A Markov Random Field Approach to Multi-scale Shape Analysis

Conglin Lu, Stephen M. Pizer, and Sarang Joshi

Medical Image Display and Analysis Group
University of North Carolina, Chapel Hill, USA

Abstract. Characterizing the geometric conformation of object complexes requires the description of certain geometric features. These features are most intuitive and provide locality if each is at a restricted range of scale. Thus a complete representation of a geometric entity, such as an object, includes descriptions at multiple scale levels, each describing a residue from the information provided at the larger scales. We have been developing a methodology using medial representations for 3D geometric entities. In this framework, we represent these entities at the following natural discrete scale levels: a whole object complex, individual objects, various object parts and sections, and fine boundary details. This has been proven to be a robust representation of the geometry scale space. It is particularly useful for probabilistic characterization that describes both the common geometry of a population of object complexes and the variation of instances within the population. Using our medial representation, we build Markov random field (MRF) models on the geometry scale space based on the statistics of shape residue across scales and between neighboring geometric entities at the level of locality given by its scale. In this paper, we present how to design MRF models on two scale levels, namely boundary displacement and object sections. This approach can be applied to various applications in medical image analysis such as image segmentation and object discrimination into classes.

1 Introduction

Analysis of the geometric conformation of complexes of one or more objects plays an important role in many computer vision and image analysis applications. For instance, geometric information can be incorporated as priors to guide image segmentation. By imposing geometric constraints on the interpretation of data, one obtains more reliable results, as opposed to making decisions based only on image intensity information. As another example, object discrimination usually involves characterizing and comparing classes that differ in their geometric conformation, e.g., their volume or their shape. In order to effectively work with the geometry of objects, one needs an efficient and accurate way of describing them.

The geometric conformation of a group of one or objects is usually represented by certain geometric features. These features are most intuitive and provide locality if each is at a restricted range of scale. To describe an object complex

at a coarse level, the relative poses of each object may be most informative. On the other hand, to characterize an individual object, one needs to go to a finer scale level and describe different sections, and each section needs to describe its fine boundary details. A complete representation should be able to provide geometric information on all relevant scales. As such, each geometric entity can be regarded as an element in the *geometry scale-space.*

The classical scale-space theory is concerned with analysis of image intensity structure across scales. The most basic scale-space is generated by local Gaussian diffusions [1, 2]. When applied to an image, this diffusion process provides descriptions of the image at different scale levels. The procedure is equivalent to solving a linear partial differential equation (PDE) with initial values. Many other classes of scale-spaces have been proposed, including those generated by various non-linear PDE's and morphological operations [3, 4], as well as spatio-temporal scale-spaces [5]. The idea has also been applied to shape analysis, where objects are deformed in a geometrically consistent manner such that a multi-scale description of objects is attained. The generic behaviors of surface evolution were studied using singularity theory [6–8]. In [9], a general "scale-based geometry" framework is proposed. Examples of multi-scale shape representation include description by singularities [10, 11] and by geometric invariants [12].

The non-linear scale spaces have been used in an attempt to reflect a notion of locality, i.e., locally relevant distance, to the spatial summaries computed by diffusion. In our study of geometric entities such as objects and boundaries, locality must be taken relative to the components of which an entity is formed. Thus an object complex must be described with respect to the distances between neighboring objects, and an object must be described with respect to the distances between the natural sections making up the object. Therefore, instead of treating scale as a continuous parameter as in the powerful PDE theory, we use discrete object-based scale levels. We follow the practice of making scale spaces from discrete levels and *residues* between levels used in the wavelet approaches [13, 14]. Thus we describe the changes in geometry across scales rather than geometric features prominent at all scales. To do so, we choose geometric primitives at each level to describe the residue from the geometric information provided at the next larger scale level. This yields a successive refinement of representations, which enables us to focus on the most relevant and localized geometric information at any particular scale level.

As just discussed, at any given scale level there is a notion of neighbors, i.e., nearby geometric entities at that scale: nearby objects, nearby object sections, etc. Neighbors at a larger scale level are typically more distant than neighbors at a smaller scale level. The neighbor relation realizes the notion of locality. This view allows a hierarchical, multi-scale description of geometric entities, in which a primitive at one scale level is seen relative both to the next larger scale level and relative to its neighbors at that scale level.

We can effect such a viewpoint by seeing object complexes or objects as members of a population of examples in the real world with fixed topology. The reason is that essentially, shape is an attribute of a class of entities. Take the

shape of liver as an example. Globally all livers have the same general geometric conformation; yet on finer scales the geometry varies significantly from one to another. In deciding the topology of the representation of livers, we need to take into account the common geometry as well as the variations among different instances, so that both of these pieces of information can be effectively described. The fixed topology enables establishing fixed correspondences across the population between entities at a scale level and between scale levels. What differs between members of the population is the quantitative, geometric parameters and not qualitative properties of structure or topology.

We have realized this methodology by describing 3D geometric entities using medial representations combined with boundary displacements, which together form a representation called *m-reps* [15, 16]. In this framework, geometric entities are explicitly represented at discrete locations at the following discrete scale levels: a whole object complex, individual objects, different object parts and sections, and boundary points. Computing the fixed topology from a population in this framework is achieved using the method of Styner [17], not discussed further here.

In this medial framework, at each scale level the geometry is represented by the configuration of certain geometric primitives that are most descriptive at that level. For example, an object as a whole is described by a global similarity transformation, whereas different sections of the object are characterized via more local transformations. This differs from the traditional scale-space in that different primitives are used at different scales in the representation.

One describes a population of geometric entities by using probabilistic models. For instance, 2D curve models were proposed in [18, 19]. For 3D objects, principal component analysis on surface points [20, 21] and spherical harmonic descriptors [22] have been studied extensively. In these models, a probability measure is put on the space of all possible deformations from a common template. The parameters of the distribution can be estimated and learned from a training data set.

The relationships between adjacent scale levels and the neighbor relationships at each scale level in our framework makes Markov random fields (MRF) the natural probabilistic approach. Given the configuration at each scale level, the geometric residue information at the next smaller scale level is described by the corresponding geometric primitives and incorporated into an MRF model. In this approach there are a small number of parameters to be estimated at each location within each scale level, so accurate parameter estimation can be achieved with limited numbers of training cases. In this paper, we discuss how to design MRF models on two scale levels, namely the boundary displacement level and the object section level.

In what follows, we first briefly describe our m-reps representation mechanism in section 2. The details of the Markov random field models are presented in sections 3 - 5. Section 6 shows some statistics on the geometric residues of hippocampi. We finish with some concluding remarks in section 7.

2 Multi-scale Shape Representation by M-reps

Two basic types of geometric object representation are boundary-based and medial-based models. Boundary-based methods describe an object by a mesh of boundary points or landmarks. While this is an intuitive and direct way of representation, it is not easy to use it to describe deformations such as elongation, bending, and widening. These kinds of geometric information is more easily described by medial-based methods, where an object is represented by one or more medial manifolds, from which the boundary of the object can be determined.

Most existing medial representations derive the medial structure from the boundary of an object [23, 11]. They are sensitive to noise and small perturbations on the boundary. This instability essentially results from the fact that the boundary includes geometric information at multiple scales. To obtain more stable descriptions, it is crucial to build medial representations in a multi-scale fashion, including a boundary displacement component.

In our framework, called m-reps, we start from a medial representation which implies a boundary. An object is described implicitly as a set of continuous medial manifolds, which can be sampled at different scales to yield discrete representations. Each sample point is called a medial atom (see Fig. 1(a)). It is a 4-tuple $\mathbf{m} = (\mathbf{x}, \mathbf{R}, r, \theta)$ consisting of

- a translation $\mathbf{x} \in \mathbb{R}^3$, specifying the position of the medial point; we can consider this translation in units of the medial width, r;
- a rotation $\mathbf{R} \in \mathrm{SO}(3)$, describing a local orthonormal frame given by $(\mathbf{n}, \mathbf{b}, \mathbf{b}^\perp)$, where $\mathbf{n}$ is the normal to the medial manifold, $\mathbf{b}$ is in the tangent plane along the direction of the fastest narrowing of the implied boundary sections, and $\mathbf{b}^\perp = \mathbf{n} \times \mathbf{b}$;
- a magnification scalar $r \in \mathbb{R}^+$, the local width, defined as the distance from the medial point to the implied boundary points;
- a 2D rotation angle $\theta \in [0, \frac{\pi}{2}]$, called the object angle, which determines the angulation of the implied boundary sections relative to $\mathbf{b}$.

Each of these 4 elements is a member of an algebraic group, a property that will be of importance later in the paper.

The two implied boundary points, $\mathbf{y}_0$ and $\mathbf{y}_1$, with relative surface normals $\mathbf{n}_0, \mathbf{n}_1$, are given by

$$\mathbf{n}_0 = \cos(\theta)\mathbf{b} - \sin(\theta)\mathbf{n}, \qquad \mathbf{n}_1 = \cos(\theta)\mathbf{b} + \sin(\theta)\mathbf{n},$$
$$\mathbf{y}_0 = \mathbf{x} + r\mathbf{n}_0, \qquad \mathbf{y}_1 = \mathbf{x} + r\mathbf{n}_1.$$

An m-rep figure is a quadrilateral mesh of medial atoms, with spacing determined through the analysis of the training population, as described in [17]. It describes a slab-like object or object part. Each atom describes a through-section of the object, and 4-adjacency in the mesh determines the neighbor relationship. A single figure object is shown in Fig. 1(b). Given an m-rep figure, a smooth boundary surface is generated by a subdivision surface algorithm [24] that interpolates the boundary positions and normals implied by each atom.
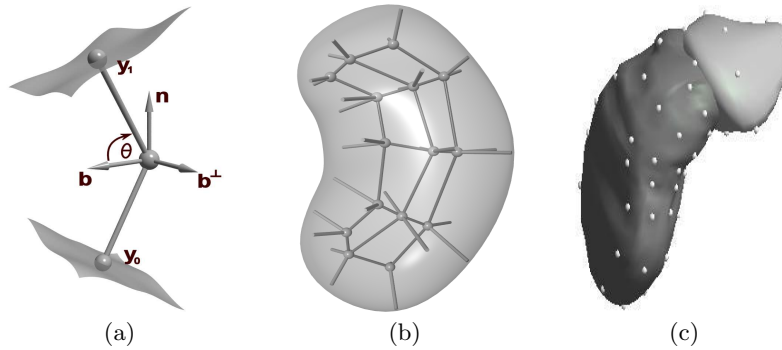
Fig. 1: M-reps. (a) A medial atom with a cross-section of the boundary surface it implies; (b) A single figure m-rep model of a kidney and its boundary surface; (c) A two-figure m-rep liver model.

In general, objects are represented by a linked figural model, together with boundary displacements. The main section of the object is represented by a main figure, and different branches, protrusions or indentations are modelled by various subfigures. The relationship among the figures can be described by a tree structure. Fig. 1(c) illustrates a two-figure m-rep model. Finally, an object complex can be described by the relative configurations of different objects.

Being a representation that contains a medial component, m-reps have the ability of representing solid 3D regions and their boundaries simultaneously. This can be exploited to describe inter-figure and inter-object relations. The m-rep representation of an object complex is obtained by describing the geometry at a hierarchy of scales. Each scale level describes the residue of geometric information provided at the next larger scale by specifying the appropriate residue transformation of the corresponding geometric primitives. A brief summary of this mechanism is given in Table 1.

Scale Level	Geometric Entity	Primitive	Residue Transformation
1	Object complex	Object complex pose	Similarity
2	Object	Object pose	Similarity
3	Figure	Main figure and subfigure poses	Main figure and subfigure transformations
4	Medial atom	Atom configuration	Atom transformation
5	Boundary vertex	Vertex position	Displacement along medially implied normal

Table 1: The scale levels of m-reps, with primitives and transformations at each scale.

As one goes from the object complex level to the atom level, a successively refined boundary representation can be derived. At each level each entity has neighbors of the same entity type. At the finest scale level, each boundary point has boundary position neighbors and moves along the medially implied normal direction to "fine tune" the description. There is a boundary tolerance associated with each scale level. At larger scales, the tolerance is higher, so details are ignored and global features are revealed; at smaller scales, the description focuses on refinements of the larger scales to describe more local geometric features.

The m-rep framework defines a geometry scale-space with a discrete scale parameter, allowing local geometric features at different scale levels to be explicitly described. Furthermore, the medial structure provides a multi-scale intrinsic coordinate system which is extremely well suited for statistical analysis of shapes, because correspondence among a population can be established systematically.

3 Probabilistic M-reps Models

3.1 Markov Random Fields

Given that any model has a fixed topology, we require a probability distribution on the space of geometric variables, i.e., real-valued random variables characterizing the primitives. These variables specify an element of the algebraic group of operations applicable to that primitive. The total number of such random variables is usually of very high dimension. The MRF approach handles this problem by characterizing geometric information through local interactions among geometric primitives at various scale levels.

An MRF model is defined with respect to a dependency graph $G = (V, E)$, which is a simply connected graph. Each vertex $v \in V$ corresponds to a random variable X_v in the model. The set of vertices that are connected to v via an edge in E is called the neighborhood of v, and is denoted by $\mathcal{N}(v)$. The completely connected subgraphs of G (including singletons) are called the cliques of G. A model P is said to be an MRF with respect to G if

$$\mathrm{Prob}\Big(X_v \Big| \text{ all other random variables }\Big) = \mathrm{Prob}\Big(X_v \big| \{X_u : u \in \mathcal{N}(v)\}\Big),$$

where $\mathrm{Prob}(\cdot|\cdot)$ denotes conditional probability. P is said to be a Gibbs distribution with respect to G if the joint probability density of $\{X_v\}$ has the form

$$p_\Theta(\{X_v : v \in V\}) = \frac{1}{Z(\Theta)} \exp\{-\sum_{C \in \mathcal{C}} A_C(X_C; \Theta)\}, \tag{1}$$

where $\mathcal{C}$ is the set of cliques of G, $X_C = \{X_v : v \in C\}$, Θ is a set of parameters, Z is a normalizing constant. Each $A_C \geq 0$ is called a potential function and depends only on those random variables whose indices are in C. The Hammersley-Clifford Theorem [26] establishes the equivalence between MRF's and Gibbs distributions with respect to the same dependency graph G. This allows one to specify an MRF by specifying the potentials in the corresponding Gibbs form.

The main advantage of the MRF approach is that the probability density to be estimated is specified by a relatively small number of parameters. These parameters can thus be efficiently learned from a training data set. One way of estimating them is to seek the parameter values that maximize the likelihood of the data. These estimates are called the maximum likelihood (ML) estimates. In general this is done via stochastic sampling algorithms, such as Markov Chain Monte Carlo (MCMC) methods [26].

3.2 Markov Random Field M-reps Models

In the Markov random field approach for our framework, the neighbor relations are between scale levels and among neighbors at any scale level. Suppose there are l scale levels indexed by $1, 2, \ldots, l$, with scale 1 being the coarsest. Let $\mathbf{z}^k$ denote the collection of geometric primitives at scale level k. At each scale level k, every primitive $\mathbf{z}^k_j$ has a value implied by a corresponding primitive at the previous larger scale, which is called the parent primitive of $\mathbf{z}^k_j$ and denoted by $\mathcal{P}(\mathbf{z}^k_j)$. For example, at the figure level, each figure implies the medial atom primitives that make it up, with position, orientations, etc. relative to the figural geometry as in the mean of the training population. Let $\Delta\mathbf{z}^k_j$ denote the inter-scale residual giving the *difference* between $\mathbf{z}^k_j$ and $\mathcal{P}(\mathbf{z}^k_j)$, where differences are taken with respect to the group operations defining the primitive, e.g., translation, rotation, magnification, and object angulation for a medial atom. Let $\Delta\mathbf{z}^k = \{\Delta\mathbf{z}^k_j\}$. At each scale level we describe local features by residual geometric information from the previous larger level by $\Delta\mathbf{z}^k$ with the Markov assumption

$$\mathrm{Prob}\Big(\mathbf{z}^k|\{\mathbf{z}^1, \ldots, \mathbf{z}^{k-1}\}\Big) = \mathrm{Prob}(\mathbf{z}^k|\mathbf{z}^{k-1}) = \mathrm{Prob}(\Delta\mathbf{z}^k), \quad \text{for } k > 1.$$

In doing so, we are describing the inter-scale-level relationship via the residuals and assuming that residuals at one scale are independent of those at other scales.
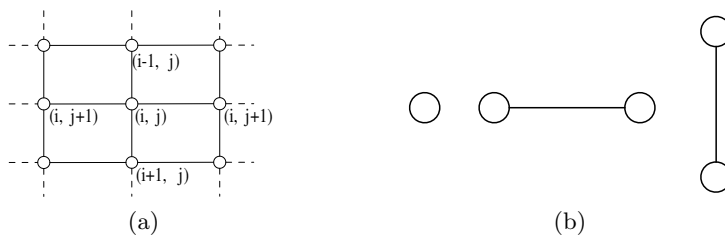

Fig. 2: The 4-neighbor structure for quad-mesh. (a) A typical node has 4 neighbors. (b) The cliques of the quad-mesh.

The residual probability distributions $\{P_k(\Delta\mathbf{z}^k)\}$ are defined as MRF models, with respect to the canonical neighborhood structure induced by the natural spatial relationship among primitives. For example, at the object level, where

the primitives are objects, the neighbors of an object are its adjacent objects. At the atom level, the primitives are medial atoms. The canonical neighborhood structure is the 4-adjacency graph induced by the quad-mesh structure, as shown in Fig. 2(a), since we sample the medial manifold by a quadrilateral array of atoms. The cliques of this dependency graph are single vertices and pairs of vertices that are connected by a horizontal or vertical edge in the quad-mesh (see Fig. 2(b)). If another sampling mesh, e.g. triangular mesh, is used, then appropriate canonical neighborhood structure can be induced similarly.

By the Hammersley-Clifford Theorem, the density of the MRF model P_k can be written in Gibbs form

$$p_k(\{\Delta \mathbf{z}^k\}) \propto \exp\{-\sum_{C \in \mathcal{C}} A_C(\Delta \mathbf{z}_C^k)\},$$

where $\mathcal{C}$ is the set of cliques, and $\Delta \mathbf{z}_C^k = \{\Delta \mathbf{z}_j^k : j \in C\}$. Our goal is to design these MRF models so that they can be specified by a relatively small number of parameters. Two sets of details need to be given:

- in defining $\Delta \mathbf{z}_j^k$, i.e., representing residual geometric information by primitives;
- in defining the potentials A_C, i.e., the form of the probability distribution and the means of estimating the parameters.

In the next two sections, we discuss these issues on two scale levels within the m-reps framework, namely the boundary level and the medial atom level.

4 MRF Models for Boundary Displacement

4.1 Model Description

For any m-rep figure, the medial manifold implies a 3D surface, which is represented by a dense set of boundary points. These are the geometric primitives at the boundary level. The medial manifold is parameterized by a u-v coordinate system, which induces a (u, v, t) coordinate system on the implied boundary. For each boundary point, (u, v) is the coordinate of the corresponding medial point, and t specifies where on the cross section the point lies. These coordinates provide correspondence between boundary points on different objects. Moreover, the description of each point includes a radius r, which is the distance between the point and the corresponding medial point, and a surface normal vector at that point.

At the boundary level, each medially implied boundary point is displaced along its normal vector to obtain a finer scale description. The displacement is measured in the unit of local radius. If the absolute amount of movement of the j-th point is d_j, then by describing distance in multiples of object width, so as to maintain magnification invariance, we define the displacement of that point to be the dimensionless variable $w_j = d_j/r_j$. This variable describes displacement as a member of the 1-dimensional group $\mathbb{R}$. The displacement field $\mathbf{w} = \{w_j\}$ on

the medially implied boundary points is the residual geometric information at this scale level. The boundary surface is determined by the medial representation together with the displacement field.

Currently we use a quad-mesh to sample the boundary, thus the canonical neighborhood structure is the 4-adjacency structure. With respect to this graph, we define Markov random field model at the boundary level on the displacement field $\mathbf{w}$. The density has Gibbs form

$$p(\mathbf{w}) = \frac{1}{Z} \exp\left\{ -\sum_i A_i(w_i) - \sum_{<i,j>} B_{ij}(w_i, w_j) \right\}, \tag{2}$$

with respect to Lebesgue measure, where $<i, j>$ denotes that points i and j are neighbors. We assume that the distribution on $\mathbf{w}$ is a zero-mean Gaussian distribution, with the particular density form

$$p_q(\mathbf{w}) = \frac{1}{Z(q_1, q_2)} \exp\left\{ -\frac{q_1}{2} \sum_{i=1}^n s_i w_i^2 - \frac{q_2}{2} \sum_{<i,j>} \frac{s_i + s_j}{2} \frac{\left(\frac{r_i + r_j}{2}\right)^2}{\|\mathbf{x}_i - \mathbf{x}_j\|^2} (w_i - w_j)^2 \right\}$$

$$= \frac{1}{Z(q_1, q_2)} \exp\left\{ -\frac{q_1}{2} \sum_{i=1}^n s_i w_i^2 - \frac{q_2}{2} \sum_{<i,j>} b_{ij}(w_i - w_j)^2 \right\}, \tag{3}$$

where q_1, q_2 are positive parameters, $Z(q_1, q_2)$ is a constant depending on q_1 and q_2, n is the number of points, r_i, s_i, and $\mathbf{x}_i$ are respectively the radius, surface area, and position associated with the i-th medially implied boundary point. The exponent above is a discrete approximation of the energy function

$$-\frac{q_1}{2} \int_{\mathcal{S}} \frac{d^2(\mathbf{x})}{r^2(\mathbf{x})} d\mathbf{x} - \frac{q_2}{2} \int_{\mathcal{S}} \|\nabla d(\mathbf{x})\|^2 d\mathbf{x},$$

where $d(\mathbf{x}), r(\mathbf{x})$ are the absolute displacement and radius at point $\mathbf{x}$, respectively. Notice that

$$p_q(w_i | \{w_j, j \neq i\}) \propto \exp\left\{ -\frac{q_1}{2} s_i w_i^2 - \frac{q_2 \sum_{<i,j>} b_{ij}}{2} \left(w_i - \sum_{<i,j>} \frac{b_{ij}}{\sum_{<i,j>} b_{ij}} w_j \right)^2 \right\}$$

This can be interpreted as putting a penalty on the amount of w_i as well as the difference between the displacement of point i and a weighted average of those of the neighboring points.

Different sections of the boundary can be modelled by the same MRF model (3) with different parameter values, which reflect the variation of boundary geometry in various sections.

4.2   Parameter Estimation

We now discuss how to estimate the parameters q_1 and q_2 in (3). Rewrite (3) as

$$p_q(\mathbf{w}) = \frac{1}{Z(q_1, q_2)} \exp\left\{ -\frac{1}{2} \mathbf{w}^T \mathbf{\Sigma}^{-1} \mathbf{w} \right\}. \tag{4}$$

Here $\boldsymbol{\Sigma}$ is the covariance matrix, and $\boldsymbol{\Sigma}^{-1} = q_1\mathbf{D} + q_2\mathbf{B}$, where $\mathbf{D}$ is an $n \times n$ diagonal matrix and $\mathbf{B}$ is an $n \times n$ symmetric, sparse matrix. The entries of $\mathbf{D}$ and $\mathbf{B}$ are determined by $\{s_i\}$ and $\{b_{ij}\}$. The normalizing constant is

$$Z(q_1, q_2) = (2\pi)^{\frac{n}{2}} |\det(q_1\mathbf{D} + q_2\mathbf{B})|^{-\frac{1}{2}}.$$

Since

$$|\det(q_1\mathbf{D} + q_2\mathbf{B})| = |\det(\mathbf{D})| \cdot |\det(q_1\mathbf{I} + q_2\mathbf{D}^{-\frac{1}{2}}\mathbf{B}\mathbf{D}^{-\frac{1}{2}})| = \prod_{i=1}^{n} s_i \prod_{i=1}^{n}(q_1 + q_2\lambda_i),$$

where $\lambda_i, i = 1, \ldots, n$ are the eigenvalues of $\mathbf{D}^{-\frac{1}{2}}\mathbf{B}\mathbf{D}^{-\frac{1}{2}}$, we have

$$\log Z(q_1, q_2) = \frac{n}{2}\log(2\pi) - \frac{1}{2}|\det(q_1\mathbf{D} + q_2\mathbf{B})|$$

$$= \frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{n}\log(s_i) - \frac{1}{2}\sum_{i=1}^{n}\log(q_1 + q_2\lambda_i).$$

Given a training data set, we seek the maximum likelihood estimates of q_1, q_2. Suppose there are M independent samples $\{\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_2, \ldots, \widehat{\mathbf{w}}_M\}$, with

$$\widehat{\mathbf{w}}_i \sim \mathcal{N}\Big(\mathbf{0}, (q_1\mathbf{D}_i + q_2\mathbf{B}_i)^{-1}\Big), \qquad i = 1, 2, \ldots, M,$$

where $\{\mathbf{D}_i, \mathbf{B}_i\}$ have the same structure as $\mathbf{D}, \mathbf{B}$. The likelihood function is

$$L(q_1, q_2) = -\sum_{i=1}^{M}\log\left(p_q^{(i)}(\widehat{\mathbf{w}}_i)\right) = -\sum_{i=1}^{M}\left(\log Z^{(i)} + \frac{1}{2}\widehat{\mathbf{w}}_i^T(q_1\mathbf{D}_i + q_2\mathbf{B}_i)\widehat{\mathbf{w}}_i\right).$$

Using the Cauchy-Schwartz inequality, we can show that the Hessian matrix $\nabla^2 L$ is negative semi-definite. Therefore, the maximum of L occurs at (q_1^*, q_2^*) such that $\nabla L(q_1^*, q_2^*) = \mathbf{0}$, which yields

$$\sum_{i=1}^{M}\sum_{j=1}^{n}\frac{1}{q_1^* + q_2^*\lambda_j^{(i)}} = \sum_{i=1}^{M}\widehat{\mathbf{w}}_i^T\mathbf{D}_i\widehat{\mathbf{w}}_i, \tag{5a}$$

$$\sum_{i=1}^{M}\sum_{j=1}^{n}\frac{\lambda_j^{(i)}}{q_1^* + q_2^*\lambda_j^{(i)}} = \sum_{i=1}^{M}\widehat{\mathbf{w}}_i^T\mathbf{B}_i\widehat{\mathbf{w}}_i, \tag{5b}$$

where $\{\lambda_j^{(i)}\}$ are the eigenvalues of $\mathbf{D}_i^{-\frac{1}{2}}\mathbf{B}_i\mathbf{D}_i^{-\frac{1}{2}}$. The above equations are solved numerically.

5   Markov Random Field Models for Object Sections

5.1   The MRF Model

The primitives at the atom level are medial atoms, which describe sections of an object. For simplicity we only consider single-figure objects. Let $\{\mathcal{A}_i\}$ denote

the set of atoms of a figure at the atom scale level. At the previous larger scale, the figural scale, the object is described by another set of medial atoms $\{\mathcal{M}_i\}$, which are the parent primitives of $\{\mathcal{A}_i\}$. The residue geometric information at the atom level is described by the differences between $\{\mathcal{A}_i\}$ and $\{\mathcal{M}_i\}$.

As discussed in section 2, each medial atom $\mathcal{A}_i$ is characterized by a 4-tuple $(\mathbf{x}_i, \mathbf{R}_i, r_i, \theta_i)$. Suppose that $\mathcal{M}_i = (\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{R}}_i, \widetilde{r}_i, \widetilde{\theta}_i)$ is the corresponding parent atom at the previous scale level. We define the atom residue to be

$$\Delta\mathcal{A}_i = \left(\frac{\mathbf{x}_i - \widetilde{\mathbf{x}}_i}{\widetilde{r}_i}, \ \widetilde{\mathbf{R}}_i^{-1}\mathbf{R}_i, \ \frac{r_i}{\widetilde{r}_i}, \ \theta_i - \widetilde{\theta}_i\right) = (\Delta\mathbf{x}_i, \Delta\mathbf{R}_i, \Delta r_i, \Delta\theta_i).$$

It is an element of the product space $G = \mathbb{R}^3 \times SO(3) \times \mathbb{R}^+ \times SO(2)$. Each component of $\Delta\mathcal{A}_i$ can be regarded as an element of the corresponding group. Let $d_E(\cdot, \cdot)$ be the Euclidean distance on $\mathbb{R}^3$, $d_R(\cdot, \cdot)$ be the Riemannian distance on $SO(3)$. The corresponding norms are denoted by $\|\cdot\|_E$ and $\|\cdot\|_R$. The distance d_A between two atom residues is defined to be

$$d_A(\Delta\mathcal{A}_i, \Delta\mathcal{A}_j) =$$
$$\sqrt{d_E^2(\Delta\mathbf{x}_i, \Delta\mathbf{x}_j) + d_R^2(\Delta\mathbf{R}_i, \Delta\mathbf{R}_j) + |\ln(\Delta r_i/\Delta r_j)|^2 + |\Delta\theta_i - \Delta\theta_j|^2}. \quad (6)$$

The corresponding norm is denoted by $\|\cdot\|_A$.

We now describe the MRF model for atom residues with respect to the canonical 4-neighbor structure. The joint probability on the atom residues has a density of the Gibbs form

$$p(\{\Delta\mathcal{A}_i\}) \propto \exp\left\{-\sum_i f_i(\Delta\mathcal{A}_i) - \sum_{<i,j>} g_{ij}(\Delta\mathcal{A}_i, \Delta\mathcal{A}_j)\right\}. \quad (7)$$

This density is with respect to the Haar measure on G. We choose the potentials f_i and g_{ij} to be quadratic functions. In particular,

$$f_i(\Delta\mathcal{A}_i) = \frac{\sigma_i}{2}d_A^2(\Delta\mathcal{A}_i, \mathcal{E}) = \frac{\sigma_i}{2}\|\Delta\mathcal{A}_i\|_A^2,$$

where $\mathcal{E} = (\mathbf{0}, \mathbf{I}, 1, 0)$, and

$$g_{ij}(\Delta\mathcal{A}_i, \Delta\mathcal{A}_j) = \frac{\tau_{ij}}{2d_A^2(\mathcal{M}_i, \mathcal{M}_j)}d_A^2(\Delta\mathcal{A}_i, \Delta\mathcal{A}_j),$$

for neighboring pairs of medial atoms. The full probability density is

$$p(\{\Delta\mathcal{A}_i\}) \propto \exp\left\{-\sum_i \frac{\sigma_i}{2}\left(\|\Delta\mathbf{x}_i\|_E^2 + \|\Delta\mathbf{R}_i\|_R^2 + |\ln(\Delta r_i)|^2 + |\Delta\theta_i|^2\right)\right.$$
$$-\sum_{<i,j>} \frac{\tau_{ij}}{2d_A^2(\mathcal{M}_i, \mathcal{M}_j)}\left(\|\Delta\mathbf{x}_i - \Delta\mathbf{x}_j\|_E^2 + \|(\Delta\mathbf{R}_i)^{-1}\Delta\mathbf{R}_j\|_R^2\right.$$
$$\left.\left. + |\ln(\Delta r_i) - \ln(\Delta r_j)|^2 + |\Delta\theta_i - \Delta\theta_j|^2\right)\right\}.$$
$$(8)$$

The conditional distribution of $\Delta\mathcal{A}_i$ given the rest of the residues has density

$$p(\Delta\mathcal{A}_i|\Delta\mathcal{A}_{\{j\neq i\}}) \propto \exp\left\{-\frac{\sigma_i}{2}\|\Delta\mathcal{A}_i\|_A^2 - \sum_{<i,j>}\frac{\tau_{ij}}{2d_A^2(\mathcal{M}_i,\mathcal{M}_j)}d_A^2(\Delta\mathcal{A}_i,\Delta\mathcal{A}_j)\right\}.$$

This model has an intuitive interpretation which is similar to the boundary displacement model (3). More precisely, the first term in the exponent penalizes the difference between $\mathcal{A}_i$ and its parent $\mathcal{M}_i$, whereas the second term penalizes $\Delta\mathcal{A}_i$ from being different to a weighted average of residues of the neighboring atoms, given the configurations of $\{\Delta\mathcal{A}_j : j \neq i\}$. Neighbors that are closer in the $\|\cdot\|_A$ metric have higher weights.

5.2 Discussion

Given a training data set, the parameters $\{\sigma_i, \tau_{ij}\}$ of the probability model (8) can be estimated using the maximum likelihood method. Since the space G of atom residues is not Euclidean, even though the potentials are quadratic, the distribution is not Gaussian. The maximum likelihood estimates of the parameters in this case are obtained by Markov Chain Monte Carlo methods. However, this is a computationally expensive procedure.

Here we present an alternative model whose parameters are easier to estimate. We start by noticing that the Riemannian distance between two rotations $\Delta\mathbf{R}_1$ and $\Delta\mathbf{R}_2$ is given by $\|\mathrm{Log}(\Delta\mathbf{R}_1^{-1}\Delta\mathbf{R}_2)\|_F$, where $\|\cdot\|_F$ is the Frobenius matrix norm, and for $\mathbf{R} \in SO(3)$,

$$\mathrm{Log}(\mathbf{R}) = \begin{cases} \mathbf{0}, & \text{if } \theta = 0; \\ \dfrac{\theta}{2\sin\theta}(\mathbf{R}-\mathbf{R}^T), & \text{if } \theta \neq 0. \end{cases}$$

Here θ satisfies $\mathrm{tr}(\mathbf{R}) = 1 + 2\cos\theta$ and $|\theta| < \pi$. When $\Delta\mathbf{R}_1, \Delta\mathbf{R}_2$ are close to identity, as in the case for atom residues, their Riemannian distance can be approximated by $\|\mathrm{Log}(\Delta\mathbf{R}_2) - \mathrm{Log}(\Delta\mathbf{R}_1)\|_F$. Now, define an invertible map L by

$$L : \Delta\mathcal{A} = \left(\Delta\mathbf{x}, \Delta\mathbf{R}, \Delta r, \Delta\theta\right) \in G \mapsto \Delta\mathcal{L} = \left(\Delta\mathbf{x}, \log(\Delta\mathbf{R}), \ln(\Delta r), \Delta\theta\right) \in \mathfrak{g}.$$

Then we can approximate the distance d_A defined in (6) on G, the space of atom residues, by the distance $d_{\mathfrak{g}}$ on the linear space $\mathfrak{g}$:

$$\begin{aligned} d_A^2(\Delta\mathcal{A}_1, \Delta\mathcal{A}_2) &\approx d_{\mathfrak{g}}^2(\Delta\mathcal{L}_1, \Delta\mathcal{L}_2) \\ &= \|\Delta\mathbf{x}_1 - \Delta\mathbf{x}_2\|^2 + \|\log(\Delta\mathbf{R}_1) - \log(\Delta\mathbf{R}_2)\|_F^2 \\ &\quad + |\ln(\Delta r_1) - \ln(\Delta r_2)|^2 + |\Delta\theta_1 - \Delta\theta_2|^2. \end{aligned}$$

Instead of (8), we define a conditional probability distribution on $\{\Delta\mathcal{L}\}$ with density

$$\begin{aligned} &p\left(\{\Delta\mathcal{L}_i\}\right) \\ &= \frac{1}{Z}\exp\left\{-\sum_i\frac{\sigma_i}{2}\|\Delta\mathcal{L}_i\|_{\mathfrak{g}}^2 - \sum_{<i,j>}\frac{\tau_{ij}}{2d_A^2(\mathcal{M}_i,\mathcal{M}_j)}d_{\mathfrak{g}}^2\left(\Delta\mathcal{L}_i,\Delta\mathcal{L}_j\right)\right\}, \end{aligned} \qquad (9)$$

where $\| \cdot \|_{\mathfrak{g}}$ is the metric corresponding to $d_{\mathfrak{g}}$. (9) induces a probability distribution on G via L^{-1}, which takes each $\Delta\mathcal{L}$ back to $\Delta\mathcal{A}$.

The model (9) is essentially a Gaussian model on the linear space $\mathfrak{g}$. The parameters $\{\sigma_i, \tau_{ij}\}$ can be estimated from a training data set by the maximum likelihood principle. In this case we can avoid MCMC methods and estimate the parameters directly.

The idea of approximating the group G by a linear space $\mathfrak{g}$ can be formalized with Lie group theory, but is beyond the scope of this paper. For a detailed description of Lie groups and their application in statistical shape analysis, see [27] and the references therein.

6 Residue Statistics of Hippocampi

In this section we present some residue statistics of a sample population of hippocampi. The data set contains 86 left hippocampi, each being represented as a single m-rep figure by a 3×8 array of medial atoms.

Fig. 3 shows some boundary statistics of the samples. The top left group of figures show histograms of all boundary points, and the other three groups show histograms of different sections of the boundary. Each section contains boundary points that have the same (u, v) coordinates. The sections whose statistics are shown are indicated by the big dots. In each case, the left plot is the histogram of boundary displacements, and the right one shows the histogram of difference in displacements between neighboring boundary points. Notice how the histograms vary according to position changes.
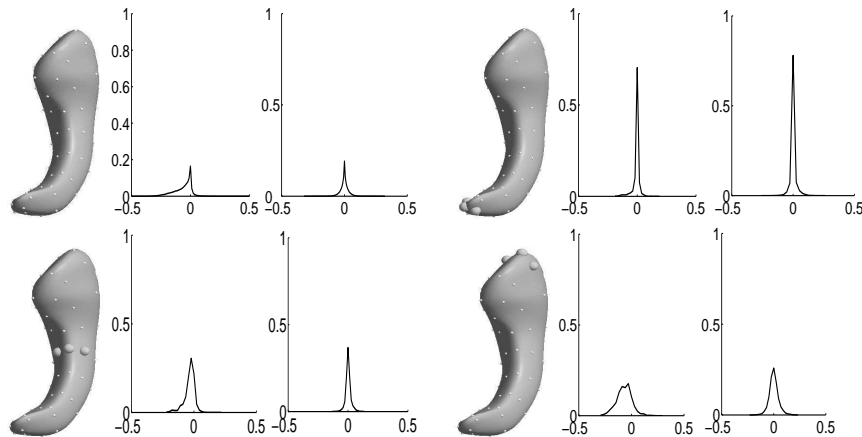

Fig. 3: Boundary displacement statistics of hippocampi. In each group, the left plot is the histogram of boundary displacement, and the right one is that of the difference of displacements between neighbors.

Fig. 4 illustrates some statistics on atom residues $\Delta\mathcal{A} = (\Delta\mathbf{x}, \Delta\mathbf{R}, \Delta r, \Delta\theta)$. The 3 rows of figures correspond to 3 different atoms. In each row, the left two plot are the histograms of $\log(\Delta r)$ and $\|\Delta\mathbf{x}\|$. The right two show histograms of $\log(\Delta r) - \overline{\log(\Delta r)}$ and $\|\Delta\mathbf{x} - \overline{\Delta\mathbf{x}}\|$, where $\overline{\log(\Delta r)}$ and $\overline{\Delta\mathbf{x}}$ are weighted averages of $\log(\Delta r)$ and $\Delta\mathbf{x}$ of the atom's neighbors, respectively. Again, the histogram varies significantly with position, indicating that different sections of hippocampus vary in different ways among the sampling population.
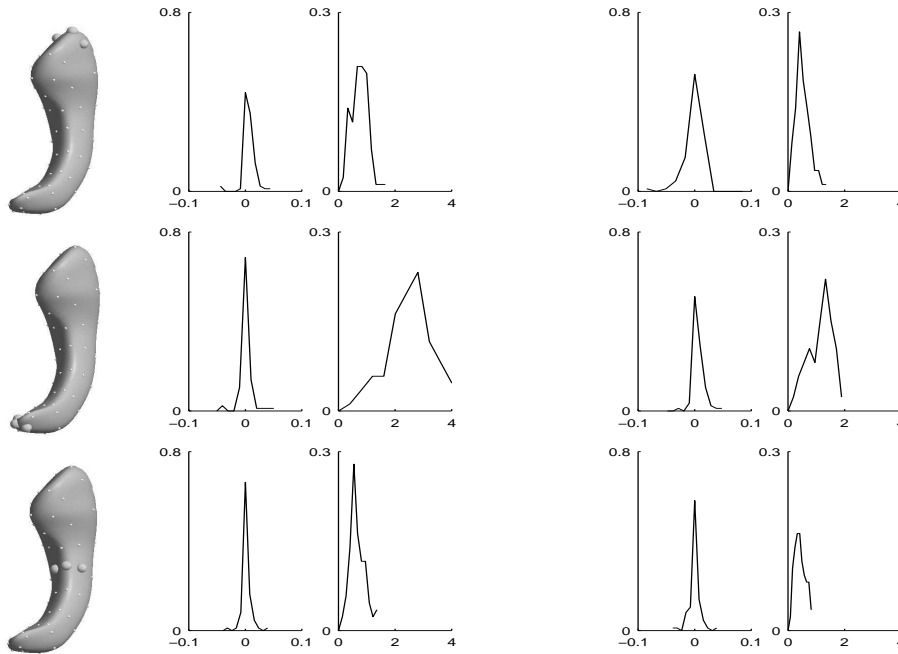

Fig. 4: Atom residue statistics of hippocampi. The big dots on the leftmost surface indicates the boundary section determined by the atom whose statistics is shown. In each row, the left two plots show histograms of $\log(\Delta r)$ and $\|\Delta\mathbf{x}\|$, the right two show histograms of $\log(\Delta r) - \overline{\log(\Delta r)}$ and $\|\Delta\mathbf{x} - \overline{\Delta\mathbf{x}}\|$, where $\overline{\log(\Delta r)}$ and $\overline{\Delta\mathbf{x}}$ are weighed averages of $\log(\Delta r)$ and $\Delta\mathbf{x}$ of the atom's neighbors, respectively.

These residue statistics are used to estimate the parameters of the corresponding MRF model.

7   Conclusions

Characterizing the geometric conformation of an object complex requires describing the geometry at multiple scale levels, so that geometric features with various degrees of locality can be described in a systematic way. We have been

developing a multi-scale shape analysis framework based on medial representations. A geometric entity is represented at the following scale levels: an object complex, individual objects, different object parts (figures), object sections (medial atoms), and boundary points. Some of these levels are directly natural and intuitive, e.g. the object and boundary levels. The representations of other levels, such as the description of object topology (figural relations) and the sampling density of medial atoms, are more meaningful and informative if they are based on a population of object instances. This is because essentially shape is not an attribute of a single object, but rather that of a class of objects.

A successive refinement of geometric representation is obtained by describing residual shape information across scales. The complete representation can be recovered from these residues. In doing so features at different scales are revealed. The geometric conformation of a population of object complexes is represented most effectively by probabilistic models, which incorporate both the common geometry and the deformations that occur among the population.

We have discussed how to build multi-scale probabilistic shape models based on both inter-scale and intra-scale residues. The relationships among them are modelled by Markov random fields, whose neighborhood structures provide a mechanism for describing local features at each scale level. We also showed how Markov random field models can be designed to describe residual geometric information on both the boundary level and medial atom level. The advantage of this approach is that it describes the rather complicated geometric information effectively by a relatively small number of parameters. Moreover, the models can be tuned based on the statistics of a training data set, thus the same type of model can be used to describe different classes of geometric entities.

We are going to apply the same idea to modelling other scale levels, so that complete multi-scale probabilistic shape models can be established. These models can be used as priors in a variety of shape analysis applications, such as image segmentation and object classification. Using maximum posterior approaches, the resulting probability distributions can be used as priors for segmentation [15] or as class probabilities in discrimination of objects by their geometry [25].

References

1. Witkin, A.: Scale-space filtering. In: Proc. Intl. Joint Conf. on Artificial Intelligence, Kalsruhe, Germany (1983) 1019–1023
2. Koenderink, J.: The structure of images. Bio. Cybern. **50** (1984) 363–370
3. Lindeberg, T.: Scale-space theory in computer vision. Kluwer Academic Publishers (1994)
4. Cao, F.: Morphological scale space and mathematical morphology. In: Scale-Space '99. Volume 1682 of LNCS. Springer-Verlag (1999) 164–174
5. Lindeberg, T.: Linear spatio-temporal scale-space. In: Proc. 1st Int. Conf. Scale-Space Theory in Computer Vision. Volume 1252 of LNCS. Springer-Verlag (1997) 113–127
6. Bruce, J., Giblin, P., Tari, F.: Families of surfaces: height functions and projections to planes. Math. Scand. **82** (1998) 165–185

7. Bruce, J., Giblin, P., Tari, F.: Families of surfaces: focal sets, ridges and umbilics. Math. Proc. Cambridge Philos. Soc. **125** (1999) 243–268

8. Lu, C., Cao, Y., Mumford, D.: Surface evolution under curvature flows. J. Visual Communication and Image Representation **13** (2002) 65–81

9. Damon, J.: Scale-based geometry for nondifferentiable functions, measures, and distributions. Parts I-III. (Preprint)

10. Kimia, B.B., Tannenbaum, A.R., Zucker, S.W.: Shapes, shocks, and deformations I: the components of two-dimensional shape and the reaction-diffusion space. Int. J. Comput. Vision **15** (1995) 189–224

11. Siddiqi, K., Bouix, S., Tannenbaum, A.R., Zuker, S.W.: Hamilton-Jacobi skeletons. Int. J. Computer Vision **48** (2002) 215–231

12. Olver, P., Sapiro, G., Tannenbaum, A.: Invariant geometric evolutions of surfaces and volumetric smoothing. SIAM J. Appl. Math. **57** (1997)

13. Mallat, S.G.: Multifrequency channel decompositions of images and wavelet models. IEEE Trans. Acoust. Speech, Signal Processing **37** (1989) 2091–2110

14. Unser, M.: A review of wavelets in biomedical applications. Proceedings of the IEEE **84** (1996) 626–638

15. Pizer, S.M., Chen, J.Z., Fletcher, P.T., Fridman, Y., Fritch, D.S., Gash, A.G., Glotzer, J.M., Jiroutek, M.R., Joshi, S., Lu, C., Muller, K.E., Thall, A., Tracton, G., Yushkevich, P., Chaney, E.L.: Deformable m-reps for 3D medical image segmentation. Int. J. Computer Vision (2003) (To appear).

16. Joshi, S., Pizer, S., Fletcher, P.T., Yushkevich, P., Thall, A., Marron, J.S.: Multi-scale deformable model segmentation and statistical shape analysis using medial descriptions. IEEE-TMI **21** (2002)

17. Styner, M., Gerig, G.: Medial models incorporating object variability for 3D shape analysis. In: IPMI '01. Volume 2082 of LNCS. Springer (2001) 502–516

18. Zhu, S.C.: Embedding Gestalt laws in the Markov random fields – a theory for shape modeling and perceptual organization. IEEE Trans. Pattern Anal. Mach. Intell. **21** (1999)

19. Lu, C.: Curvature-based multi-scale shape analysis and stochastic shape modeling. PhD thesis, Brown University (2002)

20. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Fifth European Conference on Computer Vision. (1998) 484–498

21. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - their training and application. Computer vision and image Understanding **61** (1995) 38–59

22. Kelemen, A., Szekely, G., Gerig, G.: Three-dimensional model-based segmentation. IEEE-TMI **18** (1999) 828–839

23. Blum, H., Nagel, R.: Shape description using weighted symmetric axis features. Pattern Recognition **10** (1978) 167–180

24. Thall, A.: Fast C^2 interpolating subdivision surfaces using iterative inversion of stationary subdivision rules. Technical Report TR02-001, University of North Carolina, Dept. of Computer Science (2002)

25. Gerig, G., Styner, M., Shenton, M., Lieberman, J.: Shape versus size: improved understanding of the morphology of brain Structures. In: Proc. MICCAI 2001. Volume 2208 of LNCS. Springer (2001) 24–32

26. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE T-PAMI **6** (1984) 721–741

27. Fletcher, P., Lu, C., Joshi, S.: Statistics of shape via principal component analysis on Lie groups. Technical Report TR02-043, University of North Carolina, Department of Computer Science (2002)