what Damon calls skeletal atoms (Chapter 3).

Frequently we are interested in the "shape space" of m-reps, that is, the set of m-reps modulo a global similarity transformation (7 dimensions: translation (3), rotation (3), magnification/contraction (1)). By analogy to shape spaces on flat (Euclidean) manifolds, we consider a space of m-reps that have been corrected by an alignment operation that minimizes the inter-object geodesic distances.



a. b. c.

Fig. 9 The interpolation of quad-meshes of spokes (Han et al., 2005). a) The uninterpolated atoms of a single figure kidney object. b) Interpolated atoms for that object. c) For the blend region of a multi-figure object (see Fig. 5). The red and green dots show the surfaces of the part of the host figure and the subfigure that are included in the object surface. The red and green curves delimit the region to be blended; on each curve sample points have associated spokes, shown in light blue and pink, on the host figure and subfigure, respectively. The yellow lines show spokes interpolated between the light blue and pink spokes; the endpoints of these spokes interpolate the blend surface between the corresponding points on red and green curves; these are shown for only four example points. The medium blue curves show similar interpolations between other pairs of corresponding points on the red and green curves.

7 The Statistical View of Objects

Instead of always talking about an individual entity, such as your liver or your car, as an object, we also frequently speak of an object, e.g., *the* automobile or *the* liver, as a population of instances, either across individuals or within an individual across time. Moreover, even a single object representation may be usefully understood as a center of a population of objects that agree with the representation to within some tolerance. The remaining question is how to represent such probability distributions on m-reps.

The first answer that may come to mind is to find the mean and principal components of the parameters describing an m-rep, since these can be used to parametrize a Gaussian probability distribution. Let us consider even the simplest situation, probabilities on a single medial atom. Because the transformations available include nonlinear ones of magnification and rotation, the linear theory of principal components is not suitable and, if used, generates geometrically illegal objects. However, the theory of principal components has been generalized by Fletcher et al. (Fletcher et al. (2004)) to the situation of symmetric spaces including nonlinear transformations. As described in Section 6, the mental leap comes from first understanding a geometric entity as a geometric transformation of a base entity (e.g., a medial atom as a transformation of the base medial atom as described above) and then considering it as a point in the of all such geometric transformations. Each point in that space corresponds to a transformation, i.e., to a geometric entity (though some of the points may be geometrically illegal). The idea is to do statistics on the collec-

tion of symmetric space points corresponding to a training population or to place probability density measures on the symmetric space.

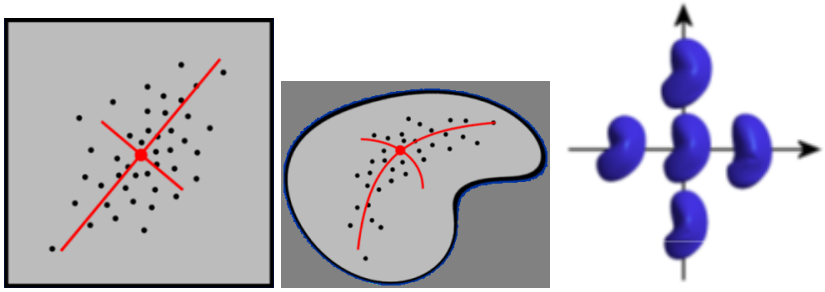


Fig. 10 Left: A flat symmetric space and Middle: a curved symmetric space, with means and principal geodesics. Right: The mean kidney m-rep is illustrated at the origin, with the m-reps ± 1 standard deviation from the mean along the two dominant principal geodesics shown along the horizontal and vertical axes.

Let us first consider the symmetric space of translations. Because addition of coordinate values is the operation of the underlying Lie group, the space of translations can be understood as a flat (Euclidean) space of transformations. That is, one can visualize the space as a flat, albeit high-dimensional, surface upon which the difference between two transformations is a vector along the straight line between the points corresponding to the two transformations (Fig. 8). The metric on that difference is the Euclidean difference between the points and is thus calculated using the Pythagorean theorem. Principal component analysis is appropriate in such spaces. In such a space the mean of a collection of geometric entities (points) is the point to which the sum of squared distances over the collection is minimum. And the subspace of the first k principal components is that k-dimensional linear subspace through the mean to which the sum of squared distances over the collection is minimum.

As illustrated in Fig. 10, this view can be generalized to the situations where the transformations defining geometric entities are nonlinear. In this case the space is curved, since straight lines between a pair of points achieved by linear interpolation of the parameters describing the points fail to stay on the surface of transformations. For example, the linear interpolation of two 3D rotation matrices or of two unit quaternions (represented as a 4-tuple of coefficients) element-by-element does not yield a 3D rotation matrix or a unit quaternion, respectively. This is because in the rotation Lie group the operation involves multiplication of the matrices or quaternions. However, if closest distance paths on this curved surface of transformations can be defined, i.e., geodesics in the space, the notions of mean and principal components can be generalized. The Fréchet mean of a collection of geometric entities (points) is the point to which the sum of squared geodesic distances over the collection is minimum. And the subspace of the first k principal components is

that k-dimensional subspace of geodesic paths through the mean to which the sum of squared geodesic distances over the collection is minimum.

The same idea can be applied to produce statistics on a tuple (Cartesian product) of transformations, i.e., to a sampled m-rep. A strength of this idea is that the different operations on an atom: hub translations, spoke rotations, and magnification of both spokes, can relate in their respectively appropriate ways. This idea can be well illustrated by considering the mean of two sampled m-reps. In producing the mean, the hubs may be translated along the vector in Euclidean 3-space between the two input hubs to the point halfway between them. At the same time, the rotations of a spoke must interpolate as rotations, i.e., on the surface of the sphere in 3-space, to the orientation halfway between the two input orientations. In addition, the lengths (magnifications) of the two spoke pairs needs to be interpolated according to its Lie group operation, namely multiplication. That is, the mean size should be the geometric mean of the two input sizes. This idea is all the stronger when it is applied to the Cartesian product of a set of atoms sampling an m-rep. Then articulations at joints are handled correctly, swellings that are local remain local, etc.

The approach just sketched depends on producing the measures on the Lie group according to which geodesics will be defined. The idea we are using at present is to make equal those differences in the different components of a medial atom that produce equivalent changes in Euclidean position in the ambient 3-space at the implied boundary, i.e., at the spoke ends. We make this choice because so much of the action in image analysis takes place at this boundary. Physical objects are visually sensed at the boundary. Image contrasts typically take place at the boundary. So even though we wish to represent the whole object interior and not just the boundary, the boundary has a special significance. Of course, if in a particular application the properties of interest take place at another place, e.g., some fixed fraction of the way from the medial locus to the boundary, i.e., some fixed value of δ, then the distance r in the following would need to be replaced by δr.

If the spokes have length r, an appropriately small translation of boundary by Δ can be accomplished by moving the hub by Δ, by changing r by Δ (multiplying r by $(1+\Delta/r)$), or by changing the spoke angle by Δ/r. Thus spoke angle changes of $\Delta\theta$ contribute $r\Delta\theta$ to distances, spoke length magnification by $1+\lambda$ contribute λr to distances, and translations of the hub by $\triangle\mathbf{p}$ contribute $|\Delta\mathbf{p}|$ to distances.

The means of computing the mean and principal components from a collection of training points on the symmetric space of sampled m-reps can now be specified. Basically, one either does trigonometry on the symmetric space surface itself, or one transfers the points to the tangent hyperplane at the present estimate of the mean via the $log_{\mathbf{p}}$ function, operates on that surface, and transfers the result back onto the symmetric space surface via the $exp_{\mathbf{p}}$ function. The second approach can be shown to work for computing the mean, as long as distances on the tangent hyperplane are taken to be the corresponding geodesic distances on the symmetric space surface. As long as the cloud of sample points is tightly clustered, the second approach produces a good approximation to either definition of principal geodesics defined on the symmetric space manifold:

1. the geodesic through the mean that minimizes the sum of squared geodesic distances to the geodesic;
2. the geodesic through the mean such that the points on that geodesic by geodesic projection of the sample point have maximum sum of square geodesic distances to the mean.

The latter two definitions, equivalent in flat space, are not equivalent in any curved space, e.g., the space of tuples of atoms (m-reps). Indeed, the approximation has certain properties preferable to either of the two definitions of principal geodesics defined on the symmetric space manifold, so we use this approximation in computing principal geodesics.

Two issues that arise when doing statistics are preliminary alignment of the training cases and setting positional correspondences among the training cases. We have discussed a method of alignment at the end of the previous section. The issue of correspondence is discussed in Section 9.

How does analysis of m-reps by geodesic alignment, Fréchet mean computation, and principal geodesic analysis compare in results to alignment of corresponding boundary point distribution models using the Procrustes method and principal component analysis of the corresponding points (the medial atom spoke ends)? Experiments on both simulations on ellipsoids and on interpatient data on real organs show cases where the principal variances and principal modes of variation differ significantly between the two forms of analysis. Indeed, in cases where the ellipsoid undergoes compositions of independent bendings, twistings, and spoke magnifications of m-reps, the analysis based on principal geodesic analysis of m-reps extracts the three basic transformations, but the principal component analysis of corresponding boundary points produces more modes, some with self-intersections and non-smooth boundaries and many of which are mixtures of the independent modes used in the ellipsoid formation. Details can be found in (Pizer et al., 2005a), in preparation.

Other probability distributions and statistical approaches can also be applied on symmetric spaces. Examples are probability distributions produced by Parzen windowing and other Gaussian mixtures and clustering approaches. Statistical approaches include discrimination by support vector machine and related methods, kernel methods, and indeed any statistical method that needs to be applied to objects or shapes.

Statistics may also be calculated on image intensities. However, image intensities, in 3D medical images in any case, only make sense in anatomic, i.e., in object-relative, coordinates. That is, for the intensities across images to correspond, whether position by position or region by region, the image space must be transformed to these coordinates before the statistics is done. This transformation may be within objects and outside near the boundary, or it may also also deal with interstitial regions between objects. Detailing of this approach is left to Chapter 9.

Finally, for deformations that take place within a particular object instance rather than between members of a population, physical models represented by partial differential equations and solved by finite element and other discrete approximations

can be aided by m-reps, in meshing (Crouch et al., 2003) and perhaps in solution via eigenanalysis directly from m-reps.

8 Discrete M-reps

To represent a continuous entity such as an m-rep in a computer, one must discretize it. The two most common means of doing this are by sampling or by producing a parametrized representation in terms of basis functions. In this section we discuss m-reps represented by sampling to produce *discrete m-reps*. In Section 10 we discuss a parametrized representation formed by m-rep splines.

The medial sheets forming a figure can be sampled by any appropriate sampling scheme: into triangular tiles, quadrilateral tiles, hexagonal tiles, etc. According to the idea that the spoke length r forms a distance metric, the spacing between the tile vertices should be approximately r-proportional. At each vertex is placed a medial atom (Fig. 4).

In a continuous m-rep the end curve is formed by atoms whose spokes have come together, i.e., the object angle is 0 (see Chapters 2 & 3)). The locus of the spoke ends of these atoms forms a crest on the boundary, i.e., a locus of a local maximum of curvature in the principal direction across the crest. Two problems arise with using such an atom of spoke-multiplicity 2 as an end atom in a discrete m-rep. First, the atoms spokes move towards each other at an infinite rate in the limit as the medial sheet end is approached (Fig. 11a), an unstable process. Second, deriving an atom from but a single point of image information is ill-conditioned, and doing this for an atom as critical to the shape as an end atom is therefore ill-advised.

Thus, as introduced in Chapter 1, we invented a new representation called *end atoms* (Fig. 11b) for figural ends that was designed to be stable in both of these senses. To avoid the infinitely fast collapse to zero of the angle between the medial spokes, it could allow only a subset of the types of ends generally allowed. At the same time, it needed to be consistent with the description used for the interior portion of the medial manifold. We therefore cut off the interior description while there was still a significant angle between the spokes and insisted that the Blum medial axis of the end portion continue straight from the place where the interior description stopped. That is, the end atom was provided with a new spoke along the bisector of the other two spokes with the new spoke incident to the crest of the implied boundary. By providing this bisector spoke with a specifiable length ηr, we provided a parameter η of crest sharpness additional to the parameters of the interior medial atoms. Even corners can thus be designated, using $\eta = 1/\cos(\theta)$.

The branching structure described in Section 5 was also developed to avoid instability in a continuous m-rep. In this representation the parts, called *hinges*, of the ends of protrusion (additive) and indentation (subtractive) subfigure sheets ride on the implied boundary of a host figure. These hinges are represented by designating a sequence of end atoms on the subfigure as hinge atoms. Knowing the hinge atoms' global position, orientation, and size parameters in the figural coordinates

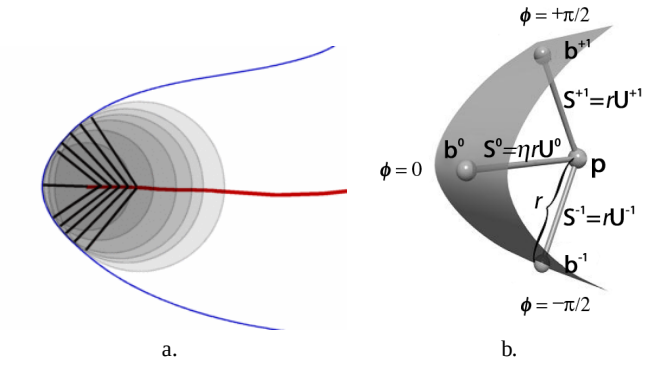


Fig. 11 Medial geometry of end atoms. a. In cross-section, the continuous relationship between a point on a medial surface and the points of contact between the disk of principal curvature inscribed at the point and the boundary of the object asymptotes at the crest point: equal steps along the medial surface result in increasing steps along the boundary. b. To account for this asymptotic relationship m-reps describe ends of figures using special end atoms with three spokes that describe the entire end-cap of the figure, symmetric about the crest at $\mathbf{b}^0$.

of the host boundary allows the subfigure to follow changes of its host figure while remaining appropriately connected to the host. The remainder of the atoms in the subfigure should then change to that member of the principal geodesic space of that subfigure for which the hinge atoms agree most closely, in geodesic distance, with their host-implied positions.

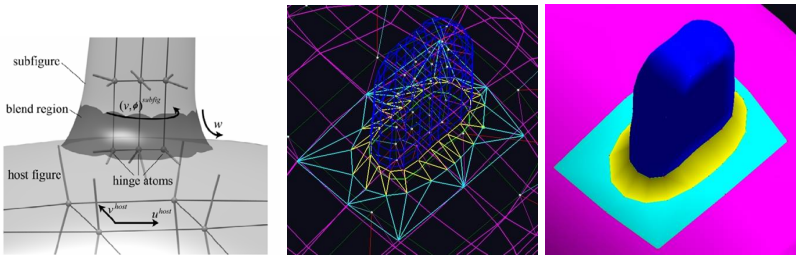


Fig. 12 A subfigure and host figure, their relation and blend, and hinge atoms.

The description of a subfigure riding on a host figure and thus intersecting the host figure requires a means of blending the seam of a subfigure with its host figure and of giving object-relative coordinates within the blend region. An approach consistent with the medial atoms of the host figure and of the subfigure, illustrated in Fig. 12, is to cut off the subfigure at a specified value of its u coordinate, produce a hole in the host figure by dilating the subfigure's intersection with the host figure in the (u, v) coordinates of the host figure, associate the positions on the two cut-off

curves based on an integrated minimum distance criterion, and between each pair of associated points interpolate the spokes. This interpolation is best accomplished using the ideas in Section 6, interpolating the hub position linearly, log of the spoke length linearly, and the spoke orientation along the spherical geodesic. The blended surface passes over the ends of the interpolated spokes. The degree of dilation and the subfigure cutoff coordinate are parameters of the process.

For speed of computation we presently approximate the surface interpolation via a grossly tiled blend and retiling and interpolative smoothing with normal agreement, using a successive subdivision algorithm described in detail in Thall (2004) (Fig. 12). This algorithm is also used to quickly approximate the boundary tile vertices of the individual figures making up an object. However, a method based on interpolating the medial atom hubs and spokes appears preferable. Such a method was briefly discussed near the end of Section 6.

The blend region needs its own figural coordinates. As illustrated in the left panel of Fig. 12, the v and ϕ coordinates of the subfigure serve to take one around the blend, δ takes one along the spokes, and we create a w coordinate that follows the interpolation from the subfigure cutoff to the edge of the hole in the host figure.

As an object deforms, not only can the individual figures that make it up change shape, but it is also useful to see subfigures as changing their overall conformation relative to their host. If they rotate or translate or magnify, they need to be seen as rotate or translate relative to the host figure, i.e., in the figural coordinates of the host. This includes not only rotation on the surface of the object but also rotation about a hinge in that surface, hinge atom by hinge atom, with a consequence for the remaining atoms within the subfigure's shape space, as described above. This provides a generalized notion of hinging even when the hinge curve is not straight.

One can handle the inter-object relationships in a similar way (Pizer et al., 2005b; Jeong et al., 2006). For each target object certain medial atoms in neighboring objects, namely those near and thus highly correlated with the target object are designated as neighbor atoms. These neighbor atoms are *augmented* to the target object atoms, and after alignment of these sets over the cases, principal geodesic analysis is applied to this augmented set, i.e., to the union of these atoms. The result is a shape space formed from the mean and set of chosen principal geodesic modes for the augmented set. For any position of the neighbor atoms, the projection of the aligned augmented set onto the shape space followed by restricting the result to the target object is the conditional mean of the target object, given the neighbor atom positions. This conditional mean is the target object value *predicted* by the neighbors. Prediction allows the target object to undergo changes sympathetic to the changes of its neighbors. This predicted value can be geodesically subtracted from the aligned target object itself to give a target object residue. A second principal geodesic analysis on these residues can be computed, yielding a shape space for the target object after prediction by its neighbors and a probability density on that space. Geodesic subtraction of the prediction and projection onto this neighbor effect shape space then yields the part of object variation that can be attributed to its own variation, as opposed to the effects of neighbor variation.

9 Correspondence of Discrete M-reps in Families of Training Cases

The principle expounded in this chapter is that a medial representation is usefully thought of as implying a boundary rather than the reverse. To compute the mean and principal geodesics associated with a family of training cases, each training case is typically given by its boundary. The problem is to produce an m-rep for each case such that the figural topology and number of medial atoms per figure are fixed and such that the individual atoms are in good correspondence across the cases. This issue of establishing correspondence is always of concern with discrete representations, but it is particularly of concern because many conformations of atoms can almost equally well describe the same boundary. One of the major strengths of discrete m-reps is its stability: moderate changes in the medial atoms can lead to only small changes in the implied boundary. But said another way, multiple arrangements of medial atoms can imply the same boundary to within some tolerance. This polymorphism can lead to problems of positional correspondence between different instances of the same object, e.g., by representing different individuals in a population or different states of the same individual.

One way to obtain correspondence is to take the point of view of Taylor, Davies, et al. (Davies et al. (2002)) that in a population of all reparametrizations of the members of the population, the one to be chosen is the one which has the tightest probability distribution. That is, degradation of correspondence is assumed to broaden the probability distribution. This point of view requires the notion of orbits of individual instance, that is, the subspace of all representations on the symmetric space that lead to the same boundary (perhaps to within some tolerance). Finding the best representative of each training case's orbit leads to a time-consuming computation.

A good step towards achieving correspondence at low cost can be obtained by depending on m-rep's edge fitting in a predictable fashion into a crest of the training case's boundary and then regularly spacing the atoms. The regular spacing can be obtained by including a measure of irregularity in an objective function to be minimized as the m-rep is fit to the object boundary. The measure of regularity that we use is the sum of the square deviations of each atom from the average of its neighbors, where the average and the deviation is computed via geodesic distances.

The objective function that we minimize to fit an m-rep of a given topology into an object is a weighted sum of the following terms.

1. The irregularity measure just mentioned.
2. The sum of squared distances between the implied boundary's tile vertices and the input object boundary. This distance function can be made large in other objects of a complex, to avoid interpenetration of the regions implied by the respective objects' m-rep sheets.
3. The sum of squared geodesic distances of the medial atoms to the corresponding atoms in a reference m-rep translated so that it shares a center of gravity with the m-rep being fit. (Rotation and scaling is also possible, using the second moments about the center of gravity.) Initially, the reference object may be taken as one of

the training objects, but after all the objects are initially fit, their mean can form a new reference object, and the fitting can be repeated.

4. The sum, over an optionally chosen set of landmarks on the reference object, of distance squared divided by tolerance squared. This helps with initialization, with getting the m-rep to fit into narrow elongated sections, and with avoiding rotation of the m-rep sheet.

As illustrated in Fig. 6, this algorithm works quite effectively, over a variety of complexes of single figure objects. A version for multifigure objects is in trial.

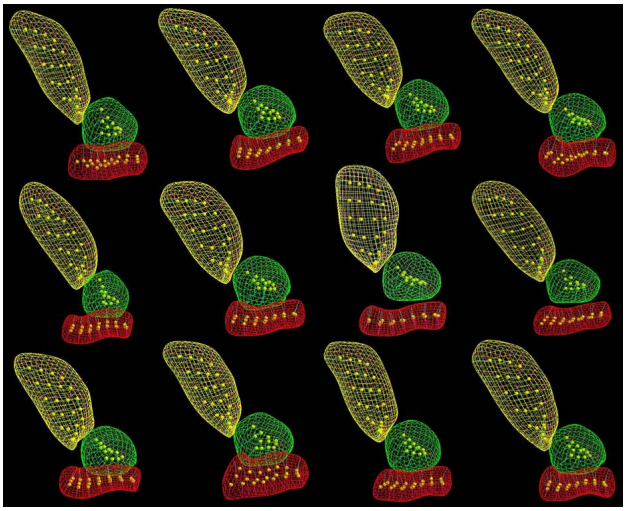


Fig. 13 An m-rep fit to 12 training cases of the bladder, prostate, rectum complex with correspondence.

10 Continuous M-Reps Via Splines or Other Basis Functions

With a parametrization of the medial sheet on (u,v), given the continuous functions of the hub locus $\mathbf{p}(u,v)$ and the spoke length value $r(u,v)$, the *envelope equations* for the family of spheres defined by these functions yield expressions restating equations 5 of Chapter 1 and the equations in proposition 7.1 of Chapter 2. The following analytic expressions for the spoke unit vectors $\mathbf{U}(u,v)$ and the spoke ends $\mathbf{b}(u,v)$ result:

$$\mathbf{U} = -\nabla r \pm \sqrt{1-|\nabla r|^2}\mathbf{N}\,, \tag{1}$$

$$\mathbf{b} = \mathbf{p} + r\mathbf{U}, \tag{2}$$

where $\mathbf{N}$ is the unit normal to the medial sheet and ∇r is the Riemannian gradient of the function r on the medial sheet (the Riemannian gradient describes the direction in which r changes fastest on the manifold $\mathbf{p}$ and its magnitude is equal to the rate of change per unit step in the tangent plane). These are given respectively by

$$\mathbf{N} = \frac{\mathbf{p}_u \times \mathbf{p}_v}{|\mathbf{p}_u \times \mathbf{p}_v|} \qquad \text{and} \qquad \nabla r = \begin{bmatrix} \mathbf{p}_u & \mathbf{p}_v \end{bmatrix} \mathbb{I}^{-1} \begin{bmatrix} r_u \\ r_v \end{bmatrix} \tag{3}$$

where $\mathbb{I}$ denotes the first fundamental form on the medial surface, given by the outer product of $\begin{bmatrix} \mathbf{p}_u & \mathbf{p}_v \end{bmatrix}$ with itself. The spokes themselves are simply $r\mathbf{U}$, the object angle $\theta(u,v)$ is given by $\cos^{-1}(|\nabla r|)$, and the frame $\mathbf{F}(u,v)$ is made from the vectors $\mathbf{N}$, $\nabla r/|\nabla r|$, and their cross-product.

Recognizing the ability to compute a continuous locus of full medial atoms from the continuous functions $\mathbf{p}(u,v)$ and $r(u,v)$, Yushkevich et al. (2003) proposed B-splines in (x,y,z,r) as a means of producing these continuous functions. Given a mesh of control points for these B-splines, the internal locus of a single-figure m-rep can be calculated. The greater challenge lay in forming the end curve and representing branches, where the vector ∇r must satisfy certain equality constraints.

Along the end curve, the equality constraint requires that $|\nabla r| = 1$, thus ensuring that the normal component of the spoke vectors $\mathbf{U}$ vanishes, and the spoke ends meet. In 2D continuous m-reps, Yushkevich was able to repose the problem as a constraint on the control points of the B-spline. In 3-D, however, the number of points at which the constraint must hold is infinite, while the number of control points is still finite; hence, the problem is overdetermined. A solution was obtained by letting the domain of the medial surface definition be the region in (u,v) space bounded by the zeroth level set of $|\nabla r| - 1$. In the B-spline framework, this solution was implemented by forcing $|\nabla r|$ to take large values on the perimeter of the unit square, making sure that $|\nabla r| < 1$ somewhere inside of the unit square, and then finding the level set to define the domain. The shape of the domain can be regulated using the spline control points, but the domain can not be fitted exactly to some prescribed curve. Figs. 14a-d illustrate how the implicit formulation makes it possible to define medial surfaces on free-form domains. Yushkevich et al. recently developed an alternative method in which the continuous m-rep can be based on any set of basis functions and the boundary of the domain can be explicitly defined; this method treats the end-curve constraint as a boundary condition of a partial differential equation (PDE) that can be solved quickly (Yushkevich et al. (2005)).

As illustrated in Fig. 15, the control points of the continuous m-rep can be adjusted to optimize the fit of the implied boundary to the boundary of a given binary image. In addition to minimizing the distance to the boundary, the optimization includes penalty and regularization terms that ensure that certain inequality constraints are satisfied and that the parametrization of the medial surface by (u,v) coordinates is more or less uniform. One of the penalty terms prevents the formation of singularities on the implied boundary by ensuring that the Jacobian of the function $\mathbf{U}(u,v)$ is positive.

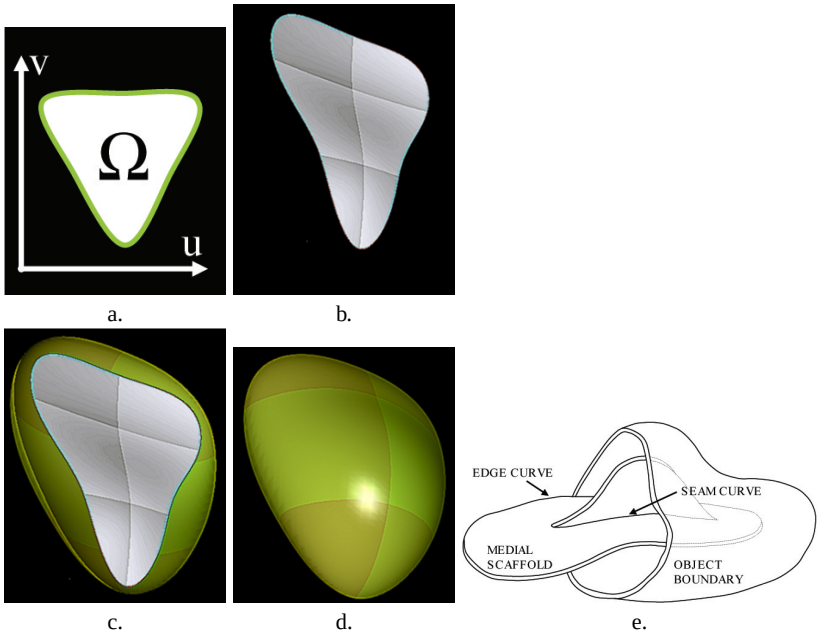


Fig. 14 **a.** A free-form domain Ω defined as the level set of the function $|\nabla r| - 1$. **b.** A B-spline medial surface interpolated on the domain Ω. **c.** One side of the boundary defined by the continuous m-rep. **d.** Both sides of the continuous m-rep's boundary. The constraint $|\nabla r| = 1$ ensures that the two sides form a closed surface. **e.** The fin-like formation of the 3D Blum medial axes poses challenges for continuous spline-based modeling.

The resulting continuous m-rep is a Blum medial locus of the object it implies. This contrasts with the locus of medial atoms produced by interpolating a discrete m-rep. That locus is a skeletal surface interpolating its medial atoms, but it is not necessarily a medial surface. The effect is that the discrete m-rep is typically able to fit a particular binary image more closely than the continuous m-rep.

The spline-based cm-reps method described above can model multifigure objects in 2D, but in 3D it is currently limited to single-figure m-reps. The alternative PDE-based method described in Yushkevich et al. (2005) appears to have the same limitation, though it does provide a means of computing the edge of the medial sheet explicitly rather than implicitly. Chapter 3 of the recently completed dissertation by Terriberry (2006) gives a method based on control points and implied control curves that not only provides all medial features explicitly but also supports medial branching. Terriberry's method also provides analytic calculation of various volume and medially implied boundary integrals useful in determining correspondences among a family of m-reps. Therefore, it appears to be an alternative for statistical m-reps worthy of exploration.

Terriberry's cm-reps provide an additional useful capability, calculation of integrals over medially implied object interiors or boundaries or parts thereof by integra-

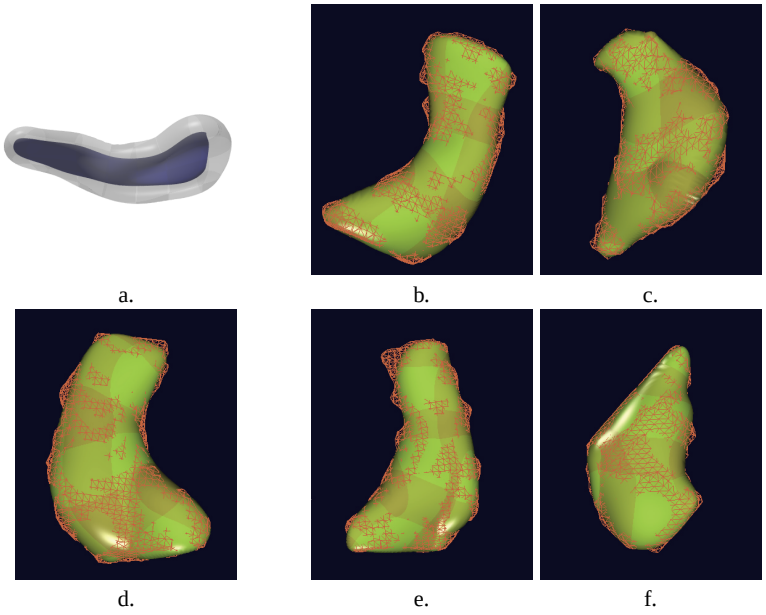


Fig. 15 Examples of fitting a continuous m-rep template to manually outlined hippocampus boundaries. **a.** A hippocampus template shown with a semi-transparent boundary. **b-f.** Examples of the fitted template's boundary (shown as green surface) displayed together with the hippocampus boundary outline (shown as orange mesh). The fit is lossy because the continuous m-rep template imposes a simple medial branching topology.

tion over the medial sheet(s) ((Terriberry, 2006), Chapter 4). Useful integrals over the object interior are moments of the medially implied object, the volume overlap between an object and a binary image, and various distances between the interiors of two objects. Useful integrals over the object boundary are the ratio of the area of a boundary patch and that of the corresponding medial sheet patch and regional averages of such geometric entities as mean and Gaussian curvature, principal curvatures, and principal directions. In particular, these and other integrals are useful in fitting a cm-res to a binary image or fitting cm-reps to multiple binary images with established correspondences over boundary points.

Both types of integration are done by pulling back the integration to the medial sheet according to the methods discussed in Chapter 3, Sections 4.1-4.4. Recall that the pullback requires calculation of the radial shape matrix S_{rad} at the hub of each spoke involved in the integration. The ability to express S_{rad} as an analytic function of the cm-rep spline coefficients allows the necessary integrations to be carried out using accurate adaptive subdivision numerical integration techniques on the medial sheet.

11 Summary and Conclusion

M-reps are designed for deformation by global and local combinations of object interior magnifications, rotations, and translations. The m-rep has the following advantages over alternative representations of objects or object complexes:

1. it inherently allows the deformation of an object to be decomposed into local translations, local twistings and bendings (rotations), and local magnifications;
2. it allows one to distinguish object deformations into along-object deviations, namely elongations and bendings, and across-object deviations, namely bulgings and attachment of protrusions or indentations;
3. it is especially designed to deal with both objects and the surrounding ensembles of objects;
4. it provides an anatomic object and object ensemble based coordinate system in terms of which to deal with the geometry-to-image match and allowing efficient determination of interpenetration of one object with another or a distant section of an object with itself;
5. it directly supports a series of object-based representation at successively smaller spatial scales and thus tolerances, allowing image analysis methods to proceed in an inherently efficient coarse-to-fine manner;
6. it deals with objects in terms of figures, i.e., protrusions and indentations that users frequently think of the objects in terms of and have names for.

A geometry and a statistics for such transformations has been devised, and a physical modeling capability is also partially developed. Among the strengths of this representation, providing efficiency, is a natural multi-scale framework made possible because of the representation's capabilities in providing within object and inter-object geometric neighborhood relationships. These capabilities can be leveraged in a wide variety of applications. Besides the image analysis applications of segmentation and hypothesis testing on shape populations described in Chapter 9, there are many other image analysis applications available and myriad applications in computer graphics and physically based modeling. Among the places of possible use in computer graphics, many of which have been piloted, are animation, texture rendering, image-based rendering, computer games, and computer-aided design. Physically based modeling using the nonlinear basis that m-reps provide is an open opportunity.

Nevertheless, m-reps, as presently designed, force the fitting of a fixed branching structure to objects in a population that may have a more complex or different branching structure. In addition, while the discretely sampled m-reps have an advantage of more tightly fitting objects over the parametrized m-reps, the discrete representation requires a more complex computational infrastructure, which is also more complex than that required by alternative non-medial object representations commonly in use. Also, like the the alternative object representations and associated statistical methods commonly in use, m-reps' statistical analysis is not prevented from including in the domain of random objects those that are geometrically

improper, e.g., ones that self-penetrate or have unsmooth boundaries, even if they are less likely than the alternatives to have such improprieties.

Acknowledgments

This work was done under the partial support of NIH grants P01 CA47982 and P01 EB02779, NSF SGER Grant CCR-9910419, and a fellowship from the Link Foundation. A gift from Intel Corp. provided computers on which some of this research was carried out. We thank Delphine Bull, Edward Chaney, Guido Gerig, A. Graham Gash, Ja Yeon Jeong, Conglin Lu, Gregg Tracton, and Joshua Stough for help with models, figures, software, and references. We are grateful to J. Stephen Marron, Keith Muller, and Surajit Ray for help with statistical methods.

References

Crouch, J., S. M. Pizer, E. L. Chaney, and M. Zaider: 2003, 'Medially Based Meshing with Finite Element Analysis of Prostate Deformation'. In: R. E. Ellis and T. M. Peters (eds.): *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Vol. 2878. Springer LNCS. [287, 340]

Davies, R., C. Twining, T. Cootes, J. Waterton, and C. Taylor: 2002, 'A Minimum Description Length Approach to Statistical Shape Modeling'. *IEEE Transactions on Medical Imaging* **21**(5), 525–537. [290]

Fletcher, P., C. Lu, S. Pizer, and S. Joshi: 2004, 'Principal Geodesic Analysis for the Study of Nonlinear Statistics of Shape'. *IEEE Transactions on Medical Imaging* **23**(8), 995–1005. [283]

Fletcher, P. T., S. M. Pizer, and S. C. Joshi": 2005, 'Shape Variation of Medial Axis Representations by a Principal Geodesic Analysis on Symmetric Spaces'. In: H. Krim and A. Yezzi (eds.): *Statistics and Analysis of Shapes*. Springer-Verlag. [280]

Grenander, U.: 1996, *Elements of Pattern Theory*. Baltimore, MD: John Hopkins University Press. [272]

Han, Q., C. Lu, G. Liu, S. Pizer, S. Joshi, and A. Thall: 2004, 'Representing Multi-Figure Anatomical Objects'. In: *International Symposium on Biomedical Imaging (ISBI)*, Vol. Catalog Number 04EX821C. pp. 1251–1254. [277]

Han, Q., S. M. Pizer, and J. N. Damon: 2005, 'Medial Atom Interpolation Via the Radial Shape Operator'. Technical Report in preparation, MIDAG, Department of Computer Science, UNC Chapel Hill. [282, 283]

Jeong, J., S. Pizer, and S. Ray: 2006, 'Statistics on Anatomic Objects Reflecting Inter-Object Relations'. In: *MICCAI Workshop on xx)*. p. xx. [289, 323]

Koenderink, J. J.: 1990, *Solid Shape*. Cambridge, Massachusetts: MIT Press. [277]

Pizer, S. M., J.-Y. Jeong, and Q. Han: 2005a, 'Principal Geodesics on M-reps vs. Principal Components on Boundary Points'. *to appear*. [286]

Pizer, S. M., J. Y. Jeong, C. Lu, K. Muller, and S. Joshi: 2005b, 'Estimating the Statistics of Multi-Object Anatomic Geometry Using Inter-Object Relationships'. In: O. F. Olsen, L. Florack, and A. Kuijper (eds.): *International Workshop on Deep Structure, Singularities and Computer Vision*, Vol. LNCS. pp. 59–70. [289]

Styner, M., G. Gerig, S. Pizer, and S. Joshi: 2003, 'Automatic and Robust Computation of 3D Medial Models Incorporating Object Variability'. *International Journal of Computer Vision* **55**(2/3), 107–122. [280]

Terriberry, T. B.: 2006, 'Continuous Medial Models in Two-Sample Statistics of Shape'. Ph.D. thesis, Department of Computer Science, University of North Carolina, Chapel Hill, North Carolina. [293, 294]

Thall, A.: 2004, 'Deformable Solid Modeling via Medial Sampling and Displacement Subdivision'. Ph.D. thesis, University of North Carolina, Chapel Hill, North Carolina. [289, 315]

Yushkevich, P., P. T. Fletcher, S. Joshi, A. Thall, and S. M. Pizer: 2003, 'Continuous Medial Representations for Geometric Object Modeling in 2D and 3D'. *Image and Vision Computing* **21**(1), 17–28. [20, 292]

Yushkevich, P., H. Zhang, and J. Gee: 2005, 'Parametric Medial Shape Representation in a 3-D via the Poisson Partial Differential Equation with Non-linear Boundary Conditions'. In: *International Conference on Information Processing in Medical Imaging*, Vol. LNCS 3565. pp. 162–173. [292, 293]

Part III
Applications

Statistical Applications with Deformable M-Reps

Anatomic Object Segmentation and Discrimination

Stephen Pizer and Martin Styner and Timothy Terriberry Robert Broadhurst and Sarang Joshi and Edward Chaney and P. Thomas Fletcher

Abstract There are many uses of the means of representing objects by discrete m-reps and of estimating probability distributions on them by extensions of linear statistical techniques to nonlinear manifolds describing the associated nonlinear transformations that were detailed in Chapter 8. Two important ones are described in this chapter: segmentation by posterior optimization and determining the significant shape distinctions that can be found in two different probability distributions on an m-rep with the same topology but from two different classes. Both uses require facing issues of probabilities on geometry at multiple levels of spatial scale. The segmentation problem requires the estimation of the probability of image intensity

Stephen Pizer
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: pizer@cs.unc.edu

Martin Styner
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: styner@cs.unc.edu

Timothy Terriberry
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: tterribe@cs.unc.edu

Robert Broadhurst
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: reb@cs.unc.edu

Sarang Joshi
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: sjoshi@cs.unc.edu

Edward Chaney
Medical Image Display & Analysis Group, University of North Carolina at Chapel Hill, USA
e-mail: edward_chaney@med.unc.edu

P. Thomas Fletcher
Department of Computer Science, University of Utah, USA
e-mail: fletcher@sci.utah.edu

distributions given the object description; we describe a way of doing that by an extension of principal component analysis to regional intensity summaries produced using the object-relative coordinates provided by m-reps. Applications of both segmentation and determination of shape distinctions to anatomic objects in medical images are described. Also described is a variant on the segmentation program used in estimating the probability density on an m-rep; this program fits an m-rep to a binary image in a way that is intended to achieve correspondence of medial atoms across the training population.

1 Introduction and Statistical Formulation

Both segmentation, i.e., extraction, of objects from images and characterization of geometric differences between classes of objects are usefully accomplished in terms of deformable shape models. In segmentation a geometric model is deformed into the image data, allowing the method to reflect an understanding of what legitimate or typical shapes are. In characterizing the differences between shapes in two different populations, the differences are measured in terms of the deformation from one shape to another. Medial models provide a useful representation of the object or complex of objects that undergoes deformation and of the deformations themselves. Moreover, statistics on medial models are useful for both applications, specifying the typicality of a shape or the population of deformations between shapes in the two classes being compared. Finally, the segmentation application requires not only statistics on the geometry, i.e., on the medial models or their deformations, but also statistics on the image intensities, given a medial model. Because these intensities are best understood statistically in object-relative coordinates, the figural coordinates provided by m-reps are an important means of producing the image intensity statistics.

In Chapter 8 the geometry of discrete m-reps and statistics on these entities were discussed. This chapter discusses the use of these geometric representations and their statistics, as well as the statistics on image intensities in figural coordinates for segmentation of anatomic objects and object complexes. It also discusses the use of these geometric representations and their statistics for statistical shape difference characterization between classes of anatomic objects or object complexes extracted from medical images, e.g., between the hippocampi or lateral ventricles of healthy and schizophrenic individuals as extracted from magnetic resonance images.

In characterizing the difference between two anatomic populations the differences need not only to be specified statistically, but also this specification needs to include *where* the differences are and *what form of deformation* occurs there, for example, whether it is a local twist or a local bend or a local swelling or a local contraction. Also, in segmentation, a coarse-to-fine, i.e., successively more local approach has serious speed advantages for any given quality of segmentation. M-reps with their coordinate systems, their provision of multiscale statistics, and their

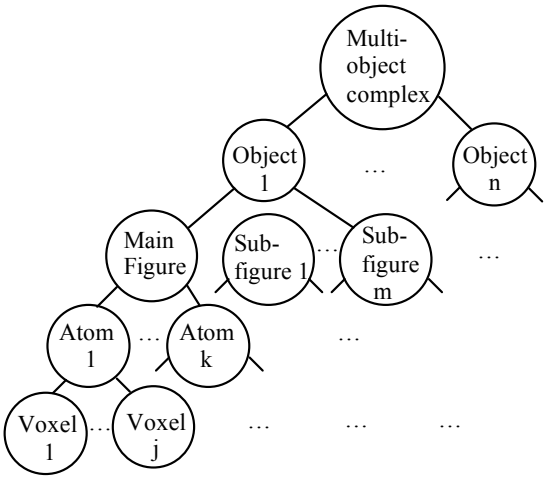


Fig. 1 A tree of objects, figures, medial atoms and voxels.

medial basis' provision of both local width and local figural orientation are well matched to these needs.

More precisely, as illustrated in Fig. 1, consider a tree of geometrical entities such that the discrete m-rep at the root of the tree describes a whole object complex and such that the children of a node describe sub-entities which taken together make up that entity. For example, if the root node describes a complex of objects, its children would respectively describe each object making up the complex. Similarly, if a node describes an object made up of figures, its children would respectively describe the figures making up the object, their children might describe individual medium atoms, and their children might describe sequences of displacement on individual voxels. In each node is a collection of atoms made up of all its children, each atom with a value. The value of a node is the atom values of all the atoms making up that node. Then deforming the entity corresponding to a node deforms all of its sub-entities, and after that deformation we may move on to the sub-entities of the node and deform them further in some order. We refer to these stages at which processing occurs as *scale-levels.*

At each scale-level other than the top of the tree, an entity **m** has as set of neighbors $N(\mathbf{m})$, that are at nearby physical positions. It is useful to think of the probabilistic relationship among entities in terms of the value of each child of a node, given the value of that node and the conditional probability of a node given the values of its neighbors. The former describe inter-scale-level differences, and the latter describe inter-neighbor differences, i.e., differences across position. This view al-

lows us to think of the problem with a Markov random field formulation in both the scale and positional dimensions.

That is, if the m-rep **n** is a child sub-entity of an m-rep **m**, and $\mathbf{m} \rightarrow \mathbf{n}$ is the value that **n** takes as a result of the deformations of its ancestors and most recently as a result of its parent **m**, we wish the conditional probability of the deformation describing the difference $(\mathbf{m} \rightarrow \mathbf{n}) \ominus \mathbf{n}$, given the parent node **m**, where the symbol $\ominus$ denotes the geodesic path between its two operands. Similarly, if $< N(\mathbf{n}) >$ describes the prediction of **n** based on its neighbors, we wish the conditional probability, $p(\mathbf{n}\ominus < N(\mathbf{n}) > | N(\mathbf{n}))$, of $\mathbf{n}\ominus < N(\mathbf{n}) >$ given the neighbor nodes $N(\mathbf{n})$. Because the essence of geometry is that entities are locally correlated, a thesis that for medial atoms of various anatomic objects is supported by our data, it is reasonable to condition $(\mathbf{m} \rightarrow \mathbf{n}) \ominus \mathbf{n}$, on only the parent node **m** and not on ancestors more distant in scale, and it is reasonable to condition $\mathbf{n}\ominus < N(\mathbf{n}) >$ on its immediate neighbors $N(\mathbf{n})$ and not on more distant entities.

In the work described here, we simplify the probabilistic formulation even further. We assume that $(\mathbf{m} \rightarrow \mathbf{n}) \ominus \mathbf{n}$ is statistically independent of **m** and that $p(\mathbf{n}\ominus < N(\mathbf{n}) > | N(\mathbf{n}))$ can be broken up into two factors, one describing the change independent of its neighbors and the other describing the interrelationship of it with its neighbors. Breaking things down according to this Markov formulation allows a segmentation or hypothesis test with final locality such that the total number of primitives at that level of locality is M (e.g., there are M voxels in the objects being segmented or at which shape differences are being tested) to operate in $\mathcal{O}(M)$ time rather than the $\mathcal{O}(M^2)$ that are required when the relation of every primitive with every other one must be dealt with.

The geodesic differences between m-reps used in the foregoing formulation are in the same symmetric space as the subtrahend and the minuend. That is, the, the geodesic differences of a collection of medial atoms is the collection of differences of the corresponding atoms, and the difference of two atoms is the Cartesian product of the corresponding components, as illustrated by the difference between interior slab atoms in the following:

1. the difference of the hub positions, which like a hub position itself is a vector in $\mathscr{R}^n$;
2. the "difference" of the spoke lengths, which is the ratio of these lengths giving the magnification of one into the other, and thus like a length itself is a scalar in $\mathscr{R}^+$;
3. the "difference" of each spoke position on the unit sphere S^2 with the corresponding spoke's position on S^2, which can be understood as a position on S^2. There are difficulties with differences of angle differences associated with having to specify a reference angle; these will not be further discussed here.

As a result, statistics on such geodesic differences can be accomplished by the same methods of computing means and principal geodesics described in Chapter 8.

Finally, consider the probabilities on differences of m-reps that are the target of statistical characterization of inter-class differences. These differences of m-reps are again in the same symmetric space as the subtrahends and minuends. One requires

methods of hypothesis testing that yield the significance of distinctions in probability distributions in this symmetric space and, as well, the location of such significant changes, for various levels of locality.

In Section 2 we introduce segmentation via posterior optimization of deformable m-reps with an overview of the approach. We find that two log probability densities are needed, one measuring the geometric typicality of an m-rep and the other measuring the match between the m-rep and an image. In Section 3 we discuss how to train the first probability density, given binary images of sample objects, and how to measure this geometric typicality on any m-rep, given this training. In Section 4 we discuss estimating the probability density on image intensities given a medial model and how to measure this probability density on any target image. In Section 5 we conclude our discussion of segmentation by specifying the segmentation scale at the smallest scale level, that of the voxel, followed by the excellent results obtained using our multi-scale method using the geometric and intensity probabilities. In Section 6 we discuss means of hypothesis testing based on m-reps for statistical characterization of shape differences between populations of objects or object complexes. Section 7 gives some examples of results using this method. In Section 8 we discuss the apparent strengths and weaknesses of the medial methods we propose for the segmentation application and characterization of shape differences application, as compared to alternative object representations. In that section we also discuss work that remains in both these methods of application of m-reps and in the formulation of m-reps themselves and their statistics.

2 Segmentation by Posterior Optimization of Deformable M-reps: Overview

Published studies by others and our own research results strongly suggest that segmentation of a normal or near-normal object (or objects) from 3D medical images in all but the simplest cases will be most successful if it uses 1) knowledge of the geometry of not only the target anatomic object but also the complex of objects providing context for the target object and 2) knowledge of the image intensities to be expected relative to the geometry of the target and contextual objects.

We use the general segmentation approach already shown by others to lead to success ((Cootes et al., 1993; Staib and Duncan, 1996; Delingette, 1999), among others; also see (McInerny and Terzopoulos, 1996) for a survey of active surfaces methods), namely deforming a geometric model by optimizing an objective function that includes a geometry-to-image match term which is constrained by or summed with a geometric typicality term. In this approach a model of the object(s) to be segmented is placed in the target image data and undergoes a series of transformations that deform the model to closely match the target object.

In computer vision an important class of methods uses explicit geometric models in a Bayesian statistical framework to provide *a priori* information used in posterior optimization to match the deformable shape models against a target image. Using

this approach, we start from a statement of the segmentation objective as finding the most probable conformation of the target object(s) **m** given the image I, i.e., of computing $\text{argmax}_{\mathbf{m}} p(\mathbf{m}|I)$. Here **m** is the geometric representation of the target object(s), in our case the tree of medial atom meshes that comprises an m-rep, and I is a tuple formed by a 3D array of image intensities. The probability density $p(\mathbf{m}|I)$ is frequently called the *posterior density*, so the method is called one of *posterior optimization* (Duda et al., 2001).

By Bayes rule, $\text{argmax}_{\mathbf{m}} p(\mathbf{m}|I) = \text{argmax}_{\mathbf{m}}[\log p(\mathbf{m}) + \log p(I|\mathbf{m})]$. Thus the geometric typicality term ideally measures the logarithm of the so-called *prior* probability density, the probability density that the candidate geometric entity exists in the population of objects, as described in Chapter 8. And the geometry-to-image match term ideally measures the logarithm of the so-called *likelihood*, the probability density that the target image values, relative to the candidate geometry, would arise in the population of images from that modality. As a fundamental means of obtaining efficiency, we optimize such an objective function for successively smaller spatial tolerances (spatial scales), where each of the spatial scale levels are object-relevant: the object complex, the object, the slab (or tube) figure, the figural section, and the voxels not only interior to the object(s) but also the voxels between them, which we call *interstitial* voxels.

The success of the deformable shape models posterior optimization approach depends on the object representation, i.e., the structural details and parameter set for the deformed model, as well as on the form of the objective function. The most common geometric representation in the literature of segmentation by deformable models is made up of directly recorded boundary locations, sometimes called *b-reps* (Cootes et al. (1993); Kelemen et al. (1999), also see papers surveyed by McInerny and Terzopoulos (1996)). Our m-reps representation (Fig. 2), principal geodesic analysis to produce its statistics, and the associated segmentation method use a medial representation intended to produce improved and/or more efficient segmentations for the reasons given in Chapter 8, Section 11. The most relevant of these advantages for this application are the efficient training of the prior it provides, its ability to provide a coordinate system in which to describe intensities probabilistically, and its inherent multi-object, multi-scale nature, which leads to effectiveness and efficiency of segmentation of single or multiple objects. However, small indentations and protrusions of anatomic objects are impractical to model medially. Our approach to solving this problem is to implement a non-medial voxel stage described in Section 5.1.

M-reps, combined with the voxel-level representation, provide their advantages over other deformable object representations at the expense of a level of complexity that required the development of special theoretical underpinnings, software, and validations. Largely automatic segmentation by large to small application of deformable m-reps has been implemented in software called *Pablo* (Pizer et al., 2005b) that accomplishes 3D segmentations in a few minutes. Software for building and training models has also been developed. The methods underlying this software and its abilities are the subject of Sections 2-5.

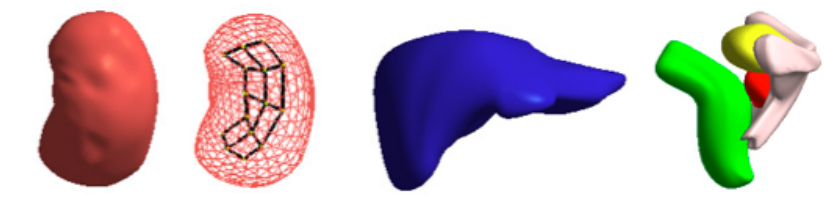


Fig. 2 M-rep modeled kidney with its medial mesh, a liver model that is made from two figures, one for each lobe, and male pelvis model made from multiple objects (two bones, bladder, rectum, prostate). The kidney model also shows the underlying representation of a sampled medial surface and a tiled boundary.

The next two sections give a more specific picture of Pablo's method (Section 2.1) and operation (Section 2.2).

2.1 Segmentation Method: Posterior Optimization for Multiscale Deformation of Figurally Based Models

Our method for deforming a model into image data typically begins with a manually chosen initial positioning of the mean model, frequently via choosing a few rough landmark positions. The segmentation process then follows a number of stages of segmentation at successively smaller levels of scale. The spatial tolerance of the resulting segmentation can be large at the largest scale level but decreases as the scale gets smaller.

As illustrated in Fig. 3, at each scale level, i.e., level of the tree shown in fig. 1, the same log prior + log likelihood objective function is optimized by geometrically transforming the entities at that scale level, using a transformation global to the respective entity. Thus, at the largest scale level, the object ensemble stage, the whole object ensemble undergoes a global transformation. At the next smaller scale level, each object making up the object ensemble separately undergoes a transformation global to it. And as the computation moves to successively smaller scale stages, successively smaller entities making up the entities at the next larger scale level, namely figures, subfigures, and medial atoms, are optimized with a transformation global to each of them. The series of optimizations concludes with a small relocation of all of the voxels in the image being optimized.

At all of these scale levels, we follow the strategy of iterative conditional modes, so the algorithm cycles among the component entities in random order until the

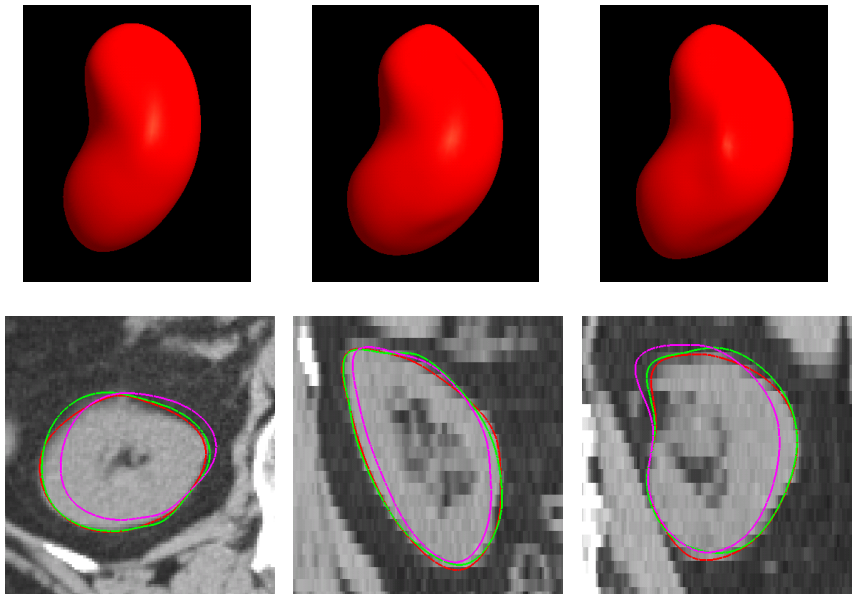


Fig. 3 Stage by stage progress of deformable m-rep segmentation of the kidney. Top: rendered 3D view, after model alignment via landmarks, the figure stage, and the figural section (atom) stage. Bottom: results on axial, sagittal and coronal CT slices. Each image compares progress through consecutive stages via overlaid curves: magenta - aligned position; green - post object stage; red - post atom stage.

group converges[1]. For example at the figural atom stage, the algorithm cycles through the atoms in random order.

At each scale level larger than the voxel scale level, the geometric transformation of the entity is made up of a typically deterministic similarity transformation and a maximum posterior warp. The similarity transform, a translation, rotation, and uniform magnification, aligns the entity to neighboring entities of the same type (objects to neighboring objects, medial atoms to neighboring atoms), except it aligns to landmarks at the largest scale. The warp is formed from a few principal geodesics (see Chapter 8) of the deformations of that entity experienced in the training data. At the voxel scale level, the optimization is over displacements per voxel of only a few voxel widths. The result is that we typically optimize 6 or fewer parameters per entity, providing efficiency and convergence of the segmentation at that scale level.

[1] The convergence properties are shared with all iterative conditional modes methods and are based on the underlying Markov random field. In practice, convergence always occurs, but sometimes the convergence is to a local maximum of the objective function rather than the desired global maximum.

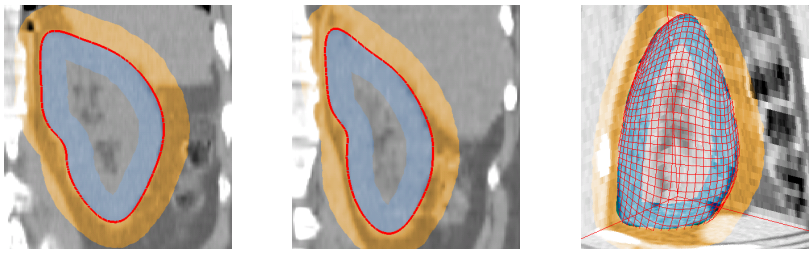


Fig. 4 Boundary-relative regions used for measuring geometry-to-image match to a kidney. a, b: Example from two different patients displayed in 2D cuts. The kidney interior region is portrayed in blue, and the kidney exterior region is portrayed in orange. c: A mesh showing in 3D the m-rep implied boundary of the kidney, and the kidney interior and exterior regions in two orthogonal cuts through the 3D image.

At each scale level we use the conjugate gradient method to optimize the log prior + log likelihood objective function. The log prior metric is detailed further in Section 3. As detailed in Section 4.2, we have implemented a way of computing the log likelihood that measures the geometry-to-image match based on probability densities on intensity distribution features in various figural-coordinate-specified regions inside and outside of the object (Fig. 4) such that each region is expected to be a constant mixture of tissue types (Broadhurst et al., 2005).

2.2 Segmentation method: user operation

M-rep-defined objects can be viewed as a boundary mesh (at any of a number of vertex spacing levels), a rendered surface, a collection of points at the aforementioned boundary vertices, or a medial atom mesh. Most users find the first two of these the most useful. Images are normally viewed in a tri-orthogonal display, with the three possible slice directions fixed to the cardinal within-image and cross-image slice directions given by the stored target image. The displayed object can be presented together with the intensity display (see Fig. 3). Moreover, we also provide a boundary display mode on the displayed slice, in which the 3D object does not appear but the curve of its intersection with the displayed slice(s) is displayed on that slice (those slices).

Using these viewing mechanisms, the user either chooses the location of preselected landmarks in the target image, which is then used as the basis of an Procrustes initialization of the model, or he or she manually initializes the chosen model by placing it in an initial position relative to the 3D image (for example, see Figs. 3-bottom row, 4c, and 9-bottom left). The initialization transform derived from the landmarks is frequently a similarity transform, but we have found it also useful

that this landmark-based transform optimize in the shape space of the principal geodesics of the object with a data-match term given by the sum of squared model landmark to image landmark squared distances, with each squared distance divided by its tolerance squared.

The landmarks on the model are chosen as a specified spoke end. These landmarks appear as colored spots on the base model in the display space. These landmarks can also be used for editing an m-rep in the middle of the optimization process or as another term in the geometry-to-image match.

The user is also given control of the values of the weights controlling the strength of the geometric typicality term in the objective function, relative to the geometry-to-image match term. However, since the two terms are now both Mahalanobis distances, the default weight of unity needs seldom be changed.

3 Training and measuring statistical geometric typicality

To be able to measure a log prior, one needs a parametrized function that one can evaluate with any m-rep for the desired object as the argument. Section 3.1 describes the means for training the parameters of this prior probability density on m-reps that is then used to measure geometric typicality of any candidate m-rep appearing in the optimization of the log posterior. This training of the prior is done by principal geodesic analysis of m-reps fit to binary images extracted from training greyscale images. Section 3.1 describes both the fitting of m-reps to binary images and how principal geodesic analysis is used at multiple scales to produce the prior probabilities needed for the various scale levels. Section 3.2 describes the means for measuring the log prior at multiple scales needed in the multiscale segmentation procedure.

3.1 M-rep model fitting and geometric statistics formation

Model-building must designate the figures making up an object or object ensemble, give the size of the mesh of each figure, and give the way the figures are related. It must also specify each medial atom in the model forming the mean object or object ensemble and the variability of these at many scale levels. Illustrated in the panels of Fig. 5 are m-rep models of a variety of anatomic structures that we have built. In the following we sketch our model building procedure, leaving the details of how we meet this challenging goal to other papers (Merck et al., 2006).

Because an m-rep is intended to allow the representation of a whole population of an anatomic object across patients, we build it based on a significant sample of instances of the object. Typically we use some tens of instances, say 50. For each instance we begin with both a 3D binary image representing the interior of the

object, typically manually segmented, and an associated 3D greyscale image (CT or MRI or another modality).

Styner et al. (2003a) describe a tool for producing m-rep models from such binary image samples, based on the principle that effective segmentation depends on building a model that can easily deform into any instance of the object that can appear in a target image. We can use this tool to compute the set of appropriate figures at a given level of approximation from a training population, or we can choose the figures based on anatomic expertise to correspond to named anatomic structures. The tool measures the level of approximation in the figure computation step via error in volume overlap (typically 98%). In either case, given the figures, the tool chooses the number of atoms in each figure as the minimal number that can fit every training instance to a given error measured by the mean absolute distance of the surfaces (typically 5% of the average radius).

More recently we have completed a stable web-sharable tool called *Binary Pablo* for fitting an m-rep model to each member of a collection of binary images and deriving the Fréchet mean and principal geodesic modes and variances (Merck et al., 2006). Once a base model is generated, we use Pablo to deform it into the binary segmented training images. The program optimizes an objective function that has an "image match" term giving an average distance between the boundary implied by the m-rep and the binary image boundary, and three geometry terms: 1) giving an average squared-distance between each atom and the geodesic average of its neighbors, thus producing a regular mesh of atoms; 2) discouraging folded objects by penalizing rS_{rad} eigenvalues $\varepsilon 1$ (see Chapter 3); 3) giving a squared-distance from a reference m-rep. The sharable version only operates for single-figure objects, but versions that fit m-reps to multi-figure objects and to multi-object complexes are available in our research toolkit.

Given the m-rep models for all the training cases (Fig. 6), we use a tool initially developed in Dam et al. (2004) and further developed by Lu (Lu et al., 2003) to compute the mean model and the principal-standard-deviation-weighted principal geodesics describing its variability. This tool uses the method of principal geodesic statistics on symmetric spaces described in Chapter 8, Section 7. As with linear statistics, each principal geodesic has an associated variance, and moving along that geodesic gives a principal mode of variation of the population of m-reps.

The statistics at one scale level need to describe the variability of the geometric entity at that scale level after the variability at the larger scale levels has been accounted for and after alignment to neighboring entities has been done. Description of this residue statistics, based on the theory of Markov random fields, is given in [Lu 2003].

With these means and a number of principal geodesics chosen to capture some fraction of the variance at that scale level, deforming a geometric entity at that scale level involves aligning the object to its neighbors and then computing the coefficients of the principal geodesics of the deformation of that entity.

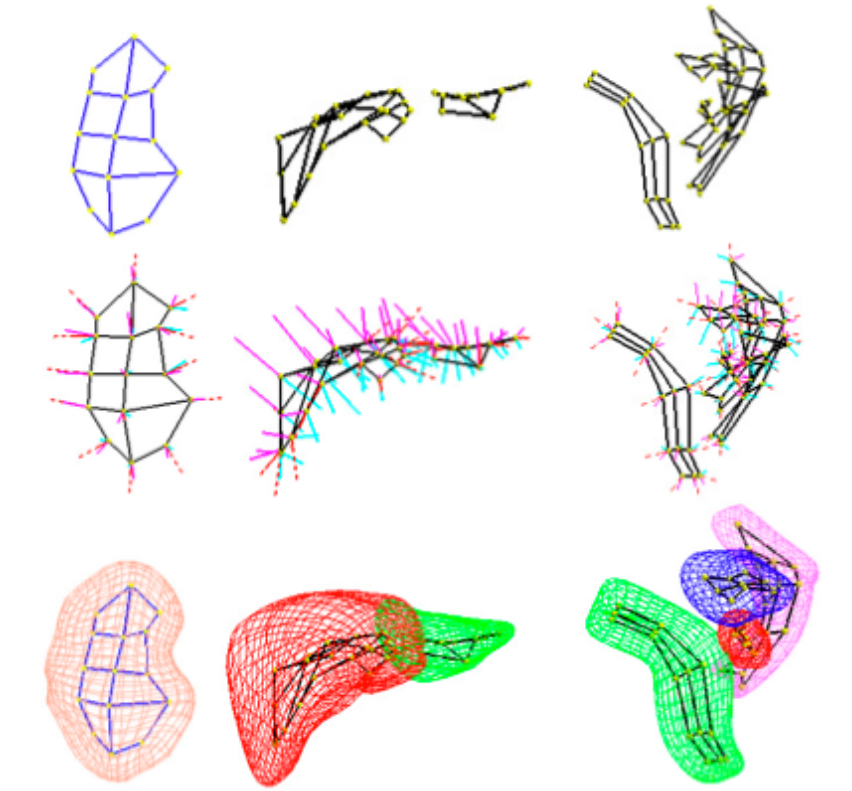


Fig. 5 M-reps for a kidney, a liver, and a male pelvis. Top row: mesh of atom hubs; middle row; mesh of medial atoms (including spokes); bottom row: the implied boundaries shown with atom mesh(es).

3.2 measuring statistical geometric typicality

The geometric typicality that we wish to use is $\log(p(\mathbf{m}))$, or in the case that $\mathbf{m}$ has neighbors $N(\mathbf{m})$, $\log(p(\mathbf{m}|N(\mathbf{m}))$. But except for an additive constant and a constant multiplier of -0.5, when the principal geodesic analysis given in Section 8.7 is used, the log probability density in the symmetric space at any scale level is just a Mahalanobis distance in a tangent space to that symmetric space. Thus, when optimizing in the space of principal geodesics, we are optimizing over the weights a_i of the projections $\mathbf{v}^i$ of the unit-variance principal geodesics onto the feature space tangent plane at the mean. For any value of these a_i, and given the variances σ_i^2 of the principal geodesics in that tangent plane that are derived in the principal geodesic

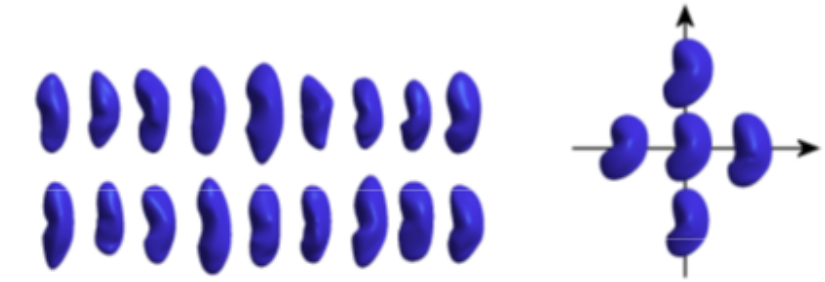


Fig. 6 Left: A subset of our population of training kidneys. Right: the mean of the population and the mean $+/-1$ standard deviation in each of the first two principal geodesic modes.

analysis, the Mahalanobis distance of $-\sum_i a_i^2$ forms the geometric typicality measure.

As discussed earlier, at all scale levels but the global one this geometric typicality metric of the relevant geometric entity needs to reflect its shape properties but also its relation to immediately neighboring peer entities. This can be accomplished with principal geodesics that were computed with augmenting atoms in adjacent objects or figures (see Chapter 8, Section 7).

Two special neighbor relations deserve comment. One is the non-interpenetration relation among very nearby (possibly abutting) objects (see the male pelvis in Fig. 2). Not only the correct relative position, orientation and size need to be reflected in the geometric typicality, but also an interpenetration of the figures needs to result in a low geometric typicality. The second neighbor relation of note is that between a protrusion or indentation *subfigure* to the "host" figure on or into which it sits (see liver in Figs. 2 and 5) or the relation between an object and a nearby, possibly abutting, object. In Chapter 8 we argued that the subfigure should ride on the boundary implied by the host's representation and be known in the figure-relative coordinates of the host. The augmentation idea mentioned as applying to nearby objects uses a similar concept. Thereby we can make measurements of typicality in terms of the position of the subfigure (or related object) relative to the host, the orientation of the subfigure relative to the host, and the size of the subfigure relative to the host. When slight modifications of the hinge atom relationship are created due to motions in symmetric spaces not maintaining the relationship of hinge atoms to their host figure boundary, we find success in simply projecting the hinge atoms back onto the host boundary along host surface normals (interpolated spokes).

4 Training and measuring statistical geometry-to-image match

Methods for training and measuring a probability density on image intensities must do so in a way respecting correspondence of locations across the population. There is much good work on correspondence, e.g., (Davies et al., 2002; Yushkevich et al., 2005), but here we suggest that correspondence be obtained through object-relative coordinates (Fig. 7). For m-reps that means that the figural coordinates provided by $\mathbf{u} = (u, v, \phi, \tau)$ within figures (see Chapter 8, Section 3) and by $\mathbf{u} = (v, w, \phi, \tau)$, within inter-figural blend regions (see Chapter 8, Section 8) provide the means of correspondence. More precisely, intensity statistics are done with respect to $I(\mathbf{u})$.

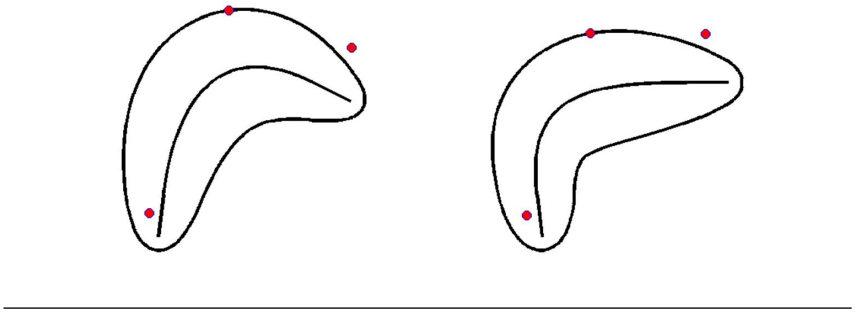


Fig. 7 Correspondence over deformation via figural correspondence. In each pair of corresponding marked points, the two points have the same value of the figural coordinates $\mathbf{u} = (u, v, \phi, \tau)$.

Recall that within an object main figure and within a subfigure outside of the blend region, (u, v) measures relative location along the medial sheet, ϕ expresses which side of the medial sheet the location is or at the end where in the transition between the sides the location is, and τ gives the fraction of the distance along the spoke from the medial end to the boundary end. For interfigural blend regions between a subfigure and a host figure v and ϕ are the cross-figure figural coordinates of the subfigure and $w \in [-1, 1]$ moves along the blend from the curve on the subfigure terminating the blend $(w = -1)$ to the curve on the host figure terminating the blend $(w = +1)$. Section 4.1 describes the computation transforming between Euclidean coordinates $\mathbf{x}$ and figural coordinates $\mathbf{u}$. Between objects one must interpolate between the figural coordinates of the nearby objects. The means of this interpolation is still a subject of research, but one of the options is described in Section 5.1.

We have used two basic methods that go from m-reps and associated greyscale images to geometry-to-image match functions on an image given an m-rep. The method we used first (Stough et al., 2004) was based, like that of the active shape method of Cootes et al. (1993), on normalized correlation between cross boundary intensity profiles and template profiles determined in training. However, in our

method the template in each profile was chosen from among a limited number of possibilities chosen by clustering profiles during training, and values needed in the normalizations of the target profiles at each boundary vertex were also determined during training, thus stabilizing the normalization. Both normalized correlation methods produce a log probability density only under the poor assumption that the profiles are uncorrelated and that the tissue mixture at a voxel in the template can be expected to be precisely the same as that in the corresponding voxel in the target image. To overcome the first weakness Ho (2004) argues for improvements based on multiscale profiles, produced by a variant on Gaussian weighting across but not along the profiles.

Either variant of this profile match method can be expected to achieve less success than our new method, which is designed to produce log probabilities without these faulty assumptions. Our experiments on kidney segmentation, sketched in Section 5.2 and detailed in (Broadhurst et al., 2006), showed the new method to give better results in practice. Thus we describe only the new method, which generates a log likelihood on discrete quantile functions from the intensities in regions relative to the m-rep. It is detailed in Section 4.2.

4.1 Transforming between figural and Euclidean coordinates

The geometry to image match term in the objective function requires object-relative image positions $\mathbf{x}(\mathbf{u})$ to be computed in large number. Thus, interpolation within $I(\mathbf{x})$ must be very efficient. In Pablo at present, this transformation $\mathbf{x}(\mathbf{u})$ is done through the mechanism of subdivision surface methods (Thall, 2004), as described below. Han is developing a more accurate method based on the interpolation of medial atoms (see Chapter 8) and is seeing how to make it adequately speedy.

In the subdivision surface method we interpolate the boundary first and consequently can interpolate medial atoms at any position on the sheet of atoms. The implied boundary is computed from the set of atom spokes connected into quadrilateral and triangular tiles both within figures and in interfigural blend regions (Figs. 8.3 & 8.13). The boundary interpolation is accomplished by a variation of the very efficient Catmull-Clark subdivision (Catmull and Clark, 1978) of the mesh of polygonal tiles. Thall's variation (Thall, 2004) of Catmull-Clark subdivision produces a limit surface that iteratively approaches a surface interpolating in position to spoke ends and with a normal interpolating the respective spokes. That surface is a B-spline at all but finitely many points on the surface. The program gives control of a tolerance on the normal and on the closeness of the interpolations.

The resulting B-spline allows the computation of both boundary positions $\mathbf{b}$ and boundary normals $\mathbf{U}$, which are spoke directions there. Interpolating the medial radius r as well as u and v at such boundary positions allows the computation of $\mathbf{x}(\mathbf{u}) = \mathbf{b} + (\tau\text{-}1)\,\mathbf{U}$.

Points $\mathbf{x}$ can also be given a figural coordinate $\mathbf{u}$ by finding the figural coordinates of the closest medially implied boundary point, using the boundary normal or

the gradient of the distance function as the spoke direction, and calculating τ from the intersection of this spoke with the sheet of hubs. This calculation, however, is fraught with danger, since the boundary may be inadequately smooth.

4.2 Geometry-to-image match via statistics on discrete regional quantile functions

4.2.1 Conceptual basis for statistics on intensity quantile functions

Any efficient geometric description does not capture all there is to say about the biology or physics of the individual being modeled. Thus for a given medially specified object or complex of objects, the variation between different images of the same class of object frequently is due not only to intensity noise but more so to the variation of the materials of which the object are made and of the variation across the object of the weights of those materials making up the materials mixture. Thus in medical images there is variation across patients of the locations of specified tissue types within and between their respective objects. This suggests that point-by-point correspondence as provided, for example, in the active shape models and active appearance models of (Cootes et al., 1993, 1999), where the probability densities are on corresponding intensity values, be replaced by probability densities on regional collections of intensities, ignoring the particular locational correspondences within these regions. In particular, this suggests probabilities on intensity summaries, such as histograms, of regions expected to have uniform mixtures of tissue types.

Our regional intensity summary based match method (Broadhurst et al., 2005; Pizer et al., 2005a) uses a region inside the object and a region outside the object (Fig. 4) and sometimes subregions of these regions defined according to figural geometry.

The feature space formed by using the bin counts of histograms of intensity provides a poor basis for probabilistic analysis. The weakness is exemplified by the fact that the average of two unimodal histograms in this form would be a bimodal histogram, rather than a unimodal histogram whose center is between the two being averaged. In the following we argue that instead representing the regional intensity collection by the curve of intensity values versus quantile (regional intensity quantile function, or *RIQF*) allows an effective log probability density to be calculated by factor analysis. Also, histogram bin counts as features suffer from quantization effects, i.e., binning errors, while discrete RIQFs do not since no arbitrary bin boundaries are selected.

The RIQF of an intensity distribution i can be shown to be the inverse of the cumulative distribution function I of i. Discretely sampling the RIQF yields the *discrete RIQF (DRIQF)*. The DRIQF is an n bin quantile function where each bin j, representing $1/n$ of the probability distribution area, stores its average image intensity i_j. Considering these values in sorted order, the DRIQF for region k can

be represented as a vector $i^k = (i_1^k, \ldots, i_n^k)$. Computing this vector requires partial sorting of the list of N intensities in the region, taking $O(Nlog(n))$ time.

The effectiveness of using standard statistical tools to construct a probability distribution of RIQFs depends on the fact that the space of RIQFs has several known linear properties related to Euclidean distance and thus mean and principal components. Euclidean distance between RIQFs corresponds to the Mallows distance (Mallows, 1972; Levina and Bickel, 2001) between the corresponding probability distributions, defined as follows. For two continuous one-dimensional distributions with cumulative distribution functions Q and R, and RIQFs $q = Q^{-1}$ and $r = R^{-1}$, respectively, the Mallows distance between them is defined as the Minkowski L_2 norm between q and r:

$$M_2(q,r) = \left(\int_0^1 |Q^{-1}(t) - R^{-1}(t)|^2 dt \right)^{1/2} = \left(\int_0^1 |q(t) - r(t)|^2 dt \right)^{1/2}.$$

The Mallows distance can be shown to measure the work required to change one distribution into another by moving probability mass, i.e, the Earth Mover's distance between the corresponding probability distributions, intuitively a good measure of difference between RIQFs. For DRIQFs q and r, the Mallows distance is defined as the L_2 norm of the vector difference between q and r:

$$M_2(q,r) = \left(\frac{1}{n} \sum_{j=1}^{n} |q_j - r_j|^2 \right)^{1/2}.$$

Location and scale changes to any probability distribution, or changes in any affine combination of the DRIQF values, are linear in the space of DRIQFs. Several families of common continuous distributions, including Gaussian, uniform, and exponential distributions, are parameterized by location and scale parameters. Thus, DRIQFs of each of these families of distributions exist in a two-dimensional linear subspace. Also, the Euclidean average (or any linear combination) of a set of DRIQFs from one of these families of distributions results in a DRIQF contained within the family and having means and standard deviations averaging (or correspondingly linearly combining) the respective means and standard deviations. For example, the Mallows distance between two Gaussian distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ is $\sqrt{(\mu_1 - \mu_2)^2 + (\sigma_1 - \sigma_2)^2}$. The average in this space of a set of RIQFs corresponding to Gaussian probability densities is a Gaussian with a mean and standard deviation equal to the average mean and standard deviation of the set of Gaussians. However, a weakness of the space is that for probability distributions composed of a mixture of multiple underlying unimodal distributions, changing the mixture amount is a nonlinear operation.

The consequence of the foregoing is that analysis of regional intensity distributions can be captured by linear statistics on their DQRIFs, which can efficiently capture variation in location and scale change. The method is not effective for dealing with multimodal probability distributions with widely separated peaks and varying

interpeak separations, but our experience is that it works well for unimodal probability distributions and even multimodal probability distributions whose peaks are not widely separated.

DRIQFs of interior and exterior regions of the bladder in 15 CT images are shown in Fig. 8. The first two principal directions of variation of the interior and exterior regions capture 95% and 97% of the variation, respectively. DRIQFs of subregions can also be constructed; this example and this discussion are only in terms of interior and exterior regions. In this example, the contribution of each voxel to the DRIQF is Gaussian weighted by its distance to the surface. This allows narrow regions to be defined that have larger capture ranges and smoother objective functions during segmentation than equivalent non-weighted regions. In each region, gas and bone intensities have been automatically removed from the distribution using a threshold, and the probability of each is independently measured. The Mallows distance is sensitive to the variation in these intensities due to their extreme intensity values compared to the differences in fat and tissue intensities.

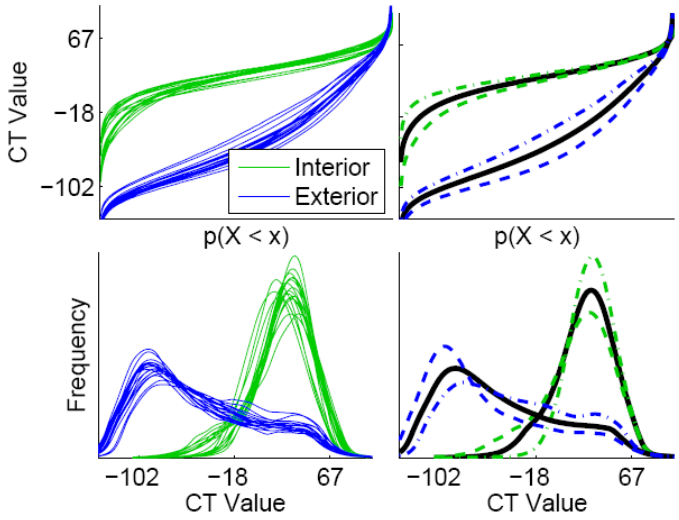


Fig. 8 Bladder DRIQFs (top) and corresponding histograms (bottom). Left: training samples; right: learned mean and ± 2 standard deviations along the first principal direction.

4.2.2 Training probability densities on regional intensity quantile functions

The probability densities on DRIQFs that we use are estimated by principal component analysis of the DRIQFs, taken as feature vectors. Then the geometry-to-image

match of the DRIQF obtained from a target image region is a Mahalanobis distance based on this principal component analysis. In the following we detail the estimation of this Mahalanobis distance function.

We model the variation in our DRIQF feature space as a multivariate Gaussian distribution. The dimension of this space is equal to the number of bins used to represent a DRIQF, which is typically 200. This often results in a high dimension, low sample size situation, which prevents us from estimating a full rank multivariate Gaussian model. Therefore, we use principal component analysis to estimate a low dimensional subspace, typically of dimension 2-4, in which we build a Gaussian model. We then measure the expected distance to this subspace by summing the remaining eigenvalues since during segmentation we expect to estimate the probability of many image regions that are not typical of the training regions. Thus, the final Gaussian model for each region is of dimension equal to the number of eigenmodes plus one.

The geometry-to-image match function is the resulting Mahalanobis distance function. Intuitively, the Mahalanobis distance of a target DRIQF is equal to the Mallows distance between probability distribution corresponding to the target DRIQF and that corresponding to the mean RIQF, modified by the standard deviations in each direction of the Gaussian model. Thus the Mahalanobis distance is a natural enhancement of the Mallows distance that accounts for the variability in the training set.

The training data on which the principal component analysis is done is formed as follows. For each training case we have a greyscale image, a binary image, and an m-rep fit to the binary image as discussed in Section 3.1. Voxel correspondences specified by m-rep based figural coordinates (Section 4.1) allows us to compute the set of DRIQFs for each object-relative image region across the training images. When determining if a voxel belongs in a region, we initially use the binary image, not the m-rep, to label voxels as being inside or outside the object. This allows us to define mean DRIQFs that correctly provide references for the Mahalanobis distances used to form the geometry-to-image match. These DRIQFs do not, however, measure the expected variation of the actual training segmentations. Therefore, we estimate the covariance of the DRIQF in each region from the DRIQF values based on m-rep region labeling minus the already computed respective mean DRIQF, which is based on binary labeling.

5 Pablo Details and Results

5.1 The Voxel-Scale Stage of Segmentation

After all of the stages of segmentation that modify the medial atoms, an m-rep has undergone transformation from the beginning model (typically the mean of the global object complex or object). Figural coordinates allow this transformation to

be understood as a diffeomorphism within all of the objects making up the complex represented by the m-rep. This warp can be interpolated into a chosen portion of 3-space including the complex, including the interstitial space between multiple objects or figures, if the complex is made up of more than one figure. A further finer scale transformation on that portion of 3-space can then be determined.

We interpolate the transformation from the objects to the surrounding 3-space, as follows. Each implied boundary position of the m-rep is understood as the tip of a particular m-rep spoke, either one of the basic representation or one interpolated from it. That spoke is from a medial atom $\mathbf{m}(1)$ at particular figural coordinates that allow it to be associated with a corresponding atom $\mathbf{m}(0)$ in the original m-rep model. Paths $\mathbf{m}(t)$, $0 \leq t \leq 1$, in the abstract space of atoms between the original atoms and the corresponding transformed atoms can be found according the mathematics in Chapter 3, Section 3.3, such that at every t the m-rep is unfolded and thus the continuous transformation of m-rep interiors is diffeomorphic. These paths can be sampled in t to produce a path of the corresponding spoke ends, and this sequence of positions can be used as a boundary condition in a landmark interpolation method. For example, one can use the thin-plate spline interpolation (Bookstein, 1991) on each of the corresponding successive pieces of the paths of all of the spoke ends. If the interstitial transformation was not diffeomorphic, as when objects slid along each other between individuals, an interpolation that allowed such transformations would need to be used.

We determine the fine scale warp to be composed with the transformation interpolated from the medial transformation using the fluid-flow warp method of (Miller et al., 1999). If the final map might not be diffeomorphic, as when regions of gas formed or were lost in the rectum or when tumors existed in the particular patient but the model was of well patients, then a warp method that permitted such situations would need to be applied.

The approach of computing a small scale space warp to be composed with a medially determined warp has the following advantages over computing the whole deformation as a space warp from an atlas. Optimizing large scale deformations is likely to be heavily affected by local minima, and in any case it is very likely to be slow as result of having to work over the combinatorially related, many small voxels. Indeed methods that have attempted to compute such warps have found it necessary to begin the process by preceding the voxel-scale warp by applying larger scale transformations such as ones based on manually chosen landmarks (Christensen et al., 1997). Using medial transformations to provide the large scale warp has advantages of being automatic, of using object-based correspondences, and of dividing itself into multiple scales, e.g., global to the object complex, object by object (with sympathetic inter-object relations), figure by figure (with sympathetic inter-figure relations), and medial atom by medial atom (reflecting inter-atom relations). Using these many scale levels produces both a much improved likelihood of convergence to the global optimum and qualitatively improved speed.

5.2 Evaluation of Segmentations

We have applied Pablo anecdotally to the segmentation of variety of organs or organ complexes. M-rep models have been built for both the liver (Han et al., 2005), a multifigure object, and the heart (Pilgram et al., 2003), a multi-object ensemble, and statistical description of these anatomic entities have been generated. Controlled evaluations, described in the next two sections, have been carried out for the following two situations: 1) the extraction of kidneys from new patients' CT scans;
2) the extraction of the bladder, prostate, rectum complex from CT scans of a patient on one day of radiation treatment given the CT scans and segmentations of that patient on the planning day and other days of treatment. The first of these involves the segmentation of a single single-figure object with statistics drawn from many other patients' images, so we refer to it as an *inter-patient segmentation*. The second involves the segmentation of a multi-object complex with the statistics describing the variation across days within a patient (*intra-patient*).

5.2.1 Inter-patient Kidney Segmentation Results

We have studied segmentation of the kidney from CT scans over a few years. An early result of evaluation of an earlier form of Pablo was described in (Rao et al., 2005). In that study we determined that averaged over a particular test sample, two humans' manual segmentations differed from each other in average surface distance over the kidney by 1.2mm. Averaged over these cases the Hausdorff distance between the two segmented kidneys was 10mm.

In a controlled study on segmentation of kidneys from 3D CT images of clinical quality, we used the sum of Mahalanobis distances described in Sections 3 and 4.2 as the objective function at the figure (object) stage. Since the log probability densities relieves the necessity of setting the relative weights of the two terms of the objective function by user control, these weights were set to unity in the study. However, at the atom stage we used the average squared-distance between each atom and the geodesic average of its neighbors, i.e, the atom irregularity penalty used in Binary Pablo (Section 3.1) for the geometric typicality, since the probability density training for the atom stage was not yet ready. This required a manually set weight on this term, which was held fixed for the experiment. The DRIQFs used in the geometry-to-image match at both stages were from Gaussian weighted regions inside and outside the kidney that had σ = 3mm. In one trial training was on 20 cases and testing was done on 19 cases. In another trial leave-one-out testing was applied, i.e., all tests with 38 training cases and 1 test case were evaluated. In the geometry-to-image match, principal components carrying 97% and 99% of the variance were used to form the inside-object and outside-object log probability densities, respectively.

For our evaluation, we first consider the segmentation result to be that leading to minimum values of the atom-stage objective function. On the 19 test CTs the segmented kidneys had average surface distances to one human segmentation that

was at least as good as found between humans in the Rao study on a different test set. More precisely, the computer vs. human segmentations differed from each other in average surface distance over the kidney by 1.2mm on the average case, and the Hausdorff distance between the two segmented kidneys was 6.8mm on the average case. In all of the test cases, the automatic segmentation was usable without editing in radiotherapy treatment planning, although a voxel-stage editing would have been considered desirable in many of the cases. In fact, the automatic segmentations are frequently judged to be superior to the manual segmentations, and they have the additional benefit of being reproducible, even if the initialization is slightly different.

In the leave-one-out experiment, with its larger training sets, the results were roughly equivalent. The results of both experiments are given in more detail in (Broadhurst et al., 2006). An atom stage with a probabilistic geometric typicality might be expected to yield a further improvement.

These results were made less impressive by the fact that the objective function optimum that was found was not always achieved when we used the initialization based on six landmarks that we anticipated using for clinical purposes, namely, the north and south poles of the kidney and the two kidney crests at the level of the renal pelvis, and two positions at that level centered between the two crest landmarks. However, only 2 of the 19 cases in the first experiment and 4 of the 39 cases in the leave-one-out experiment would have required editing for clinical purposes.

5.2.2 Intra-patient Multi-object Male Pelvis Segmentation Results

As illustrated in Fig. 9, we have built a model for the multi-object ensemble of bladder, prostate, and rectum in the male pelvis. We have fit this multi-object model into a few dozen binary segmentations of these organs from fraction-by-fraction[2] CT images in five patients' cases, and after alignment of the prostate based on the two landmarks of the urethral entrance into and exit from the prostate, and after alignment of the bladder based on two polar and four equatorial landmarks, we have built statistical descriptions of the variability of the these objects across fractions within each particular patient. As well, we have built DRIQF statistics as described in the previous section, but here for 6 regions: interior and exterior regions for each organ. For the prostate and for the bladder, we also evaluated the use of approximately 200 overlapping regions to produce the exterior DRIQFs.

Finally, we have used these statistics to segment the prostate and the bladder from the CT images in other fractions in a leave-one-out experiment. (The rectum was represented as a tubular m-rep, and successful segmentation of the rectum was done in a separate experiment.) The initialization was done using the aforementioned landmarks. Since we are nearly ready to apply our method of principal geodesic analysis on medial atom residues and factor analysis on DRIQFS in local atom-relative object regions, we have optimized at the object stage only. The best results are produced when using the 200 exterior regions for DRIQFs. These results show

[2] A fraction is the radiation treatment on a given day.

segmentations that have a median, over the cases, of the intersection/average volume overlap to a human segmentation of 93% for both the bladder and prostate and a median, over the cases, of the average closest point distance to the human segmentation of 1.13 mm for the bladder and 0.99 mm for the prostate. The numbers for the prostate, comparing segmentation based on statistics from a human who produced the training manual segmentations to the that human's result in the left-out-case, should be compared to the numbers comparing another observer's manual segmentation of the prostate to that of the training observer in one of the five patients' set of 16 multi-day CTs (Foskey et al., 2005). The agreement of the two humans' segmentations was 81% volume overlap and 1.9 mm average closest point surface separation.

When our segmentation was not as good as we wished, there were two explanations. First, in many of the segmentations of the bladder, a smaller scale refinement was necessary. We expect this to be accomplished when the log posterior optimizing atom stage is applied. Second, in a few cases the bladder initialization based on prostate landmarks was not adequate, but with a bladder-based initialization the segmentation was improved in a majority of cases.

This multi-object segmentation has been adapted for the clinical situation of adaptive radiotherapy by training the object principal geodesics by a pooling of aligned deviations from the mean of other patients. The results, which will soon be published with the details of the method, are comparable to those reported above. Also, we expect shortly to report results of the atom-stage refinements of these segmentations.

Moreover, we are presently investigating having each object's change at the object scale level be divided into an m-rep change $\Delta\mathbf{m}^{self}$ independent of neighboring objects and an m-rep change $\Delta\mathbf{m}^{ngbr}$ reflecting the effect on the object of changes in neighbor atoms (Jeong et al., 2006). The neighbor-induced change is statistically described using the method of augmented object descriptions and prediction described in Chapter 8, Section 8. $\Delta\mathbf{m}^{ngbr}$ is decomposed as a conditional mean of the object, given designated neighbor atoms in its neighbors, plus a neighbor-effect residue with its own probability density. Probability densities on $\Delta\mathbf{m}^{self}$, on the augmented object, and on the neighbor effect residue are estimated by repetition of successive principal geodesic analyses. In segmentation the posterior is successively optimized with the prior iteratively in succession being the *self* probability density and the *neighbor residue* probability density, respectively. Initial results from statistical analysis on the bladder, prostate, rectum object complex are biologically reasonable, but it remains to test this approach by segmentations that use the self and neighbor residue probability densities.

5.2.3 Speed of Computation

The speed of a 3D segmentation on a Pentium IV, 1.7GHz computer subdivides as follows.

- Preprocessing computations take less than 1 second.

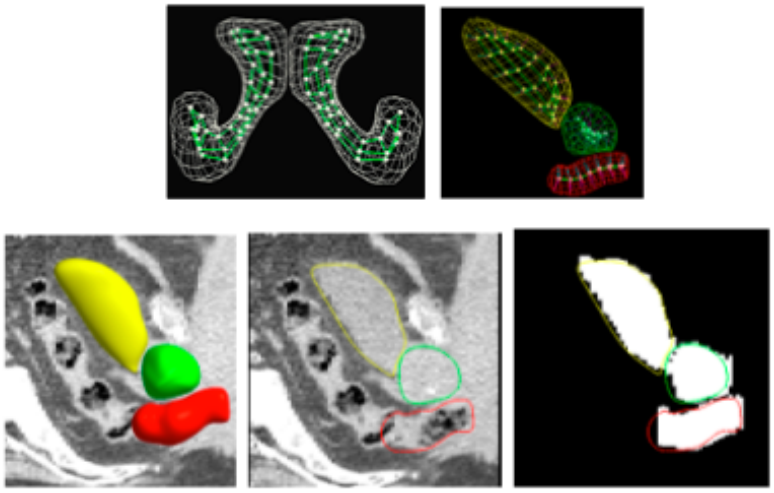


Fig. 9 M-reps for segmenting the male pelvis in CT images in later radiotherapy fractions. Top left: m-rep for pubic bones, used to register the later day fraction images with the first day fraction. Top-right: the m-rep for the bladder, prostate and rectum. Bottom left: a visualization of the bladder, prostate rectum m-rep's implied boundaries relative to a slice of the associated 3D CT image. Bottom middle and right: the segmentation result in a later fraction, shown in one of the image slices first vs. the greyscale CT image and then vs. the human segmentation shown in white.

- The largest scale stage (the object complex stage for an ensemble, the object stage for a single multifigure object, the figure stage for a single-figure object) takes a about 5 seconds per iteration and on the average requiring 20 iterations for a total time of about 3 minutes to determine the geometric warp coefficients.
- When the smaller scale medial stages are appropriately re-programmed, the same numbers will apply to each subfigure stage and about double for a full pass through the atoms at the atom stage, modulo the number of iterations required.
- The voxel displacement stage has not been timed, but it is expected to operate in under a minute.

Thus the total time for a kidney segmentation will typically be 7 minutes to segment a single-figure object.

While the method's speed has already benefited strongly from moving much of the computation from the deformation stage to the model building stage, there is still much room for speedup of integer multiples by more medial levels of coarse to fine, by medial deformation measured directly from the atoms without resort to the implied boundary, by having the gradients of the objective function relative to the changing parameters needed by the optimization steps be computed analytically

rather than with numerical derivatives (shown in initial tests to more than double the speed), and just by more careful coding.

6 Hypothesis Testing for Localized Shape Differences between Groups

We now focus on the quantitative morphologic assessment of structures between groups of human subjects. Our examples are individual brain structures in neuroimaging. Conventional methods study only volumetric changes, which explain intuitively global atrophy or dilation of structures. On the other hand, structural changes at specific locations are not sufficiently reflected in volume measurements. Statistical shape difference testing has thus become of increasing interest. Its potential to precisely locate morphological changes and to discriminate between groups makes it a good choice for studying pathological morphologic processes due to disease, as well as neuro-developmental processes. For example, we may wish to understand the shape differences in the hippocampus, caudate nucleus, cerebral ventricle complex in the brain between control patients and schizophrenics, or we may be interested in the differences of the hippocampus between 2-year olds who will develop autism and 4-year olds who will develop autism.

We focus in this section on the *hypothesis testing* of *whether* and *where* there are m-rep shape differences between the groups. We will discuss both global tests and truly local tests. Hypothesis testing applications using other medial descriptions have been proposed by Golland et al. (1999) and Bouix et al. (2005).

We call the group designator C, which is numbered from 1 to the number of studied groups. Each group C_i consists of the objects of a sample of N_i cases. We assume that the objects or object complexes have been aligned across all cases, with the same alignment applied for the cases in both classes. The discrete m-rep objects are described as a tuple of medial atoms. The first idea is either to take all the atoms together and do a global test by studying the multivariate tuple of atoms $\times$ the 8 or 9 parameters per atom. Such a test can be powerful but will fail to localize the differences found to a particular collection of locations (i.e., parameters).

The alternative is to do a local (for a particular parameter of a particular atom) test on each atom parameter, at each position. We will use the term *location* to refer to such a combination of parameter and atom. The first idea might be to design a statistical test separately on any such location, and then to repeat that test over all atoms $\times$ parameters. However, the atoms are all correlated, and the parameter values are all correlated. To avoid unintended looseness in the threshold for rejecting the null hypothesis for any parameter, the threshold for rejection has to be adjusted for each parameter in a way reflecting the correlations. In Section 6.1.2 we describe a non-parametric permutation method to deal with this problem for individual parameters.

Section 6.2 will then focus on testing the full m-rep atom parameters jointly in symmetric space at a fixed scale. Finally, Section 6.2.2 will discuss why even atom

by atom testing is not adequately local to the regions determined by the atoms and how to more appropriately handle locality.

6.1 Tests in Euclidean space

6.1.1 Univariate Tests in Euclidean space

We may test a particular location (see Fig. 10). Here we focus on the two parameters, local position and radius (Styner et al., 2003b, 2004) of a particular atom. We first compute the group average objects by averaging the position and radius for each medial atom across each group. The overall average location is then computed as the average over all group average locations and serves as a template for computing univariate shape distance measurements. The signed position and thickness differences to the template are computed separately for the specified atom. The sign of the position difference is computed using the direction of the template medial surface normals.

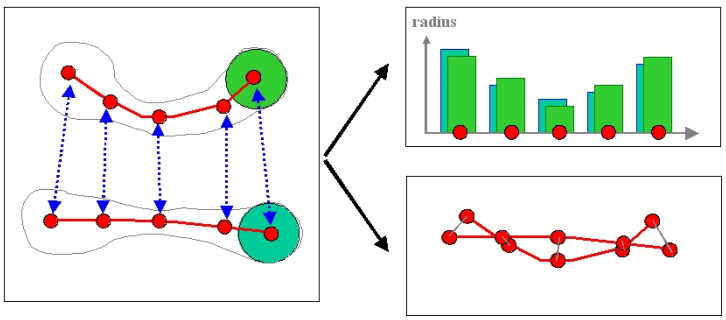


Fig. 10 Scalar m-rep shape difference (schematically in 2D) of 2 m-rep objects: Differences in the radius (top graph) and position (lower graph) are studied separately in Euclidean space. The properties express different kinds of underlying processes (growth vs. deformation) that are assumed to be statistically independent in the scalar testing.

Global shape analysis is computed by analyzing summarizing features such as the mean, median or other quantile measurements of the local differences across each object by standard statistical hypothesis tests. The choice of the feature evidently influences the outcome of the tests. The statistical tests mainly include parametric mean difference tests based on the Students-*t* distribution, and non-parametric mean difference tests, as well as parametric analysis of variance tests (ANOVA).

Local shape analysis does not need a summarizing feature as it is a truly local test. It is computed by testing each medial atom independently with a standard statistical hypothesis test. This results in a significance value (*P*-value) for each parameter and medial atom. We can represent this significance in a 3D visualization using color and size of spheres at the atom positions of the overall average object. This visualization, called a medial statistical significance map, allows one to locate significant shape differences between the groups in an intuitive but not truly local fashion (see Section 6.2.2). However, this raw significance map is incorrectly optimistic in regard to false-positive error rate because the atoms as well as the individual parameter values of a single atom are correlated, leading to the *multiple comparison problem*, a topic of active research in the field of shape analysis (Worsley et al., 1996; Nichols and Holmes, 2001).

The raw significance map can be corrected for this multiple comparison problem using a uniformly sensitive, non-parametric permutation test approach (Pantazis et al., 2004) described in the next section. This results in a corrected significance map. In contrast to the raw significance map, which is a quite optimistic estimate of the real significance map, the corrected significance map is a somewhat pessimistic estimate, as discussed in the next section.

6.1.2 Multivariate Permutation tests in Euclidean spaces

The permutation tests we describe here localize regions (atom indices or parameters thereof) in objects that exhibit statistically significant morphological variation among two population groups while controlling the risk of any false positives, as long as the object features are in a Euclidean space. We find local thresholds that control the false-positive error rate and at the same time achieve uniform sensitivity among all locations.

We assume we have two groups of local parameter sets, group A and group B. Each parameter set represents either shape parameters or differences of shape parameters. We want to test the two groups for difference in the means at each location. Permutation tests are a valid and tractable approach for such an application. Our null hypothesis is that the distribution of the parameter set at each element is the same for every subject regardless of the group. Permutations among the two groups satisfy the exchangeability condition, i.e., they leave the distribution of the statistic of interest unaltered under the null hypothesis. Given n_1 members of the first group a_k, $k = 1 \ldots n_1$ and n_2 members of the second group b_k, $k = 1 \ldots n_2$, we can create $M \leq \binom{n_1+n_2}{n_2}$ permutation samples. A value of M from 20000 and up should yield results that are negligibly different from using all permutations (Edgington, 1995).

The complex set of steps needed to test the null hypothesis that the two groups have the same probability distributions is illustrated in Fig. 11. We take the reader through this process step by step. For each permutation sample j, we compute a difference metric $\boldsymbol{T}_j$ between the groups, with elements T_{ij}. For univariate Euclidean parameters the absolute distance between the means of the groups is often used:

$$T_{ij} = \left|\hat{\mu}_{aij} - \hat{\mu}_{bij}\right| \tag{1}$$

where i is the location index, j the permutation index. If we wish to sense locations at which differences of collections of parameters at the locations are signficant, we can use difference metrics for multivariate, Euclidean or non-Euclidean parameters, as long as the difference metric itself is in Euclidean space, such as the multivariate Hotelling T^2 test statistic for the collection: $T^2 \propto D^2 = (\hat{\mu}_a - \hat{\mu}_b)^T \hat{\Sigma}^{-1}(\hat{\mu}_a - \hat{\mu}_b)$, where $\hat{\Sigma}$ is the pooled sample covariance. In $\mathbb{R}^n$ this statistic is invariant to coordinate transformations and is uniformly the most powerful test with this property (see Anderson (1958) for a derivation). We cannot use this statistic directly on the multivariate combination of all atoms and parameters due to its inability to provide sensing of location.

In Fig. 11 it is assumed that we are given a target division of the data into two groups, A and B. To achieve uniform sensitivity across all locations, the parameter (or group of parameters) value T_{ij} at each location is first transformed to a uniformly distributed probability density value on [0,1], making all locations comparable. This is applied both to the test grouping, producing T^p_{i0}, and as illustrated in the bottom row of Fig., 11, it is also applied to the random permutations derived from the union of groups A and B, producing the T^p_{ij}. We can compare T^p_{ij} for each parameter i within each permutation j to produce a conservative summary statistic S_j for each permutation. Across the permutations the distribution of this summary statistic produces a common threshold S^{th} for each of the respective probability-normalized local parameters T^p_{i0}, as illustrated in the top row. The justification and specification of this scheme now follows.

The conservative summary statistic that we use for each permutation is the smallest probability density value over all locations i. We may then use the empirical distribution of this conservative summary statistic to extract thresholds that control the false positives to a desired level.

This method depends on having a form of normalization in the statistic T_{ij} that makes the locations comparable. A suitable normalization scheme is based on computing p-values, i.e., at each spatial location we compute the empirical distribution across permutations and then replace the statistic T_{ij} for each permutation sample with its p-value T^p_{ij}. The normalized metric T^p_{ij} is then guaranteed to have a uniform distribution on [0,1] under H_o for each i.

We can use the normalized data to define a local threshold map that controls the false-positive error rate to a desired level, say $\alpha = 5\%$, when applied to the original data. If the conservative summary statistic of the local parameters is $S_j = min_i\{T^p_{ij}\}$ over all locations i and $\hat{F}_S$ is the empirical cumulative distribution function of S, the appropriate global thresholds for a level α test would be $\hat{F}_S^{-1}(\alpha)$. For example, if we choose a threshold that leaves 5% of the area of the empirical distribution on the left side of S_j, then we have 5% probability of one or more false positives across all locations. This threshold S^{th} can be directly applied to T^p_{i0} (the statistic formed by probability-normalizing the original data with permutation index $j = 0$). Since the statistic T_{ij} is normalized separately for each location i, the same S^{th} corresponds

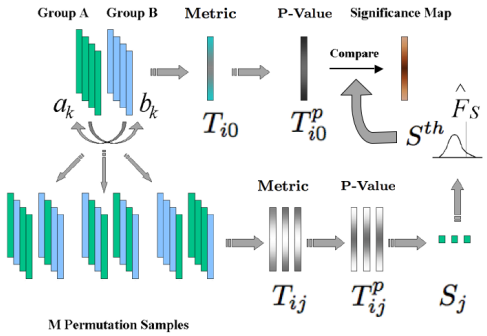


Fig. 11 Illustration of the permutation scheme. In the bottom row we create M permutation samples from the original data. We let j index the permutations and let i index the locations. For each permutation and location we compute the group difference metric T_{ij}, which is probability-normalized into T^p_{ij}. The data is then summarized across all locations to create the conservative summary statistic S_j over all locations. The empirical distribution of S_j, called $\hat{F}_S$ is used to define a global threshold S^{th} which for each location is applied to the probability-normalized test statistic obtained from the division to be tested, into groups A and B.

to different values of local thresholds $p_i^{-1}(S^{th})$ of the unnormalized statistic T_{i0} at different locations.

Due to the use of the minimum p-value statistic across the whole surface, this correction scheme is focused only on controlling the rate of false positives at the given level α (commonly $\alpha = 0.05$) across the surface. No similar control of the rate of false negatives is available with this scheme. As the local significance level correctly controlled for false negatives can be anywhere between the raw p-value and the p-value corrected with our scheme, this corrected significance map is a pessimistic estimate of the true significance map.

6.2 Tests in Symmetric Spaces

The ideas in the previous section must be generalized to the non-Euclidean feature spaces appearing in m-reps and their symmetric space. We can derive permutation tests for equality of means of two groups using elements of the symmetric space. The sample means of each group, $\hat{\mu}_a$ and $\hat{\mu}_b$, can be computed using the techniques described in Chapter 8. Replacing T_{ij} from (1) with

$$T_{ij} = d(\hat{\mu}^*_a, \hat{\mu}^*_b) \tag{2}$$

yields a natural extension of local tests to symmetric spaces.

This provides a way to produce tests for a single aspect of the m-reps, such as position or radius of a particular atom, independently of the others, but typically we require a multivariate test using all of the parameters of one or more atoms simultaneously. We cannot fall back on Hotelling's T^2 test because it applies only to the linear case. Instead we can apply a transformation that forms new features from marginal probabilities, handling differing degrees of variability or correlation and making the test independent of magnification.

6.2.1 Global Multivariate Permutation Tests in Symmetric Spaces

We must now generalize the desirable properties of Hotelling's T^2 test to a nonparametric, nonlinear setting. One seemingly attractive option is to perform statistics in the tangent plane as is done with principal geodesic analysis, since its linearity means Hotelling's T^2 test can be applied directly. However, with two samples, the question that arises is *which* tangent plane, since there is a different one around each sample's mean, and there may be no unique map between them.

The other conceptual problem is that if one follows geodesics past the *cut locus*—the set of points where two or more geodesics cross—then points on the manifold no longer have a single well-defined representative in the tangent plane. Instead of addressing these problems, we take a more general approach, which only requires that our objects lie in a metric space.

Our approach is based upon a general framework for nonparametric combination introduced by Pesarin (2001). The general idea is to perform a set of partial tests, each on a different aspect of the data, and then combine them into a single summary statistic, taking into account the dependence between the variables and the true multivariate nature of the data. When performing the partial tests, we require that each distribution has the same structure around the mean—equivalent to the assumption of a common covariance required by Hotelling—and test for a difference of means. More precisely, following the idea described in the previous section, we map each feature to its marginal probability and use these probability values as features.

The following two sections describe the details.

Partial Tests. We compute test statistics T_{ij} as before, where as before i indexes the model parameters and j is the permutation index. We now turn to the case where we have Q test statistics: one for each of the parameters in our shape model. Let $\mu_{a,i}$ and $\mu_{b,i}$ be the means of the ith model parameter for each group. Then we wish to test whether any hypothesis $H_{1,i} : \{\mu_{a,i} \neq \mu_{b,i}\}$ is true against the alternative, that each null hypothesis $H_{0,i} : \{\mu_{a,i} = \mu_{b,i}\}$ is true. The partial test statistics $T_{ij}, i \in 1 \ldots Q, j \in 1 \ldots M$ are defined analogously to (2).

It can be shown that each of our mapped features T_{ij}^p has the properties of being significant for large values, consistent, and marginally unbiased, as defined in (Pesarin, 2001). Given that, Pesarin shows that a suitable combining function

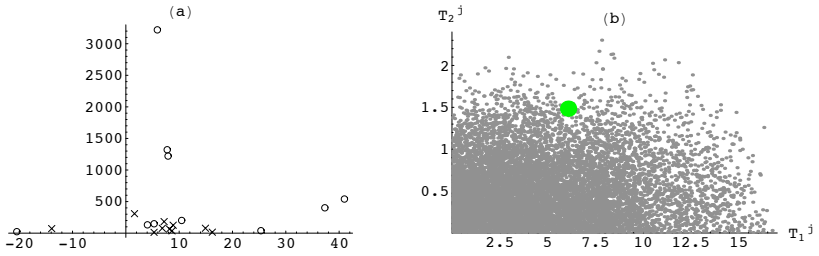


Fig. 12 The observed data and test statistics for our simple example. (a) shows the distribution of our two samples, with $\times$'s for the first and $\circ$'s for the second. (b) shows the distribution of the partial test statistics under permutation. The large dot indicates the location of the observed data point.

(described in the next section) will produce an unbiased test for the global hypothesis H_0 against H_1.

Since each of our tests are restricted to the data from a single model parameter and we have assumed that the distributions around the means in each group are identical, they are marginally unbiased. We cannot add an explicit test for equality of the distributions about the mean, as then the test for equality of means would be biased on its outcome.

To illustrate these ideas, we present a simple example, which we will follow through the next few sections. We take two samples of $n_1 = n_2 = 10$ data points from the two-dimensional space $\mathbb{R} \times \mathbb{R}^+$, corresponding to a position and a scale parameter. The samples are taken from a multivariate normal distribution by exponentiating the second coordinate, and then scaling both coordinates by a factor of ten. They are plotted together in Fig. 12a. They have a common covariance (before the exponentiation), and the two means are slightly offset in the second coordinate.

We construct $Q = 2$ partial test statistics using (2) for each coordinate, and evaluate them using Monte Carlo simulation with $M = 10,000$ permutations.

The results are shown in Fig. 12b. The first partial test value lies in the middle of the distribution, while the second lies near the edge. However, the scale of the first test is much larger, because no logarithm is involved in its metric.

Multivariate Combination. Given the partial tests from the previous section, we wish to combine them into a single test, while preserving the underlying dependence relations between the tests. This is done in the following manner. We apply the same M permutations to the data when computing each of partial tests, and we then compute a p-value statistic, T_{ij}^p as in Section 6.1.2. It is critical to use the same permutations for each partial test, as this is what captures the nature of the joint distribution.

We now wish to design a combining function to produce a single summary statistic, T'_j, from each p-value vector $\boldsymbol{T}_j^p$. For one-sided tests, this statistic must be monotonically non-increasing in each argument, must obtain its (possibly infinite) supre-

mum when any p-value is zero, and the critical value S^{th} must be finite and strictly smaller than the supremum. If these conditions are satisfied along with those on the partial tests from the previous section, T_j' will be an unbiased test for the global hypothesis H_0 against H_1 (Pesarin, 2001).

Our combining function is motivated by the two-sided case (with signed distances), where we can use the Mahalanobis distance. Thus we need to transform the uniformly distributed p-values to a random variable that is normally distributed with mean zero and standard deviation 1. This is straightforwardly accomplished by applying the inverse of the cumulative density function for that Gaussian after subtracting $\frac{1}{2M}$. The extra $\frac{1}{2M}$ term keeps the values finite when the p-value is 1, and it is negligible as M goes to infinity. That is, we compute a $\boldsymbol{U}_j$ vector for each permutation, where $U_{ij} = \Phi^{-1}(T_{ij}^p - \frac{1}{2M})$, $j \in 1 \ldots M$, and Φ is the cumulative distribution function for the standard normal distribution. The map via the p-values and the Φ function gives the statistics known distributions that are directly comparable.

Arranging the $\boldsymbol{U}_j$ vectors into a single $M \times p$ matrix $\mathbf{U}$, we can estimate the covariance matrix $\hat{\Sigma}_U = \frac{1}{M}\mathbf{U}^T\mathbf{U}$ and use the Mahalanobis statistic: $T_j' = (\boldsymbol{U}_j)^T \hat{\Sigma}_U^{-1} \boldsymbol{U}_j$. In the event that the data really is linear and normally distributed, the $\hat{\Sigma}_U$ matrix converges to the true covariance as the sample size goes to infinity (Pallini and Pesarin, 1992), making it asymptotically equivalent to Hotelling's T^2 test. Even if the sample size is small, the matrix Σ_U is well-conditioned regardless of the number of variables, since it is the covariance over the M permutations.

Typically, our distances are not signed, so we are interested in a one-sided test. In this case, we use the positive half of the standard normal cumulative distance function, $U_{ij} = \Phi^{-1}(1 - \frac{1}{2}(T_{ij}^p - \frac{1}{2M}))$ and assume the $\boldsymbol{U}_j$ distribution is symmetric about the origin. This assumption, however, implies that the covariance between $U_{i_1 j}$ and $U_{i_2 j}$ is exactly zero when $i_1 \neq i_2$. The diagonal entries of $\hat{\Sigma}_U$ are 1 by construction, so $\hat{\Sigma}_U = I$, the identity matrix. The fact that the p-values of the partial tests are invariant to scale obviates the need for arbitrary scaling factors. Thus, our one-sided combining function is

$$T_j' = (\boldsymbol{U}_j)^T \cdot \boldsymbol{U}_j \ . \tag{3}$$

The normality of the partial test statistics is not required. Also, even though the marginal distributions of the $\boldsymbol{U}_j$ vectors are normal, the joint distribution may not be. Therefore, we must use the empirical distribution of T_j' in order to compute the final p-value of the global test: $T_0'^p$. It is this nonparametric approach that corrects for correlation among the tests, even without explicit diagonal entries in the covariance matrix.

We return to our example from the previous section. The $\boldsymbol{U}_j$ vectors are plotted in Fig. 13a, along with the $\alpha = 0.05$ decision boundary, and our sample is shown to lie outside of it. As can be seen, equal power is assigned to alternatives lying at the same distance from the origin in this space. Fig. 13b shows this boundary mapped back into the space of the original p-values.

The entire procedure is very similar to procedures used in correction for multiple tests described in the previous sections. However, instead of trying to find a local

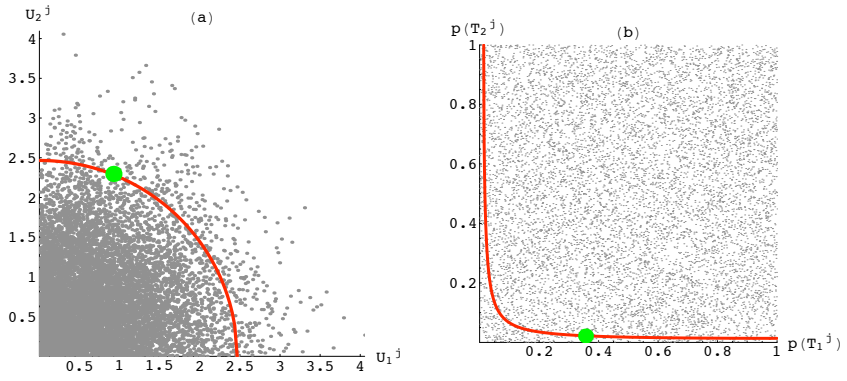


Fig. 13 The empirical distribution of our example plotted against the decision boundary at $\alpha = 0.05$. (a) The distribution of the $\boldsymbol{U}_j$ vectors, where the cutoff is a circle centered around the origin. (b) The distribution of the original p-values with the decision boundary pulled back into this space. The observed value is shown as the large dot in both plots.

threshold for each test individually, we carve out a region of the multivariate T_{ij}^p space that contains some particular fraction, e.g., 5%, of the data to label as significant. We lose the ability to say *which* test is significant but gain power in the cases where multiple statistics independently signal significant differences.

6.2.2 Local Multivariate Tests in Symmetric Spaces

A test on all of the geometric primitives (e.g., medial atom) taken together is not truly a large scale test, for it confuses correlation with spatial scale. A test on each geometric primitive is not truly a small scale test, for it will respond equally well to a variation with large spatial scale as to one with a small scale. The Markov assumption on geometric neighbors allows the separation of scales by removing the correlation of neighboring elements from an element. In particular, if we can estimate the best predictor of a primitive by its neighbors and subtract that predictor from the primitive, the resulting residue provides the entity to test for significant variation *at the specified locality*.

This idea can be used for primitives such as objects or figures, but we are presently working to test it at the scale of the medial atom. Using the ideas in Section 1, the hypothesis testing would thus be done on each geodesic difference of the interpoland from the atom. However, we are still working on this form of local test, so the following section simply tests the atoms, one by one, rather than their residues.

7 Applications of Hypothesis Testing to Brain Structure Shape Differences in Neuro-Imaging

This section presents two application of hypothesis testing of m-rep objects. In the first application, scalar hypothesis testing of individual medial parameters was employed (see Section 6.1.2) for analyzing hippocampal shape in schizophrenia. In the second application, hypothesis testing in the symmetric space (see Section 6.2) was employed for analyzing ventricular shape in healthy twins and in schizophrenia.

7.1 Hippocampus study in Schizophrenia

In the study presented in this section, we investigated the shape of the hippocampus structure in the left and right brain hemisphere in schizophrenic patients (SZ, 56 cases) and healthy controls (Cnt, 26 cases). The hippocampus is a gray matter structure in the limbic system and is involved in processes of motivation and emotions. It also has a central role in the formation of memory. Hippocampal atrophy has been observed in studies of several neurological diseases, such as schizophrenia, epilepsy, and Alzheimer's disease. The goal of our study was to assess shape changes between schizophrenic patients and the control group.

The subjects in this study all have male gender and the same handedness. The two populations are matched for age and ethnicity. The hippocampi were segmented from inversion-recovery-prepped SPGR MRI datasets (resolution: $0.9375 \times 0.9375 \times 1.5mm$) using a manual outlining procedure based on a strict protocol and well-accepted anatomical landmarks (Duvernoy, 1998). The segmentation was performed by a single clinical expert (Schobel et al., 2001) with intra-rater variability of the segmented volume measurements at 0.95. Spherical harmonic (SPHARM) coefficients were computed using a sampling of 2252 points, and the results were normalized via a rigid-body Procrustes alignment and a scaling to unit volume. The m-rep model was built on the aligned full population including the objects of all subjects on both sides, with the right hippocampi mirrored at the interhemispheric plane prior to the model generation. The resulting m-rep model has a single figure topology and a grid sampling of 3×8 medial atoms, in total 24 atoms. The range of the average distance error between the fitted m-rep's boundary and the original boundary was between 0.14mm and 0.27mm (mean error 0.17mm), less than half the voxel size of the original MRI, so the medial shape analysis should capture the shape changes in the image data.

The template for the medial shape analysis was determined by the overall average structure. As the two populations are not equal in size, we computed the overall average as the average of the population averages. Due to age variation in both populations, the shape difference values were corrected for age influence, using a linear least squares model.

The global shape analysis in Table 1 shows a strong trend in the m-rep position analysis on the left side. The m-rep thickness analysis is significant for neither hippocampus. This suggests a deformation shape change in the hippocampus between the schizophrenic and the control group. The results of the m-rep position analysis shows a stronger significance than the SPHARM-PDM analysis that was also carried out. Additionally to the mean difference, several quartile measures (Median, 75% and 95%) were analyzed and produced structurally the same results.

Global Analysis	M-rep Thickness	M-rep Position
Left	$p = 0.722$	$p = 0.0513$
Right	$p = 0.751$	† $p = 0.0001$

Table 1 Results of global shape analysis (average across the surface/medial manifold): Table of group mean difference *p*-values between the schizophrenic and control group (†: significant at $\alpha = 0.05$ significance level).

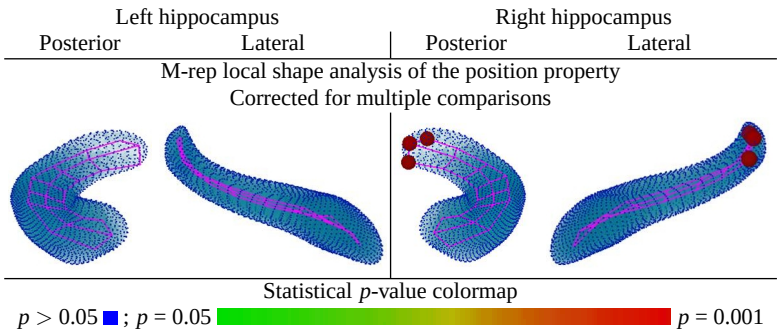


Fig. 14 Statistical maps of the local shape analysis from posterior and lateral views, corrected for multiple comparisons. The m-rep analysis shows the statistical significance at each medial atom using both the color and the radius of spheres placed at the atom positions. The main area of significance is located at the hippocampal tail. The corrected results are overly pessimistic.

The m-rep local position shape analysis (Fig. 14) yields significant changes that are in roughly the same position, mainly in the hippocampal tail, as shown by SPHARM-PDM shape analysis and by distance maps of the averages. No significance was found in the local m-rep thickness analysis. Similar to the outcome of the global analysis, the local m-rep position analysis shows a stronger significance than the SPHARM-PDM analysis. The local shape differences are mainly located at the right hippocampal tail, with near significance in the left hippocampal tail. By

inspecting the average structures of the two groups, we further find that the hippocampal tail region of the control group in our study is more bent than the one of the schizophrenic group.

7.2 Lateral Ventricle Study of Healthy and Schizophrenic Twins

The data for these experiments comes from a twin pair schizophrenia study conducted by Weinberger et al. (2001). High resolution $(0.9375 \times 0.9375 \times 1.5\ \text{mm}^3)$ MRIs scans were acquired from three different subject groups: 9 healthy monozygotic twin pairs (MZ), 10 healthy dizygotic twin pairs (DZ), and 9 monozygotic twin pairs with one twin discordant for schizophrenia and one twin unaffected (DS). See Fig. 15 for some examples. A fourth group of 10 healthy non-related subject pairs (NR) was constructed by matching unrelated members of the two healthy groups. All four groups were matched for age, gender, and handedness. A tenth healthy, monozygotic twin pair was discarded due to possible brain shape changes attributed to major head trauma suffered by one of the twins in a car accident at age seven. A tenth twin pair discordant for schizophrenia was discarded due to hydrocephaly in the unaffected twin.

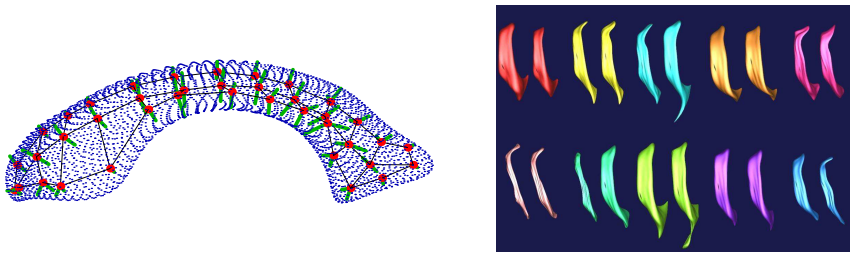


Fig. 15 Left: An example m-rep of a left lateral ventricle. The mesh vertices and off-shooting spokes make up the medial atoms. The shape the m-rep was fit to is shown as a point cloud surrounding it. Right: Ventricle pairs from five monozygotic twin pairs (top) and five dizygotic twin pairs (bottom).

The left and right lateral ventricles were segmented using automatic atlas based tissue classification (van Leemput et al., 1999) and 3-D connectivity. An automatic morphological closing operation was applied to ensure a spherical topology. An area-preserving map was used to map them to a sphere, after which they were converted to a spherical harmonics representation (SPHARM) (Brechbühler et al., 1995). Correspondence on the boundary was established using the first order harmonics (Gerig et al., 2001). Point Distribution Models (PDMs) were constructed by uniformly sampling the boundary at corresponding points. The m-rep models were constructed using a robust method that ensures a common medial topology (Styner

et al., 2003a). For our data, this consists of a single medial sheet with a 3×13 grid of medial atoms, which provides 98% volume overlap with the original segmentations.

From this data set, we wish to determine if the twin pairs that were more closely related had smaller variations in shape. We also wish to see if the shape variations between the discordant and the unaffected twins in the schizophrenic pairs is similar to the normal variation between healthy monozygotic twins. For this purpose, we use the partial test statistics:

$$T_{ij} = \frac{1}{n_1} \sum_{k=1}^{n_1} d(a_{ki}^{1*}, a_{ki}^{2*}) - \frac{1}{n_2} \sum_{k=1}^{n_2} d(b_{ki}^{1*}, b_{ki}^{2*}) \; . \tag{1}$$

Here (a^1, a^2) form the twin pairs for one group, while (b^1, b^2) form the twin pairs for the other. The partial tests are applied separately to all three components of the medial atom location, $\underline{\mathbf{x}}$, as well as the radius and two spoke directions. This gives six partial tests per medial atom, for a total of $p = 3 \times 13 \times 6 = 234$, much larger than the sample size. Each is a one-sided test that the variability in group 2 is larger than that in group 1.

For consistency with previous studies (Styner et al., 2005), all shapes were volume normalized. After normalization, we also applied m-rep alignment, as described by Fletcher et al. (2004), to minimize the sum of squared geodesic distances between models in a medial analog of Procrustes alignment. First, the members of each twin pair were aligned with each other, and then the pairs were aligned together as a group, applying the same transformation to each member of a single pair.

In order to ensure invariance to rotations, we had to choose data-dependent coordinate axes for the $\underline{\mathbf{x}}$ component of each medial atom. Our choice was the axes which diagonalized the sample covariance of the displacement vectors from one twin's atom position to the other at each site. While this had some influence on the results, the general trend was the same irrespective of the axes used.

For each pair of twin groups, we generated $M = 50,000$ permutations, and computed their p-value vectors. Following Section 6.2.1, these were mapped into $\boldsymbol{U}_j$ vectors, from which the empirical distribution of the combined test statistic T'^k from (3) was estimated, producing a single global p-value.

The results are summarized in Table 2. For comparison, we list the results of a previous study which used a univariate test on the average distance between corresponding points on the PDMs (Styner et al., 2005). While we note that the significance of a p-value on an experimental data set is not a useful metric for comparing different methods, it is interesting to see the differences between the two. Our tests give a consistent ranking: MZ $\approx$ DS $<$ DZ $\approx$ NR, which is fully transitive. The boundary study, however, finds a significant difference between DZ and NR, but fails to identify the difference between DS and DZ.

We also performed local tests, to identify specific medial atoms with with strong differences. A multivariate test was conducted using our procedure on the 6 components of each atom, and the results were corrected for multiple tests using the minimum p-value distribution across the shape described in Section 6.1.2. The results are shown in Fig. 16.

	Our Study		Boundary Study (Styner et al., 2005)	
	Left	Right	Left	Right
MZ vs. DS	0.12	0.38	0.28	0.68
MZ vs. DZ	† 0.00006	† 0.0033	† 0.0082	† 0.0399
MZ vs. NR	† 0.00002	† 0.00020	† 0.0018	† 0.0006
DS vs. DZ	† 0.020	† 0.0076	0.25	0.24
DS vs. NR	† 0.0031	† 0.00026	† 0.018	† 0.0026
DZ vs. NR	0.16	0.055	† 0.05	† 0.016

Table 2 *p*-values for paired tests for the difference in the amount of shape variability in groups with different degrees of genetic similarity. Results from our method are in the first two columns, while results from a previous study (Styner et al., 2005) are in the last two for comparison. Groups are: monozygotic (MZ), monozygotic twins with one twin discordant for schizophrenia (DS), dizygotic (DZ), and non-related (NR). (†: significant at the $\alpha = 0.05$ significance level)

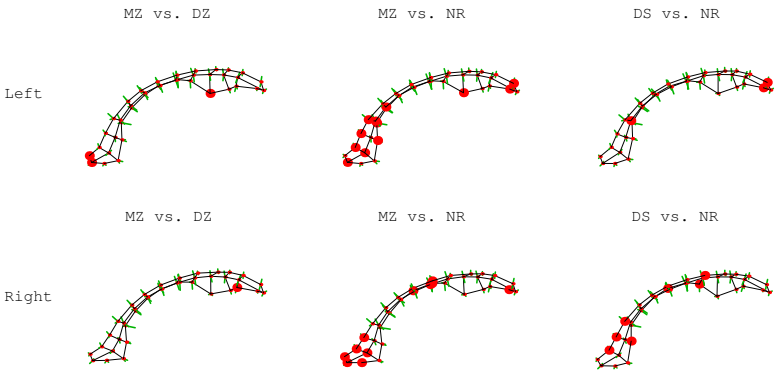


Fig. 16 Results for local tests for the difference in shape variability in groups with different degrees of genetic similarity. Atoms with differences significant at the $\alpha = 0.05$ level are shown in a larger size. Tests not shown had no significant local differences.

8 Discussion and Future Work

8.1 Are M-reps Effective?

The main objective of this chapter was to describe m-reps based methods for 3D medical image segmentation and for statistical characterization of differences of anatomic shapes seen in populations of medical images. M-reps have been used both to capture knowledge of object geometry and to give a basis of the positional correspondences needed in training and measuring geometry-to-image match log probabilities. As well, they have allowed efficient, multiscale operation in both training the probabilities and applying them. It has been our expectation that they provide

more stable estimates of modes of variation and the associated principal variances for a given number of training samples than alternative bases for geometric statistics, and we have some early results suggesting that this is the case, but this is yet to be proven.

In addition, much more than other geometric representations, m-reps have provided a means of yielding probability distributions whose samples were very unlikely to be geometrically improper, avoiding illegal interpenetrations and creases. Checks on geometric propriety via S_{rad} has avoided creasing or near creasing and the improved estimates of boundary normals without lowering the tightness of boundary fits to training binary images. DRIQF statistics based on these fits have led to improved segmentations.

The success of m-reps as an object representation designed for statistical uses should be judged by the success of the applications. Within the class of deformable models methods that might be considered to provide comparable segmentation accuracy, robustness and low interaction requirements, m-reps based segmentations are among the speedier.

In terms of accuracy and robustness, the 3D segmentation method based on m-reps have produced single-figure object, viz. kidney, segmentations that are competitive with human manual segmentation and are, to our knowledge, the most accurate kidney segmentations reported in the literature. The same can be said of the initial multi-object segmentations of male pelvis objects in CT images using intra-patient statistics, though given the serious challenge of this problem, further work must be done before the method can be tested on many patients and its results compared to the results of alternative methods for segmentation of these objects. Moreover, while the apparatus for segmentation of multi-figure objects exists and has been tried on simple test cases, real application and testing of such segmentation is yet to be done.

Of course, when comparing m-reps to other object representations that are being used for segmentation via deformable models, the issue is not simply whether m-reps are as good or better than the alternatives, but whether they are enough better to justify the complexities of the medial representation. Controlled, quantitated validation on a variety of objects by multiple methods in competition needs to be carried out before this can be judged.

We are in the process of making the following improvements to our deformable m-reps segmentation method and software:

1. Sensing and reporting locations on the segmented object that do not have the expected level of geometry-to-image match, so that the user can take actions of relocating that object section and then restart the segmentation.
2. Bringing to routine usability a posterior optimizing atom stage as well as the option of computing a small scale diffeomorphism both in the target object(s) and in the interstitial space between objects in place of the small scale boundary displacement.
3. Developing a form of our software intended for clinical use and thus being as automatic as possible, and making all interactions in clinical terms.

The m-rep hypothesis testing tools have been applied to several studies in neuroimaging and have shown to provide meaningful results. The main advantage of our m-rep hypothesis testing tools over boundary based testing tools is the identification of different types of processes using the different m-rep atom properties. This leads to results that are more intuitively interpretable. In several studies of the hippocampus, the caudate and the lateral ventricles, we have shown that the overall results correlate well between medial and boundary description, but also that our m-rep analysis is able to capture additional information not seen in the boundary analysis.

Our current hypothesis testing tools are based on a true multivariate permutation test approach for hypothesis testing in direct products of metric spaces. The resulting test does not require a priori scaling factors to be chosen, and captures the true multivariate nature of the data. It is well-defined even in the high-dimensional, low-sample size case. The method has been developed for m-reps, though it is suitable for any type of metric data, including potentially categorical data. An important area for future research is the design of suitable partial tests to use in each space. Because they cannot be broken into smaller pieces than a single component of the direct product, the distance to the mean and similar tests are limited in the types of distributions they can describe. For example, the distance from the mean can only characterize an isotropic distribution on the sphere. This would allow us to relax our assumption of identical distribution about the mean.

Even though our hypothesis testing tools have matured to a degree that they can be employed routinely in neuro-imaging studies, there are several limitations to our current tools making the following enhancements to our methods necessary:

1. Developing a combined analysis of multiple objects in order to capture correlated differences of the shape in neighboring brain structures such as the lateral ventricle and the caudate.
2. Enhancing the analysis scheme to incorporate several layers of scale starting at the global multi-object scale down to the local single atom scale.
3. Incorporating statistical models of patient covariates such as gender, age and medication in the permutation test algorithm. The current method incorporates covariates by correcting atom parameters independently using least-squares linear regression.

8.2 Other M-rep Uses and Properties

In a separate paper (Crouch et al., 2003) we have shown how the space parametrization provided by m-reps also allows the interior of the object to be divided into mesh elements useful for efficient mechanical modeling of intra-patient motion of anatomic structures due to such interventions as intrarectal imaging probes. The measures of mechanical energy computed in this approach could be used for segmentation of a patient whose segmented m-rep from an earlier (e.g., planning) image can be used as the model for a segmentation in a later (e.g., intra-treatment) image.

M-reps provide one means of modeling objects and collections of objects; boundary representations (b-reps) are a common alternative means of such object modeling. They share the limitations of all object modeling methods, namely that a single object model will not serve for a class of objects with mixed topologies at the figural level. However, because they explicitly model the interfigural relations, they have special weaknesses when these relations are variable over the population of objects. For example, an m-rep for a right kidney and a separate left kidney will not perform well for a horse-shoe kidney, in which the kidneys are joined. For such mixed classes, a separate m-rep is required for each exemplar. Another issue shared with other object models is instability for nearly spherically or circularly symmetric objects. In such cases the nearly degenerate geometry creates computational instabilities in discriminating among the three major axes which in turn can cause an m-rep to "flip" during deformation in the image data in an unstable manner. However, m-reps share with other object models the particular strength of resolving these orientational instabilities via the relations among objects.

M-reps' special abilities relative to b-reps derive from their explicit representation of object orientation changes such as twisting and bending and of object size changes such as widening and narrowing. Thus statistics on rectal widenings due to gas, on the variability in the relative pose of the two lobes of the liver, and on the orientation of the bladder relative to the prostate are very effective in m-reps terms. The limitations not of m-reps by themselves but of m-reps with statistics come in situations when the orientational or magnificational relationships are very variable. Thus, like b-reps m-reps are well suited to complex slabs and tubes such as the cerebral cortex or the intestine, and both are well suited to intra-patient variations of these structures over time. But because in the population of humans the variability of the folding structure of the cerebral cortex is high and the variability of the curvature of the intestine is high, statistics on m-reps is a weak tool over that population for these structures.

Because m-reps represent the interior of objects, they lose their effectiveness in image situations where only one side of an object appears in an image, and they have weakness relative to b-reps in situations where one side of an object boundary is statistically stable but the other side has great variability. In that situation b-reps can ignore the unstable or unimaged side, whereas m-reps inherently must represent both sides together.

M-reps allow statistics by providing a fixed topology of sheets and their branching. As presently designed, populations that are not well modeled by fixed topology m-reps together with voxel scale refinements will require a different geometric representation.

Acknowledgments

This work was done under the partial support of NIH grants P01 CA47982 and P01 EB02779 and a grant by the Stanley Foundation. A gift from Intel Corp. provided computers on which some of this research was carried out. We thank Xifeng Fang,

Qiong Han, Ja Yeon Jeong, Joshua Levy, Conglin Lu, Derek Merck, Joshua Stough, Gregg Tracton, Guido Gerig, A. Graham Gash,and Delphine Bull for help with models, figures, software, and references. We are grateful to J. Stephen Marron, Keith Muller, and Surajit Ray for help with statistical methods, to Jeffrey Lieberman, Julia Fielding, and Valerie Jewells for medical images, and to members of the UNC Neuroimage Analysis Laboratory and Julian Rosenman for manually segmenting images.

References

Anderson, T. W.: 1958, *An Introduction to Multivariate Statistical Analysis*. New York: John Wiley & Sons, Inc. [328]

Bookstein, F. L.: 1991, *Morphometric Tools for Landmark Data*. Cambridge University Press. [2, 320]

Bouix, S., J. Pruessner, D. L. Collins, and K. Siddiqi: 2005, 'Hippocampal Shape Analysis Using Medial Surfaces'. *Neuroimage* **25**(4), 1077–89. [325]

Brechbühler, C., G. Gerig, and O. Kübler: 1995, 'Parameterization of Closed Surfaces for 3-D Shape Description'. *Computer Vision, Graphics, and Image Processing: Image Understanding* **61**, 195–170. [336]

Broadhurst, R., J. Stough, S. Pizer, and E. Chaney: 2006, 'A Statistical Appearance Model Based on Intensity Quantile Histograms'. In: *IEEE International Symposium on Biomedical Imaging (ISBI)*. pp. 422–425. [315, 322]

Broadhurst, R. E., J. Stough, S. M. Pizer, and E. L. Chaney: 2005, 'Histogram Statistics of Local Model-Relative Image Regions'. In: O. F. Olsen, L. Florack, and A. Kuijper (eds.): *International Workshop on Deep Structure, Singularities and Computer Vision*, Vol. LNCS. pp. 71–82. [309, 316]

Catmull, E. and J. Clark: 1978, 'Recursively generated B-spline surfaces on arbitrary topological meshes'. *Computer Aided Design* **10**, 183–188. [315]

Christensen, G. E., S. C. Joshi, and M. I. Miller: 1997, 'Volumetric Transformation of Brain Anatomy'. *IEEE Transactions on Medical Imaging* **16**(6), 864–877. [320]

Cootes, T., C. Beeston, G. Edwards, and C. Taylor: 1999, 'A Unified Framework for Atlas Matching Using Active Appearance Models'. In: *Information Processing in Medical Imaging (IPMI)*, Vol. 1613. New York, pp. 322–333, Springer. [316]

Cootes, T. F., A. Hill, C. J. Taylor, and J. Haslam: 1993, 'The Use of Active Shape Models for Locating Structures in Medical Images'. In: *IPMI*. pp. 33–47. [305, 306, 314, 316]

Crouch, J., S. M. Pizer, E. L. Chaney, and M. Zaider: 2003, 'Medially Based Meshing with Finite Element Analysis of Prostate Deformation'. In: R. E. Ellis and T. M. Peters (eds.): *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Vol. 2878. Springer LNCS. [287, 340]

Dam, E., P. T. Fletcher, S. M. Pizer, G. Tracton, and J. Rosenman: 2004, 'Prostate Shape Modeling based on Principal Geodesic Analysis Bootstrapping'. In:

C. Barillot, D. Haynor, and P. Hellier (eds.): *Medical Image Computing and Computer-Assisted Intervention (MICCAI) Conference*, Vol. 3217(2). pp. 1008–1016, Springer-Verlag. [311]

Davies, R., C. Twining, T. Cootes, J. Waterton, and C. Taylor: 2002, 'A Minimum Description Length Approach to Statistical Shape Modeling'. *IEEE Transactions on Medical Imaging* **21**(5), 525–537. [314]

Delingette, H.: 1999, 'General Object Reconstruction Based On Simplex Meshes'. *International Journal of Computer Vision* **32**(2), 111–146. [2, 305]

Duda, R., P. Hart, and D. Stork: 2001, *Pattern Classification*. New York: John Wiley and Sons. [306]

Duvernoy, H. M.: 1998, *The human hippocampus, Functional Anatomy, Vascularization and serial Sections with MRI*. Springer. [334]

Edgington, E. S. (ed.): 1995, *Randomization Tests*. Academic Press. [327]

Fletcher, P. T., C. Lu, S. M. Pizer, and S. C. Joshi: 2004, 'Principal Geodesic Analysis for the Study of Nonlinear Statistics of Shape'. *IEEE Transactions on Medical Imaging* **23**, 995–1005. [337]

Foskey, M., B. Davis, L. Goyal, S. Chang, E. Chaney, N. Strehl, S. Tomei, J. Rosenman, and S. Joshi: 2005, 'Large deformation 3D image registration in image-guided radiation therapy'. *Physics in Medicine and Biology* **50**, 5869–5892. [323]

Gerig, G., M. Styner, D. Jones, D. R. Weinberger, and J. A. Lieberman: 2001, 'Shape Analysis of Brain Ventricles Using SPHARM'. In: *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*. pp. 171–178. [336]

Golland, P., W. Grimson, and R. Kikinis: 1999, 'Statistical Shape Analysis Using Fixed Topology skeletons: Corpus Callosum Study'. In: *Information Processing in Medical Imaging*. pp. 382–388. [325]

Han, Q., S. M. Pizer, D. Merck, S. Joshi, and J.-Y. Jeong: 2005, 'Multi-figure Anatomical Objects for Shape Statistics'. In: G. Christensen and M. Sonka (eds.): *Information Processing in Medical Imaging (IPMI)*, Vol. 3565. pp. 701–712. [321]

Ho, S.: 2004, 'Profile Scale Spaces for Statistical Image Match in Bayesian Segmentation'. Ph.D. thesis, University of North Carolina, Chapel Hill, North Carolina. [315]

Jeong, J., S. Pizer, and S. Ray: 2006, 'Statistics on Anatomic Objects Reflecting Inter-Object Relations'. In: *MICCAI Workshop on xx*. p. xx. [289, 323]

Kelemen, A., G. Székely, and G. Gerig: 1999, 'Elastic Model-Based Segmentation of 3D Neuroradiological Data Sets'. *IEEE Transactions on Medical Imaging* **18**(10), 828–839. [306]

Levina, E. and P. Bickel: 2001, 'The Earth Mover's Distance is the Mallows Distance: Some Insights from Statistics'. In: *ICCV 2001*. pp. 251–256. [317]

Lu, C., S. Pizer, and S. Joshi: 2003, 'A Markov Random Field Approach to Multiscale Shape Analysis'. In: L. D. Griffin and M. Lillholm (eds.): *Fourth International Conference On Scale Space Methods in Computer Vision*, Vol. LNCS 2695. pp. 416–431. [311]

Mallows, C.: 1972, 'A Note on Asymptotic Joint Normality'. *Annals of Mathematical Statistics* **43**(2), 508–515. [317]

McInerny, T. and D. Terzopoulos: 1996, 'Deformable models in medical image analysis: a survey'. *Medical Image Analysis* **1**(2), 91–108. [305, 306]

Merck, D., G. Tracton, S. Pizer, and S. Joshi: 2006, 'A Methodology for Constructing Geometric Priors and Likelihoods for Deformable Shape Models'. Technical report, MIDAG, Department of Computer Science, UNC Chapel Hill. [310, 311]

Miller, M. I., S. C. Joshi, and G. E. Christensen: 1999, 'Large Deformation Fluid Diffeomorphisms For Landmark and Image Matching'. In: A. W. Toga and J. Mazziotta (eds.): *Brain Warping*. pp. 115–131, Elsevier. [320]

Nichols, T. E. and A. P. Holmes: 2001, 'Nonparametric Permutation Tests for Functional Neuroimaging: A Primer with Examples'. *Human Brain Mapping* **15**, 1–25. [327]

Pallini, A. and F. Pesarin: 1992, 'A Class of Combinations of Dependent Tests by a Resampling Procedure'. In: K. H. Jöckel, G. Rothe, and W. Sendler (eds.): *Bootstrapping and Related Techniques*, Vol. 376 of *Lecture Notes in Economics and Mathematical Systems*. Berlin, pp. 93–97, Springer-Verlag. [332]

Pantazis, D., R. Leahy, T. Nichol, and M. Styner: 2004, 'Statistical Surface-Based Morphometry using a Non-parametric Approach'. In: *Int. Symposium on Biomedical Imaging(ISBI)*. [327]

Pesarin, F.: 2001, *Multivariate Permutation Tests with Applications in Biostatistics*. Chirchester: John Wiley & Sons, Ltd. [330, 332]

Pilgram, R., P. T. Fletcher, S. M. Pizer, O. Pachinger, and R. Schubert: 2003, 'Common Shape Model and Inter-individual Variations of the Heart using Medial Representation: a pilot study'. Technical report, Institute for Medical Knowledge Representation and Visualization, University for Health Informatics and Technology, Tyrol, Austria. [321]

Pizer, S., J. Jeong, R. Broadhurst, S. Ho, and J. Stough: 2005a, 'Deep Structure of Images in Populations via Geometric Models in Populations'. In: O. Olsen, L. Florack, and A. Kuijper (eds.): *International Workshop on Deep Structure, Singularities and Computer Vision (DSSCV)*, Vol. 3753. pp. 48–58, Springer LNCS. [316]

Pizer, S. M., P. T. Fletcher, S. Joshi, A. G. Gash, J. Stough, A. Thall, G. Tracton, and E. L. Chaney: 2005b, 'A Method and Software for Segmentation of Anatomic Object Ensembles by Deformable M-Reps'. *Medical Physics* **32**(5), 1335–1345. [306]

Rao, M., J. Stough, Y.-Y. Chi, K. Muller, G. S. Tracton, S. M. Pizer, and E. L. Chaney: 2005, 'Comparison of Human and Automatic Segmentations of Kidneys from CT Images'. *International Journal of Radiation Oncology, Biology, Physics* **61**(3), 954–960. [321]

Schobel, S., M. Chakos, G. Gerig, H. Bridges, H. Gu, H. Charles, and J. Lieberman: 2001, 'Duration and Severity of Illness and Hippocampal Volume in Schizophrenia as Assessed by 3D-Manual Segmentation'. *Schizophrenia Research* **49**(1-2), 165. [334]

Staib, L. H. and J. S. Duncan: 1996, 'Model based Deformable Surface Finding for Medical Images'. *IEEE Transactions on Medical Imaging* **15**(5), 1–13. [305]

Stough, J., S. M. Pizer, E. L. Chaney, and M. Rao: 2004, 'Clustering on Image Boundary Regions for Deformable Model Segmentation'. In: *International Symposium on Biomedical Imaging (ISBI)*, Vol. Catalog Number 04EX821C. pp. 436–439, IEEE. [314]

Styner, M., G. Gerig, S. C. Joshi, and S. M. Pizer: 2003a, 'Automatic and Robust Computation of 3-D Medial Models Incorporating Object Variability'. *International Journal of Computer Vision* **55**, 107–122. [311, 336]

Styner, M., G. Gerig, J. Lieberman, D. Jones, and D. Weinberger: 2003b, 'Statistical Shape Analysis of Neuroanatomical Structures Based on Medial Models'. *Medical Image Analysis* **7**(3), 207–220. [326]

Styner, M., J. Lieberman, D. Pantazis, and G. Gerig: 2004, 'Boundary and Medial Shape Analysis of the Hippocampus in Schizophrenia'. *Medical Image Analysis* **8**(3), 197–203. [326]

Styner, M., J. A. Lieberman, R. K. McClure, D. R. Weinberger, D. W. Jones, and G. Gerig: 2005, 'Morphometric Analysis of Lateral Ventricles in Schizophrenia and Healthy Controls Regarding Genetic and Disease-Specific Factors'. *Proceedings of the National Academy of Science* **102**(12), 4872–4877. [337, 338]

Thall, A.: 2004, 'Deformable Solid Modeling via Medial Sampling and Displacement Subdivision'. Ph.D. thesis, University of North Carolina, Chapel Hill, North Carolina. [289, 315]

van Leemput, K., F. Maes, D. Vandermeulen, and P. Seutens: 1999, 'Automated Model-based Tissue Classification of MR Images of the Brain'. *IEEE Transactions on Medical Imaging* **18**, 897–908. [336]

Weinberger, D. R., M. F. Egan, A. Bertolino, J. H. Callicott, V. S. Mattay, B. K. Lipska, K. F. Berman, and T. E. Goldberg: 2001, 'Prefrontal Neurons and the Genetics of Schizophrenia'. *Biological Psychiatry* **50**, 825–844. [336]

Worsley, K. J., S. Marrett, P. Neelin, A. C. Vandal, K. J. Friston, and A. C. Evans: 1996, 'A unified statistical approach for determining significant signals in images of cerebral activation'. *Human Brain Mapping* **4**, 58–73. [327]

Yushkevich, P., H. Zhang, and J. C. Gee: 2005, 'Parametric Medial Shape Representation in 3-D via the Poisson Partial Differential Equation with Non-Linear Boundary Conditions'. In: G. Christensen and M. Sonka (eds.): *IPMI 2005*, Vol. 3565. pp. 162–173, Springer. [314]