Multi-scale Shape Modeling by Markov Random Fields

Conglin Lu, Stephen M. Pizer, and Sarang Joshi

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

Abstract. The geometric conformation of an object complex is usually described by certain geometric features, which are most intuitive and provide locality if each is chosen to be at a restricted range of scale. A complete representation of a geometric entity, such as an object, includes descriptions at multiple scale levels. We have been developing a multi-scale framework for describing 3D geometric entities based on medial representations. At each scale level, an entity is represented by a collection of geometric primitives that describes a residue from the information provided at larger scales. The local geometric variations within a scale level are reflected by the difference in configurations of individual primitives. The primitives can be identified as elements of appropriate transformation groups. To characterize the common geometry of a population of object complexes as well as the variation of instances among the population, we build Markov random field (MRF) models based on both inter-scale and intra-scale residues, which are described in terms of the metrics on primitive transformations. In this paper, we present how to describe residues and 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 object complexes plays an important role in many medical imaging applications. For instance, geometric information can be incorporated as priors to guide image segmentation. By imposing geometric constraints, one obtains more reliable interpretation of images, as opposed to making decisions based solely on image intensity [1]. As another example, object discrimination usually involves characterizing and comparing classes that differ in their geometric conformation, e.g., their volume or their shape [2].

A complete representation of geometric entities should be able to describe them at multiple scale levels, so that geometric features with various degrees of locality, i.e., locally relevant size and distance, can be extracted. In our study of geometric entities such as objects and boundaries, locality must be taken with respect to the components of which an entity is formed: the relevant sizes and distances for an object complex are determined by various objects in it, while

those for an object must be determined by the natural sections making up the object, and so on. At each scale level, each feature summarizes the geometric information contained in certain spatial domain of size relevant to that scale. The differences in description between scale levels reflect different levels of detail. Also, the relevant distance within a scale level induces 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. This neighbor relation, together with the spatial extent of features and levels of detail in description, realize the notion of locality. Based on this view, we use object-based scale levels and describe *residues* between scales as in the wavelet approaches [3, 4]. Thus we describe changes in geometry across scales rather than geometric features prominent at multiple scales. This yields a hierarchical, multi-scale description of geometric entities, in which an entity at one scale level is seen relative both to the next larger level and relative to its neighbors at that level. As we will show, using this approach one can design shape models that are rich in geometric information yet relatively easy to work with, and hence achieve efficiency both in training the models and in applying them to shape analysis tasks such as segmentation.

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 medial representation , we need to take into account the common structure as well as the variations among different instances, so that both of these pieces of information can be effectively described. The fixed topology enables one to establish fixed correspondences between geometric primitives across the population. As a result, what differs among members of the population is the quantitative, geometric parameters and not qualitative properties of structure or topology.

We have been developing a methodology for describing 3D entities using medial representations combined with boundary displacements, which together form a representation called *m-reps* [1, 5]. In this framework, a geometric entity is represented at discrete scales and locations. At each scale it is described by a set of geometric primitives and their relative transformations. Determining the fixed topology from a population is achieved using the method described in [6].

A population of entities with similar geometric conformation can be effectively described by probabilistic models. For instance, principal component analysis [7, 8] and spherical harmonic descriptors [9] of 3D objects 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 measure are estimated and learned from a training data set. In our framework, the relationships between adjacent scale levels and among intra-scale neighbors make Markov random fields (MRF) the natural probabilistic models. We define MRF m-reps models so that they incorporate both inter-scale and intra-scale residue

information, which are expressed in terms of the appropriate metrics induced by the geometric primitives. With this approach, a reasonably small number of parameters are sufficient in describing the model at each scale level, so accurate parameter estimation can be achieved with limited numbers of training cases. In this paper, we describe MRF models on two scale levels, namely the boundary displacement level and the object section level.

In what follows, we briefly describe m-reps in section 2. Details of the MRF models are presented in section 3. Section 4 shows some results on estimating MRF models for hippocampi. We finish with some discussion in section 5.

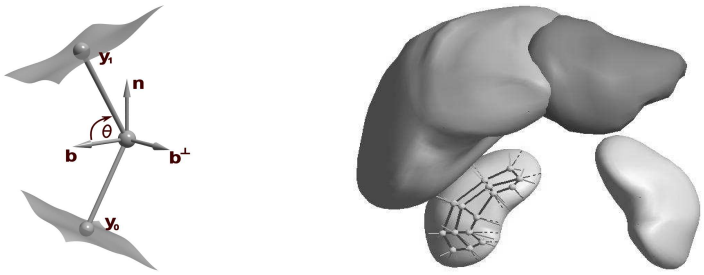
2 Multi-scale Shape Representation by M-reps

Medial-based representations [10, 11] provide a method of explicitly describing geometric deformations such as elongation, bending, and widening. To obtain stable medial and boundary structures, it is important to build them in a multi-scale fashion, including a boundary displacement component. In our framework, called m-reps, at all but the boundary scale level, an object is described by a set of continuous medial manifolds, which is sampled to yield discrete representations. Each sample point is called a medial atom, which describes a through section of an object (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 (defined below);
- a rotation $\mathbf{R} \in \mathrm{SO}(3)$, describing a local orthonormal frame $(\mathbf{n}, \mathbf{b}, \mathbf{b}^\perp)$, where $\mathbf{n}$ is the normal to the medial manifold, $\mathbf{b}$ is 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 $\theta \in SO(2)$, called the object angle, which determines the angulation of the implied boundary sections relative to $\mathbf{b}$.

As such, each medial atom can be identified as an element of the product group $G = \mathbb{R}^3 \times SO(3) \times \mathbb{R}^+ \times SO(2)$. The two implied boundary points are specified as $\mathbf{y}_i = \mathbf{x} + r\mathbf{n}_i, i = 0, 1$, where $\mathbf{n}_0$ and $\mathbf{n}_1$ are the two respective surface normals given by $\mathbf{n}_{0,1} = \cos(\theta)\mathbf{b} \pm \sin(\theta)\mathbf{n}$.

An m-rep figure is a quadrilateral mesh of medial atoms, with spacing determined through the analysis of the training population [6]. It describes a slab-like object or object part. The 4-adjacency in the mesh determines the atom neighbor relationship. A smooth boundary surface of a figure is generated by a subdivision surface algorithm [12] that approximates the boundary positions and normals implied by each atom. Objects are generally represented by a linked figural model, together with boundary displacements. A main figure describes the main section of an object; various subfigures, each of which described by a single medial sheet, represent different branches, protrusions or indentations. Finally, an object complex is described by the configurations of individual objects. The inter-figure and inter-object relations can be effectively described by m-reps,



(a) (b)

Fig. 1: M-reps. (a) A medial atom with a cross-section of the boundary surface it implies; (b) A liver-kidney complex, including two single-figure kidney objects and a liver object with a main figure (light) and a subfigure (dark). Each figure is represented as a quadrilateral mesh of medial atoms.

since solid 3D regions and their boundaries are represented simultaneously. Fig. 1(b) illustrates an example of this hierarchy of representations.

In this multi-scale framework, residual geometric information between scale levels is specified by the appropriate residue transformations of the corresponding geometric primitives. Within a scale each primitive has neighbors of the same type. See Table 1 for a brief summary. A successively refined boundary representation can be derived as one goes from coarse to fine scale levels. At the finest level, each boundary point moves from its medially implied position along the medially implied normal direction to "fine tune" the description.

Scale level	**Primitive**	**Inter-scale residue transformation**	**Intra-scale neighbors**
Object complex	Object complex pose	Similarity	—
Object	Object pose	Similarity, relative to object complex level	Adjacent objects
Figure	Main figure and subfigure poses	Main figure and subfigure transformations, relative to object level	Adjacent figures
Medial atom	Atom configuration	Atom transformation, relative to figure level	Adjacent medial atoms
Boundary	Boundary vertex position	Displacement from medially implied position	Adjacent boundary vertices

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

The m-rep framework allows geometric features at different positions and scale levels to be explicitly described. Furthermore, the medial structure, determined by a training population, provides a multi-scale intrinsic coordinate system [13] that is extremely well suited for statistical analysis of shapes, because correspondence among a population can be established systematically.

3 Markov Random Field M-reps Models

Given that any model has a fixed topology, we determine a probability distribution on the space of random variables characterizing the geometry of primitives. The total number of such random variables is usually very high. The Markov random field approach [14, 15] handles this by characterizing global geometric information through local interactions among various geometric primitives.

An MRF model characterizes a collection of random variables, the relationship among which is described by a simply connected *dependency graph* $G = (V, E)$. Each vertex $v \in V$ corresponds to a random variable X_v in the model. The *neighborhood* of v, denoted by $\mathcal{N}(v)$, contains those vertices that are connected to v via an edge in E. 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 its conditional probabilities satisfy

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

Also, P is said to be a Gibbs distribution with respect to G if the joint probability density of $\{X_v\}$ has the form

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

where $\mathcal{C}$ is the set of cliques of G, $X_C = \{X_v : v \in C\}$, Θ is a set of parameters, $Z(\Theta)$ is a normalizing constant (depending on Θ). Each $A_C \geq 0$ is called a *potential* and depends only on those random variables whose indices are in C. The Hammersley-Clifford Theorem [14, 15] 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, which can then be learned from a training data set.

In m-reps, there are natural neighbor relations between and within scale levels. Suppose the scale levels are indexed by $1, 2, \ldots, l$, scale 1 being the coarsest. Let $\mathbf{z}^k = \{\mathbf{z}_j^k\}$ denote the collection of geometric primitives at level k. Each primitive $\mathbf{z}_j^k$ has a value implied by the previous level, which is called the parent primitive of $\mathbf{z}_j^k$ and denoted by $\mathcal{P}(\mathbf{z}_j^k)$. 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 $\Delta\mathbf{z}_j^k$ denote the inter-scale residual giving the *difference* between $\mathbf{z}_j^k$ and $\mathcal{P}(\mathbf{z}_j^k)$, 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. At level k, we describe residual geometric information from level $k-1$ by $\Delta\mathbf{z}^k = \{\Delta\mathbf{z}_j^k\}$, with a Markov assumption

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

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

The residue probability distributions $\{P_k(\Delta \mathbf{z}^k)\}$ are defined as MRF models, with respect to the canonical neighborhood structure induced by the spatial relationship among primitives. For example, at the medial atom level, 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 an edge in the quad-mesh (refer to Fig. 2(b)). If another sampling mesh, e.g. triangular mesh, is used, then appropriate canonical neighborhood structure can be induced similarly.

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

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

$$p_k(\{\Delta \mathbf{z}^k\}) \propto \exp\{-\textstyle\sum_{C \in\ \mathcal{C}} A_C(\Delta \mathbf{z}^k_C)\}, \tag{1}$$

where $\mathcal{C}$ is the set of cliques, and $\Delta \mathbf{z}^k_C = \{\Delta \mathbf{z}^k_j : j \in C\}$. Our approach to designing these models is to identify the residue primitives $\Delta \mathbf{z}^k_j$ as elements of appropriate transformation groups, and to define the potentials A_C in terms of the metrics on these groups. In the rest of the section, we discuss these issues on the boundary level and the object section level within the m-reps framework.

3.1 MRF Models for Boundary Displacement

For an m-rep figure, the medially implied surface is represented by a dense set of boundary points, which are the geometric primitives at the boundary level. They are parameterized by an object-intrinsic coordinate system induced by the medial manifold. Associated with each medially implied boundary point $\mathbf{y}$ are a radius $r(\mathbf{y})$, which is the distance between $\mathbf{y}$ and the corresponding medial point, and a surface normal vector $\mathbf{n}(\mathbf{y})$ at $\mathbf{y}$.

At the boundary level, each medially implied boundary point $\mathbf{y}$ moves along $\mathbf{n}(\mathbf{y})$, yielding a finer scale description. To maintain magnification invariance, we measure displacement in multiples of object width, i.e., if the amount of movement of the $\mathbf{y}$ is $d(\mathbf{y})$, then we define the displacement of $\mathbf{y}$ to be the unitless variable $w(\mathbf{y}) = d(\mathbf{y})/r(\mathbf{y})$, which is a member of the one-dimensional additive group $\mathbb{R}$. The displacement field $\mathbf{w} = \{w_j\}$ on $\{\mathbf{y}_j\}$ provides the residual geometric information at this scale level.

Currently we use a quad-mesh to sample the boundary, thus the canonical neighborhood structure is the 4-adjacency structure in Fig. 2. With respect to this graph, we define the boundary level model as a zero-mean Gaussian MRF model on the displacement field with density

$$p_q(\mathbf{w}) = \frac{1}{Z(q_1,q_2)} \exp\Big\{ -\frac{q_1}{2}\sum_i s_i w_i^2 - \frac{q_2}{2}\sum_{<i,j>} b_{ij}(w_i - w_j)^2 \Big\}, \qquad (2)$$

where $< i, j >$ denotes that points i and j are neighbors, $q_1, q_2 > 0$ are parameters, and $Z(q_1, q_2)$ is a normalizing constant. We choose $\{s_i, b_{ij}\}$ so that the exponent above is a discrete approximation of the energy function

$$-\tfrac{q_1}{2}\int_{\mathcal{S}} \tfrac{d^2(\mathbf{y})}{r^2(\mathbf{y})}\mathrm{d}\mathbf{y} - \tfrac{q_2}{2}\int_{\mathcal{S}} \|\nabla d(\mathbf{y})\|^2 \mathrm{d}\mathbf{y}.$$

The model (2) incorporates both inter-scale and intra-scale residues at the boundary level. Notice that the conditional density for a particular w_i is

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

This can be interpreted as putting penalties on the amount of w_i as well as on 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 (2) with different parameter values, which reflect the variation of boundary displacement field in various sections.

The parameters q_1 and q_2 in (2) can be estimated by the maximum likelihood method. Given M independent training objects with displacement fields $\widehat{\mathbf{w}}_1, \widehat{\mathbf{w}}_2, \ldots, \widehat{\mathbf{w}}_M$, and assuming $\widehat{\mathbf{w}}_i$ is a sample from the distribution with density $p_q^{(i)}$, we seek the parameter values q_1, q_2 such that the likelihood function

$$L(q_1, q_2) = \textstyle\sum_{i=1}^{M} \log\big(p_q^{(i)}(\widehat{\mathbf{w}}_i)\big)$$

is maximized. 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 the gradient $\nabla L(q_1^*, q_2^*) = \mathbf{0}$. This equation is solved numerically.

3.2 MRF Models for Object Sections

Object sections are described at the atom scale level, where the primitives are the medial atoms $\{\mathcal{A}_i\}$. The parent primitives $\{\mathcal{P}(\mathcal{A}_i)\}$ are the medial atoms describing the object at the previous larger scale, the figural scale. The residue geometric information is described by the differences between $\{\mathcal{A}_i\}$ and $\{\mathcal{P}(\mathcal{A}_i)\}$. As discussed in section 2, a medial atom $\mathcal{A}_i$ is characterized by a 4-tuple $(\mathbf{x}_i, \mathbf{R}_i, r_i, \theta_i)$. Suppose $\mathcal{P}(\mathcal{A}_i) = (\widetilde{\mathbf{x}}_i, \widetilde{\mathbf{R}}_i, \widetilde{r}_i, \widetilde{\theta}_i)$. We define the atom residue to be

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

It is an element of the group $G = \mathbb{R}^3 \times SO(3) \times \mathbb{R}^+ \times SO(2)$. Let $d_R(\cdot,\cdot)$ and $d_2(\cdot,\cdot)$ be the Riemannian distance on $SO(3)$ and $SO(2)$, respectively, with the corresponding norms denoted by $\|\cdot\|_R$ and $\|\cdot\|_2$. The distance d_G, with corresponding norm denoted by $\|\cdot\|_G$, is defined to be

$$d_G(\Delta\mathcal{A}_i, \Delta\mathcal{A}_j) = \sqrt{\|\Delta\mathbf{x}_i - \Delta\mathbf{x}_j\|^2 + d_R^2(\Delta\mathbf{R}_i, \Delta\mathbf{R}_j) + |\ln(\Delta r_i/\Delta r_j)|^2 + d_2^2(\Delta\theta_i, \Delta\theta_j)}. \quad (3)$$

We assume the MRF model for atom residues with respect to the canonical 4-neighbor structure has a density of the Gibbs form

$$p\big(\{\Delta\mathcal{A}_i\}\big) \propto \exp\Big\{ -\sum_i \frac{\sigma_i}{2}\|\Delta\mathcal{A}_i\|_G^2 - \sum_{<i,j>} \frac{\tau_{ij}}{2} d_G^2\big(\Delta\mathcal{A}_i, \Delta\mathcal{A}_j\big)\Big\}, \quad (4)$$

where $< i, j >$ denotes that atom i and atom j are neighbors, $\{\sigma_i, \tau_{ij}\}$ are positive parameters. This density is with respect to the invariant measure on G. The conditional distribution of $\Delta\mathcal{A}_i$ given the rest of the residues has density

$$p(\Delta\mathcal{A}_i|\Delta\mathcal{A}_{\{j\neq i\}}) \propto \exp\Big\{ -\frac{\sigma_i}{2}\|\Delta\mathcal{A}_i\|_G^2 - \sum_{<i,j>} \frac{\tau_{ij}}{2} d_G^2\big(\Delta\mathcal{A}_i, \Delta\mathcal{A}_j\big)\Big\}.$$

Intuitively, the first term in the exponent penalizes the difference between $\mathcal{A}_i$ and its parent $\mathcal{P}(\mathcal{A}_i)$, 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 $\{\Delta\mathcal{A}_j : j \neq i\}$. Clearly, the model incorporates information on both inter-scale and intra-scale residues at the atom level.

Given a training data set, the parameters $\{\sigma_i, \tau_{ij}\}$ of the probability model (4) 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. Notice that the Riemannian distance between two 3D rotations $\Delta\mathbf{R}_1$ and $\Delta\mathbf{R}_2$ is given by $\|\mathrm{Log}(\Delta\mathbf{R}_1^{-1}\Delta\mathbf{R}_2)\|_F/\sqrt{2}$, where $\|\cdot\|_F$ is the Frobenius matrix norm, and for $\mathbf{R} \in SO(3)$,

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

Here ϕ satisfies $\mathrm{tr}(\mathbf{R}) = 1 + 2\cos\phi$ and $|\phi| < \pi$. 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 $\|\mathrm{Log}(\Delta\mathbf{R}_2) - \mathrm{Log}(\Delta\mathbf{R}_1)\|_F$. Similarly, the distance between $\Delta\theta_1$ and $\Delta\theta_2$ in $SO(2)$ is $|\Delta\theta_1 - \Delta\theta_2|$ when they are both close to 0. Define an invertible map L by

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

For $\Delta\mathcal{A}_1, \Delta\mathcal{A}_2$ close to the identity of G, the distance d_G defined in (3) can be approximated by the distance $d_{\mathfrak{g}}$ on the linear space $\mathfrak{g}$:

$$\begin{aligned} d_G^2(\Delta\mathcal{A}_1, \Delta\mathcal{A}_2) &\approx d_{\mathfrak{g}}^2(\Delta\mathcal{L}_1, \Delta\mathcal{L}_2) \\ &= \|\Delta\mathbf{x}_1 - \Delta\mathbf{x}_2\|^2 + \frac{1}{2}\|\mathrm{Log}(\Delta\mathbf{R}_1) - \mathrm{Log}(\Delta\mathbf{R}_2)\|_F^2 \\ &\quad + |\ln(\Delta r_1) - \ln(\Delta r_2)|^2 + |\Delta\theta_1 - \Delta\theta_2|^2. \end{aligned}$$

Now instead of the model (4), we define a probability distribution on $\{\Delta\mathcal{L}\}$ with density

$$p\big(\{\Delta\mathcal{L}_i\}\big) = \frac{1}{Z}\exp\Big\{-\sum_i \frac{\sigma_i}{2}\|\Delta\mathcal{L}_i\|_{\mathfrak{g}}^2 - \sum_{<i,j>}\frac{\tau_{ij}}{2}d_{\mathfrak{g}}^2\big(\Delta\mathcal{L}_i, \Delta\mathcal{L}_j\big)\Big\}, \quad (5)$$

where $\|\cdot\|_{\mathfrak{g}}$ is the norm corresponding to $d_{\mathfrak{g}}$. (5) induces a probability distribution on G via L^{-1}, which takes each $\Delta\mathcal{L}$ back to $\Delta\mathcal{A}$.

The model (5) is essentially a zero-mean Gaussian model on the linear space $\mathfrak{g}$. Again, the parameters $\{\sigma_i, \tau_{ij}\}$ can be estimated from a training data set by the maximum likelihood principle, but in this case we may avoid MCMC methods and estimate the parameters directly.

It should be pointed out that two of the components of $\Delta\mathcal{L}$, namely $\mathrm{Log}(\Delta\mathbf{R})$ and $\Delta\theta$, lie in bounded domains, so the Gaussian distribution in (5) is not exactly the desired distribution. However, it is a close approximation, because typically the values of those components are very close to the origin, and the Gaussian measures of the complements of the bounded domains are negligible.

In models (4) and (5), the variations of atom residues are essentially measured by their norms. There is no distinction between individual components, i.e., translation, rotation, etc. In case when distinctions have to be made, we need to modify either the definition of the norm or the specification of the models by, for example, putting appropriate weights on the components.

The idea of approximating the group G by a linear space $\mathfrak{g}$ can be formalized with Lie group theory. In fact, G is a Lie group, $\mathfrak{g}$ is the corresponding Lie algebra, and the mapping L^{-1} is the exponential map. The details are beyond the scope of this paper, but can be found in [16] and the references therein. Also, Gaussian distributions on the Lie group G are discussed more formally in [17].

4 Experiment on Hippocampi

In this section we present some results on learning the MRF models (2) and (5) for hippocampi. The training data set contains 86 binary segmented images for left hippocampi. A common hippocampus model, represented as a single m-rep figure by a 8×3 array of medial atoms, is deformed into each image. Fig. 3 shows a few sample models. The indexing of the atoms are such that the 1-st row is at the tail of a hippocampus, and the 8-th row is at the head.

In each case, a global transformation is applied to the common model first, then each individual atom undergoes a further deformation to fit the image. The

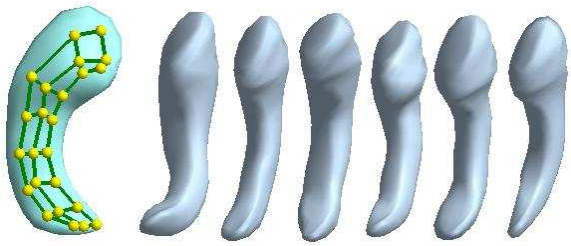
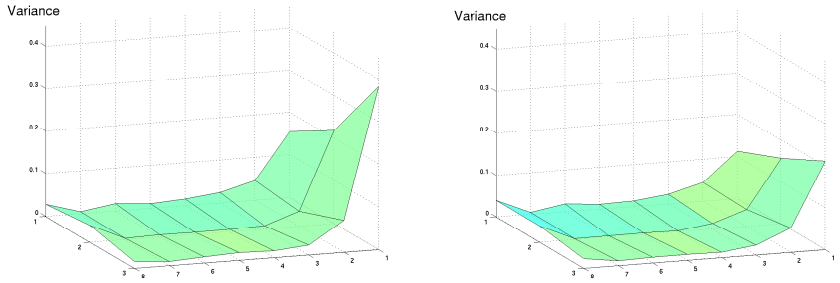


Fig. 3: Sample hippocampi. The leftmost one shows the medial mesh. In the picture the top part of each hippocampus is the "head", and the bottom part is the "tail".

atom residues are the differences in atom configurations between the two stages. They are used to obtain the ML estimates for the parameters $\{\sigma_i, \tau_{ij}\}$ in the MRF model (5). Fig. 4 shows both the sample and the estimated variances of the atom residues. We can see that the tail has the most variation, and the middle section has the least.



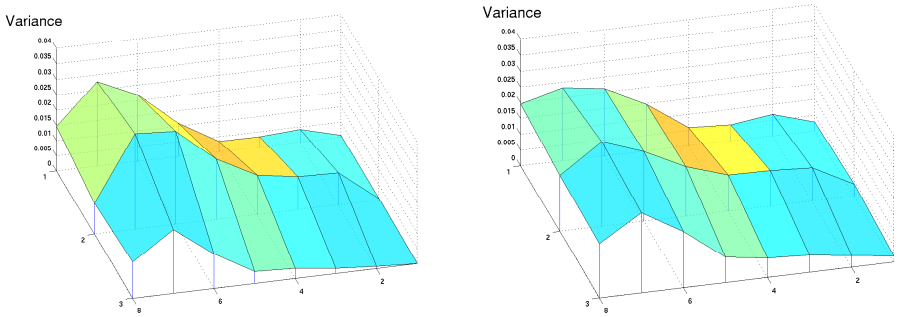
(a) Sample variances (b) Estimated variances

Fig. 4: Atom residue variation for hippocampi. Each grid point corresponds to a medial atom. The tail has the highest variance.

At the boundary scale, the medially implied boundary points of each object are displaced, yielding a displacement field. Each medial atom describes two opposite parts of the boundary surface (see Fig. 1(a)). In this way we divide the surface of a hippocampus into 48 sections, 24 on each side of the medial manifold. We assume each section is described by the MRF model (2) with appropriate parameters q_1, q_2, which are estimated based on the displacement fields. Fig. 5 shows both the sample and the estimated variance of displacements on one side of the hippocampus surface. The head part has the most variance in this case.

5 Discussion

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



(a) Sample variances (b) Estimated variances

Fig. 5: Boundary displacement variation for one side of hippocampi by sections. Each grid point represents a section. The head has more variation.

developing a multi-scale shape analysis methodology based on medial representations. In this framework, 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 and positions can be explicitly described.

We have discussed how to build multi-scale probabilistic shape models based on both inter-scale and intra-scale residues. These residues are incorporated into Markov random fields, whose neighborhood structures provide a mechanism for describing geometric features at different scale levels and positions. We also showed how MRF models can be designed to describe residual geometric information on both the boundary level and medial atom level. The basic idea is to identify the geometric primitives as elements of appropriate transformation groups, and to specify the potentials of the MRF models in terms of the metrics on these groups. These models can be tuned based on the statistics of training data, thus the same type of model can be used to describe different classes of geometric entities. The advantage of this approach is that it describes the rather complicated geometric information effectively by a relatively small number of parameters, so training and applying these models can be done efficiently.

Our basic assumption is that the complete configuration of geometric entities can be effectively captured by residues and described by MRF models with simple, local neighbor structures. Empirical evidence shows that this is a reasonable assumption in a variety of situations, although whether this is the case in general remains to be seen. Of course, one can always design MRF's with more complicated and global neighborhood structures, but learning and applying these models become more difficult accordingly.

We are going to extend the same idea to modelling other scale levels, so that complete multi-scale probabilistic shape models can be obtained. Using maximum posterior approaches, the resulting probability distributions can be applied as priors to segmentation or as class probabilities to discrimination of objects by their geometry.

Acknowledgement

This work is supported by National Cancer Institute grant P01 CA47982.

References

1. 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).
2. 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
3. Mallat, S.G.: Multifrequency channel decompositions of images and wavelet models. IEEE Trans. Acoust. Speech, Signal Processing **37** (1989) 2091–2110
4. Unser, M.: A review of wavelets in biomedical applications. Proceedings of the IEEE **84** (1996) 626–638
5. Joshi, S., Pizer, S., Fletcher, P.T., Yushkevich, P., Thall, A., Marron, J.S.: Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. IEEE-TMI **21** (2002)
6. Styner, M., Gerig, G.: Medial models incorporating object variability for 3D shape analysis. In: IPMI '01. Volume 2082 of LNCS. Springer (2001) 502–516
7. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Fifth European Conference on Computer Vision. (1998) 484–498
8. 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
9. Kelemen, A., Szekely, G., Gerig, G.: Three-dimensional model-based segmentation. IEEE-TMI **18** (1999) 828–839
10. Blum, H., Nagel, R.: Shape description using weighted symmetric axis features. Pattern Recognition **10** (1978) 167–180
11. Siddiqi, K., Bouix, S., Tannenbaum, A.R., Zuker, S.W.: Hamilton-Jacobi skeletons. Int. J. Computer Vision **48** (2002) 215–231
12. Thall, A.: Fast C^2 interpolating subdivision surfaces using iterative inversion of stationary subdivision rules. (midag.cs.unc.edu/pub/papers/Thall_TR02-001.pdf)
13. Pizer, S.M., Fletcher, P.T., Thall, A., Styner, M., Gerig, G., Joshi, S.: Object models in multi-scale intrinsic coordinates via m-reps. Image and vision computing (2003) To appear.
14. Besag, J.: Spatial interaction and the statistical analysis of lattice systems. J. Royal Statistical Society, Series B **36** (1974) 192–236
15. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE T-PAMI **6** (1984) 721–741
16. Fletcher, P., Lu, C., Joshi, S.: Statistics of shape via principal component analysis on Lie groups. (midag.cs.unc.edu/pub/papers/CVPR03_Fletcher_StatsShape.pdf)
17. Fletcher, P.T., Joshi, S., Lu, C., Pizer, S.M.: Gaussian distributions on Lie groups and their application to statistical shape analysis. (Submitted to IPMI '03)