International Journal of Computer Vision 75(3), 387–404, 2007

DOI: 10.1007/s11263-007-0045-0

Statistical Multi-Object Shape Models

CONGLIN LU, STEPHEN M. PIZER, SARANG JOSHI AND JA-YEON JEONG
Medical Image Display & Analysis Group, University of North Carolina, Chapel Hill, USA
coa05lu@yahoo.com or lu@cs.unc.edu

Received July 18, 2005; Accepted February 15, 2007

First online version published in March, 2007

Abstract. The shape of a population of geometric entities is characterized by both the common geometry of the population and the variability among instances. In the deformable model approach, it is described by a probabilistic model on the deformations that transform a common template into various instances. To capture shape features at various scale levels, we have been developing an object-based multi-scale framework, in which a geometric entity is represented by a series of deformations with different locations and degrees of locality. Each deformation describes a residue from the information provided by previous steps. In this paper, we present how to build statistical shape models of multi-object complexes with such properties based on medial representations and explain how this may lead to more effective shape descriptions as well as more efficient statistical training procedures. We illustrate these ideas with a statistical shape model for a pair of pubic bones and show some preliminary results on using it as a prior in medical image segmentation.

Keywords: statistical shape models, multi-object shape models, deformable model, multi-scale shape analysis, medial representation

1. Introduction

In many shape analysis areas, including deformable models (Grenander, 1995; Kass et al., 1987), it is desirable to consider shape as a property of a *population* of geometric entities, such as object sections (e.g., the left lobe of livers), objects (e.g., livers), or groups of objects (e.g., liver-kidney complexes). As such, it describes both the common geometric features and the geometric variability among instances of the population. With the deformable model approach, a representative instance S_0, usually referred to as the *model* or *template*, is chosen for a given shape, say **S**. The geometry of S_0 is essentially specified by transformations on certain canonical geometric primitives, for instance, the positions of boundary/surface points or positions and boundary normal directions of medial structures. The overall transformation is a "summary" of the shape geometry and captures the typical geometric conformation of **S** at various levels of detail, including the geometry of objects, parts, sections, and their relative configurations. Any instance of **S**, say S, is then defined via a further geometric deformation F applied to S_0, i.e. $S = F(S_0)$. The map

$$\begin{aligned} \psi &: \mathcal{F} \to \mathcal{G} \\ F &\mapsto F(S_0) \end{aligned} \tag{1}$$

establishes a correspondence between the space of all such geometric transformations, $\mathcal{F}$, and the set of all geometric entities, $\mathcal{G}$. We can regard an instance of the shape S as a random sample from an underlying probability measure ν on G. Intuitively, ν puts soft constraints on "what an instance of the shape S should look like". Because of the correspondence ψ, the geometric variability among shape instances, which is captured by ν, can be effectively described via variability in geometric transformations applied to the template. One way to define ν, then, is to introduce a probability measure μ on $\mathcal{F}$ and to let ν be the probability measure induced by the mapping ψ. This is the approach we will adopt in this paper. The measure μ can be effectively used as shape priors to make statistical inferences on what one observes (Grenander et al., 1991; Zhu, 1999; Pizer et al., 2003a; Gerig et al., 2001).

Existing geometric representations include representation by dense sample points, landmarks, diffeomorphisms on displacement velocity fields (Joshi, 1997), distance functions or their level sets (Tsai et al., 2003), skeletons (Siddiqi et al., 2002), shocks (Kimia et al., 1995), multigrid and scale-space based methods (Lu et al., 2002), etc. To develop statistical models, the geometric representation also needs a notion of correspondence among different shape instances. For example, if we use boundary

points to describe the contour of a hand, then the thumbs in different hands should be described by the same set of points. Without such a correspondence the statistical model would not be robust. Examples of statistical shape models are deformable templates (Grenander et al., 1991), point distribution models (Cootes et al., 1998, 1995), spherical harmonic descriptors (Kelemen et al., 1999) for 3D objects, to name a few. Most of these methods do not respect object-based scale levels, so they lack the ability to explicitly describe important geometric information such as locality, inter-object relations, etc. The statistical models typically face the high-dimension-low-sample-size (HDLSS) problem: on the one hand, the number of parameters needed to accurately describe the geometry is usually very large; on the other hand, the available training samples are often limited, as is the case in many medical imaging applications.

Many working with geometric probabilities in medical image analysis have dealt with multi-object shape statistics by concatenating the descriptors for all the objects and doing global statistics on the resulting tuple. Examples are Cootes et al. (1999) and Tsai et al. (2003). These do not have the locality of the object or smaller. A few investigators have produced statistics that are hierarchical by scale and include a scale describing the individual objects making up the complex but not having the smaller scale descriptions be residues from other scales (Vaillant and Davatizikos, 1999; Kapur et al., 1998), but Davatzikos et al. (2003) has taken this approach using residues. However, objects change not only on their own but due to their neighbors, and yet when describing an object statistically, few have attempted to describe the inter-relationships between objects statistically. Pohl et al. (2005) have attempted to describe inter-relationships. They represent these via the distance function to objects' boundaries, to which, they agree, linear statistical methods are not well-suited. Moreover, they have not described inter-relationships in a hierarchical fashion, or they have used hierarchy but no statistical inter-relationships (Pohl et al., 2005). The approach we attempt is simultaneously hierarchical by scale using inter-scale residues and describes objects not only by themselves but also in terms of their inter-relationships with other objects.

In order to efficiently characterize a population of geometric entities, it is important to realize that geometric deformations have various *localities* and we believe they should be incorporated in shape description (Yushkevich et al., 2001). In our study of geometric entities such as objects and boundaries, locality is relative to object scale level and must be taken with respect to the components of which an entity is formed. For example, at the object complex level, a global deformation is one that applies to *all* objects in it, while a local deformation applies only to some of the objects (but as a whole); at the individual object level, being a local deformation means it applies only to certain natural sections making up the object. The geometric features described by the deformations at each object scale level thus have certain sizes and distances that are relevant to that level. The differences in description between scale levels reflect different levels of detail. Moreover, 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. This neighbor relation, together with the spatial extent of features and levels of detail in description, realize the notion of locality and form the basis of our representation. In the context of describing a population of entities, locality reflects the spatial correlation between various geometric components of different scales among the population.

This notion of locality suggests we use object-based scale levels and describe *residues* within and between scales similar to the wavelet approaches (Mallat, 1989; Unser, 1996). We argue that it is more advantageous not to describe a geometric entity $\mathcal{S}$ by a single transformation F, bur rather represent it in a coarse to fine manner by decomposing F into a series of, say K, deformations, i.e., $F = F_K \circ F_{K-1} \circ \cdots \circ F_1$ (Pizer et al., 2005). In this way geometric features at various object scale levels can be explicitly described by the corresponding transformations at that level. For each $k = 1, \ldots, K$, $\mathcal{S}_k = F_k \circ F_{k-1} \circ \cdots \circ F_1(S_0)$ is the representation of S at step k, and the deformation F_k describes a *residual* transformation relative to S_{k-1}, measured in terms of both across-scale and within-scale primitive relationships. In other words, the smaller scale features are described as *residues* of larger scale features. At any particular object scale level, we will assume that each residual feature tends to be more correlated with those that are close by, compared to those that are further away. Any strong correlation between entities that are not close to each other is a global feature and is supposed to have already been accounted for by the deformation at the (previous) larger scale level. This is an important assumption that we make in developing statistical shape models, an assumption based on measurements we have made in certain sample populations. Consider, for example, the shape of a pair of pubic bones (see Fig. 1(b)). In this case F_1 may be a single global level deformation that is applied to the both bones simultaneously, F_2 may represent a *further* deformation of the left bone, F_3 may represent a collection of even further local deformations of the left bone, each applied to a very small part of it, and so on.

In addition to providing more intuitive and accurate geometric descriptions, the above multi-step approach also has the potential of leading to more efficient probabilistic models. We believe that by taking advantage of locality one can effectively characterize each deformation by a much smaller number of parameters. Because

the process of estimating these parameters suffers a lot less from the HDLSS problem, it might be possible to obtain more reliable and stable estimations of statistical shape models from limited training samples.

A valid statistical shape model needs to have certain generative and predictive power. It should be able to capture most of the variations in a object population and produce instances that are similar to training samples and are *geometrically proper* (Pizer et al., 2005), meaning that these instances preserve topology, do not have geometric singularities and have no or very little inter-penetration among objects.

We have been developing a methodology for describing 3D entities using medial representations combined with voxel scale displacements of the object interior and boundary, which together form a representation called *m-reps* (Pizer et al., 2003a; Joshi et al., 2002). 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. For a population of geometric entities, we establish correspondence among instances by fixing the topology of our representation. As a result, what differs among members of the population is the quantitative, geometric parameters and not qualitative properties of structure or topology. One method for determining the fixed topology from a population is described in Styner and Gerig (2001).

We describe shapes at multiple object scale levels to provide locality and efficiency in statistical analysis. The relationships between adjacent scale levels and among intra-scale neighbors make Markov models the natural choice. In this framework, we learn a series of probability models, each of which can be effectively characterized by a reasonably small number of parameters and thus can be estimated with limited number of training cases.

In what follows, we briefly describe m-reps in Section 2. Details of the multi-object m-rep models are presented in Sections 3–5. Section 6 shows an example of using statistical shape models to segment pubic bones. We finish with some discussion in Section 7.

2. Multi-Scale Representation with M-reps

The approaches described in this paper may be used with a variety of object representations, but the properties of medial representations are helpful for realizing the approaches. Medial-based representations (Blum and Nagel, 1978; Siddiqi et al., 2002; Pizer et al., 2003b) provide a method of explicitly describing geometric deformations such as elongation, bending, and widening of object interiors. To obtain stable medial and boundary structures, it is important to build them in a multi-scale fashion, including a voxel scale interior and boundary displacement component. In our framework, called m-reps, at all but the voxel scale level, an object is described by a set of continuous medial manifolds, which are 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)), and it is at this locality that the geometric deformation can be applied. An internal atom (i.e., one that is not on the mesh boundary) is represented as a four-tuple $A = (p, r, U^{+1}, U^{-1})$ consisting of

- a translation $p \in \mathbb{R}^3$, specifying the position of the medial point; we can consider this translation in units of the medial width r (defined below);
- a magnification scalar $r \in \mathbb{R}^+$, defined as the distance from the medial point to the implied boundary points b^{+1}, b^{-1} (the local width);
- $U^{+1}, U^{-1} \in S^2$ are two unit vectors pointing from the medial position to the two implied boundary points b^{+1}, b^{-1}. Equivalently, they are the surface normal vectors at b^{+1} and b^{-1};

As such, in m-reps each internal medial atom can be identified as an element of the eight-dimensional space $G = \mathbb{R}^3 \times \mathbb{R}^+ \times S^2 \times S^2$. For atoms on the boundary of the mesh, there is an extra parameter (with value in $\mathbb{R}^+$. See Pizer et al. (2003b)). For simplicity, in this paper we assume all atoms are elements of G.

In our current scheme, an m-rep *figure* is a sheet of medial atoms represented by a quadrilateral mesh of atoms with spacing determined through the analysis of the training population (Styner and Gerig, 2001). It describes a slab-like object or object part. The four-adjacency in the mesh determines the atom neighbor relationship (see Fig. 1(b)). A smooth boundary surface of a figure is generated by an algorithm (we presently use a subdivision surface method (Thall, 2004)) that approximates the boundary positions and normals implied by each atom. An m-rep object corresponds to a geometric object and is generally represented by a linked figural model: a main figure describes the main section of an object; various sub-figures, each of which described by a single medial sheet, represent different branches, protrusions or indentations. Finally, an object complex is made up of individual objects. With m-reps, the inter-object, inter-figure, and inter-atom relations can be effectively described by appropriate sets of atom transformations, which are in turn describable by basic transformations such as translation, scaling, and rotation. Figure 1(b) illustrates an example of this hierarchy of representations for a pair of pubic bones.

The m-rep framework allows geometric features at different positions and scale levels to be explicitly described. Furthermore, the medial structure, which is determined by a training population, provides a multi-scale intrinsic coordinate system (Pizer et al., 2003c). This establishes

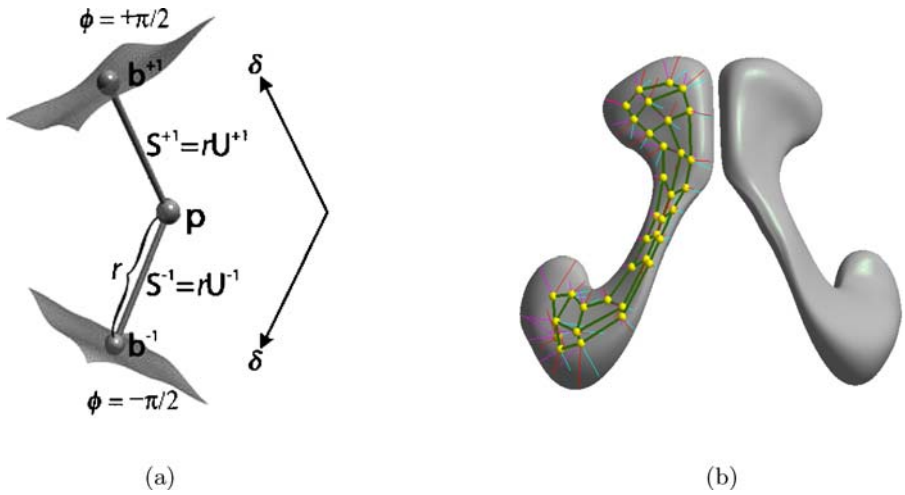


Figure 1. M-reps. (a) A medial atom with a cross-section of the boundary surface it implies; (b) The hierarchy made by the two-bone complex. Each bone is represented by one m-rep figure. Each figure contains 36 atoms (only those in the right bone are shown).

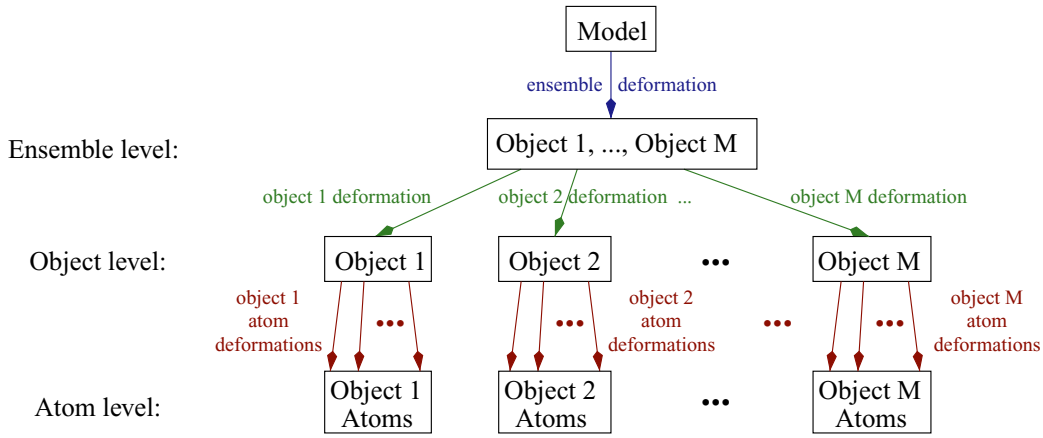


Figure 2. Hierarchical representation of an object ensemble consisting of single-figure objects.

correspondence among a population and makes m-reps extremely well suited for statistical analysis of shapes. The issue of how to choose the spacing of the medial atoms so that all the members of a population are in optimal correspondence is beyond the scope of this paper. The ideas in Merck et al. (2005) and Davies et al. (2002) are certainly relevant.

In this paper, we consider object ensembles that contain only single-figure objects. The hierarchy of representation of such a geometric entity instance is obtained by applying a series of deformations to the shape model (template), illustrated in the diagram in Fig. 2. After the atom level, there can be a separate boundary displacement level where each medially implied boundary point moves along its associated surface normal direction to "fine tune" the representation. We will not discuss this step in detail here. A Markov random field model for boundary displacement fields can be found in Lu et al. (2003).

Notice that in Fig. 2, the number of medial atoms used at each level of representation does not change. For instance, object 1 is described by the same number of atoms at the ensemble, object, and atom level. In other words, this is not a type of multi-grid representation. Instead, the number of deformations that are applied increases as one goes from coarse level to fine level. For example, suppose for $i = 1, 2, \ldots, M$, object i has n_i atoms. Then at the ensemble level, there is one transformation that acts on all $\sum_{i=1}^{M} n_i$ atoms; at the object level, there are M different deformations; and at the atom level, there are $\sum_{i=1}^{M} n_i$ atom deformations. This set of deformations is what we need to model in order to characterize a shape.

3. Probabilistic Models for Object Ensembles

Let $\mathcal{S}_0$ be the shape model for an object ensemble. The underlying geometric entity, considered as a random

quantity, is denoted by S. In the examples in this paper, we assume it can be represented by a an m-rep model consisting of single-figure objects. Let A_i denote the i-th medial atom of S, assuming all the atoms are indexed by an index set $\mathcal{I}$. For any index set $\mathcal{J}$, $A_{\mathcal{J}} := \{A_i | i \in \mathcal{J}\}$. Therefore, the shape S is just $A_{\mathcal{I}}$. For simplicity, we identify A_i with $A_{\{i\}}$ when the index set has only one element, namely i.

Let $A^0_{\mathcal{I}}$ be the shape model. A random sample of the geometric entity $A_{\mathcal{I}}$ can be obtained by applying a random deformation F to $A^0_{\mathcal{I}}$. We describe this deformation by a series of, say K, random deformations, $F_1, F_2, \ldots, F_K$. The representation of $A_{\mathcal{I}}$ at the k-th step is

$$A^{\mathrm{k}}_{\mathcal{I}} = F_k \circ F_{k-1} \circ \cdots \circ F_1\big(A^0_{\mathcal{I}}\big) = F_k\big(A^{\mathrm{k}-1}_{\mathcal{I}}\big), \quad \text{for } k = 1, 2, \ldots, K.$$

The sequence $A^0_{\mathcal{I}}, A^1_{\mathcal{I}}, \ldots, A^{\mathrm{K}}_{\mathcal{I}}$ are representations of $A_{\mathcal{I}}$ at different scale levels and provide a series of approximations to $A_{\mathcal{I}}$. Each representation $A^{\mathrm{k}}_{\mathcal{I}}$ corresponds to a node of the tree in Fig. 2. If step k is an atom deformation step, F_k actually describes the net effect of a collection of individual atom deformations.

Our goal is to describe the shape by defining a probability measure on $A_{\mathcal{I}}$, which is usually considered as the shape prior. By virtue of our multi-step representation, we instead seek to define a joint probability distribution $\Pr(A^1_{\mathcal{I}}, A^2_{\mathcal{I}}, \ldots, A^{\mathrm{K}}_{\mathcal{I}})$. The marginal distribution of the final step, $\Pr(A^{\mathrm{K}}_{\mathcal{I}})$, is an approximation to the "true" distribution $\Pr(A_{\mathcal{I}})$ (Strictly speaking, these probability distributions are defined on appropriate spaces of deformations. The mapping $A^{\mathrm{k}-1}_{\mathcal{I}} \to A^{\mathrm{k}}_{\mathcal{I}}$ determines F_k.) Notice that we may equivalently define the joint distribution as follows:

$$\Pr\big(A^1_{\mathcal{I}}, A^2_{\mathcal{I}}, \ldots, A^{\mathrm{K}}_{\mathcal{I}}\big) = \Pr\big(A^1_{\mathcal{I}}\big) \prod_{k=2}^{K} \Pr\big(A^{\mathrm{k}}_{\mathcal{I}} \big| A^{\mathrm{k}-1}_{\mathcal{I}}, \ldots, A^1_{\mathcal{I}}\big). \quad (2)$$

At a first glance, this joint distribution seems to be more complicated than $\Pr(A_{\mathcal{I}})$ and thus harder to estimate. However, there are several reasons why this alternative may be a better approach.

Firstly, because geometric features have intrinsic scales associated with them, it is advantageous to describe a shape at multiple scale levels, so that features with various degrees of locality can be revealed. This also enables one to investigate and answer intuitive questions such as "which part of the object tends to vary more" or "how likely it is for one object to be bigger than a neighboring object". In many applications, especially in medical image analysis, these kinds of questions are very crucial for users. The series of probability distributions on the right hand side of Eq. (2) allows one to explicitly describe features at different scale levels.

Secondly, the multi-step description may be more effective in certain applications. One such example is the image segmentation of prostates. Because of the variability in prostate geometry and the lack of contrast in image intensity from 3D CT images, experience of ours and others suggests that the accuracy of a direct automatic segmentation has so far been inadequate. To achieve better results, we can exploit the geometric relationship between the prostate and the surrounding organs. For instance, we may build a statistical model for an object ensemble that contains two pubic bones and the prostate. The bones have better image contrast and are easier to segment. Once they are in place, the statistical inter-object relations between the prostate and the bones can be used to predict where the prostate should be, giving us a better chance of locating it. Probabilistically, this multi-step process is best characterized by a marginal probability measure of the bones and a conditional probability measure of the prostate given the bones (ref. Eq. (2)).

Thirdly, to estimate $\Pr(A_{\mathcal{I}})$ directly, we are likely to run into the HDLSS problem. On the other hand, we can estimate $\Pr(A^1_{\mathcal{I}}, \ldots, A^{\mathrm{K}}_{\mathcal{I}})$ by estimating the set of marginal and conditional distributions given by the right hand side of Eq. (2). Under certain Markov assumptions (explained in detail later), each conditional distribution may be specified by a much smaller number of parameters and thus may be estimated more accurately and efficiently.

In order to use Eq. (2), an explicit order in which different objects and scales are described has to be chosen. In other words, we need to decide how to index the objects in the ensemble and how different nodes of the tree in Fig. 2 are traversed. Obviously, our estimation of the probability distribution would depend on these decisions. In many situations, such as in the earlier prostate example, the geometric variability and/or image quality for some objects (e.g., the bones) are more stable than those of others (e.g., the prostate). As some previous work has suggested (Kapur et al., 1998; Davatzikos et al., 2003), since stability provides efficiency, we choose to describe objects in the order of decreasing stability, i.e. objects with less combined variation in geometry and image contrast should be described before those with more. In the prostate example, since we have more confidence in describing and identifying bones, we may choose to first describe them all the way through the atom level, then go to the object level for the prostate—which ends up in a depth-first order to traverse the tree in Fig. 2. It is certainly possible that there is more than one sensible order, and comparisons on the effects of choosing different orders need to be carried out. We will not discuss the details of these aspects further here. In the rest of this paper, we assume an appropriate order for the objects has been

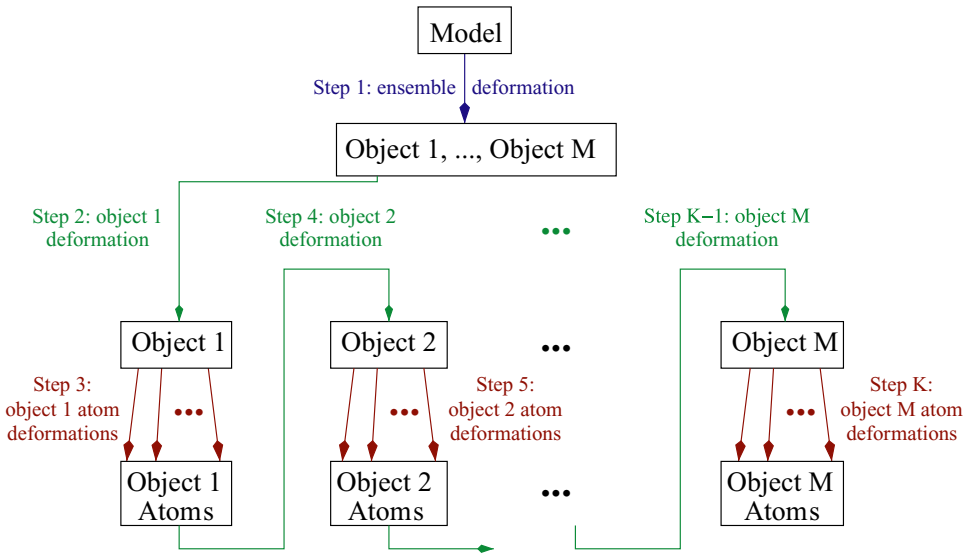


Figure 3. Depth-first representation of an object ensemble consisting of single-figure objects.

chosen, and we describe them in a depth-first manner, as shown in Fig. 3.

As an example, the two-bone shape model (see Fig. 1(b)) can be described in the following sequence of steps: the two-bone ensemble, the right bone (object 1) object level, the right bone atom level, the left bone (object 2) object level, and the left bone atom level. Therefore in this case, $M = 2$ and $K = 5$ (refer to Fig. 3.)

4. Markov Models for Object Ensembles

As we have seen, instances of a geometric entity $S = A_{\mathcal{I}}$ can be regarded as random samples drawn from an underlying population, according to certain probability distribution defined on it. This probability can be characterized by $\Pr(A^0_{\mathcal{I}}, \ldots, A^{\mathrm{K}}_{\mathcal{I}})$, where $A^0_{\mathcal{I}}, \ldots, A^{\mathrm{K}}_{\mathcal{I}}$ are descriptions of $A_{\mathcal{I}}$ at different levels. The shape model $A^0_{\mathcal{I}}$ is supposed to be fixed, although it needs to be optimized. Thus in essence we need to define $\Pr(A^1_{\mathcal{I}}, \ldots, A^{\mathrm{K}}_{\mathcal{I}} | A^0_{\mathcal{I}})$, which for simplicity we will denote by $\Pr(A^1_{\mathcal{I}}, \ldots, A^{\mathrm{K}}_{\mathcal{I}})$.

4.1. *Description of Inter-Object Relations*

The deformation of one object usually has an effect on other objects in the ensemble. Some of this effect is described by the ensemble deformation. We will assume that the residual deformations are more local, in the sense that the residual change of one object is only strongly correlated to changes of portions of some other neighboring objects. Inter-object relations are described via object-level residues from the ensemble level and neighbor residues. Below we explain how this is done using m-reps.

We represent inter-object relations by an augmentation-prediction mechanism that is described in more detail in Pizer et al. (2005). The basic idea is that for each object, say object k, we select from other objects a small subset of the atoms whose residual deformations are most correlated with those in object k. This group of selected atoms are called the *augmenting atoms* for object k. When describing the deformation on an object, we actually think of it as a deformation applied on both the object and its augmenting atoms. The rest of the atoms on other objects are deformed by a so-called *predicted deformation* which is intended to lower the overall shape change of these objects.

Suppose the shape template $A^0_{\mathcal{I}}$ is fixed. For $k \geq 1$, let $F_k : A^{\mathrm{k}-1}_{\mathcal{I}} \to A^{\mathrm{k}}_{\mathcal{I}}$ be the k-th step deformation. At the first step, a global deformation F_1 is applied to all the atoms in $A^0_{\mathcal{I}}$ to yield $A^1_{\mathcal{I}}$. At each of the following steps ($k \geq 2$), there is a deformation f_k that is either an object-level deformation or a collection of atom-level deformations on the atoms of a particular object, say object m_k (refer to Fig. 3), *plus* its augmenting atoms. Since we describe objects in sequential order, we can assume the atoms of object 1 through object $m_k - 1$ *do not change* at this step. Therefore, the augmenting atoms are in objects $m_k + 1$ through object M. F_k thus has three different types of effects on $A^{\mathrm{k}-1}_{\mathcal{I}}$:

- object 1 through object $m_k - 1$ do not deform;
- object m_k and its augmenting atoms are deformed by f_k;
- the rest of the atoms in object $m_k + 1$ through object M deform according to the predicted deformation Pred_k, as a result of f_k.

If we define the index sets for these three sets of atoms by

$$\begin{aligned}\mathcal{L}^k &:= \bigcup_{m=1}^{m_k-1} \mathcal{I}_m,\\ \mathcal{M}^k &:= \mathcal{I}_{m_k} \bigcup \{\text{indices of augmenting atoms}\},\\ \mathcal{N}^k &:= \mathcal{I} \setminus \big(\mathcal{L}^k \cup \mathcal{M}^k\big),\end{aligned}$$

where "\" denotes set difference, then in step k, $A_{\mathcal{L}^k}$ do not change, $A_{\mathcal{M}^k}$ deform according to f_k, and $A_{\mathcal{N}^k}$ deform according to Pred_k, which is based on (a) the previous-step relationship between $A^{k-1}_{\mathcal{M}^k}$ and $A^{k-1}_{\mathcal{N}^k}$ and (b) the current-step deformation on the augmenting atoms.

For a concrete example, consider the two-bone example. In this case F_1 is a deformation on atoms of both bones. When $k = 2$, $m_k = m_2 = 1$. This is the the right bone (object 1) object step. We choose a subset of atoms on the left bone whose residual deformations are most highly correlated with those of the right bone atoms. These are the *augmenting atoms* at step $k = 2$. By modeling the correlation between the right bone and the left bone augmented atoms, we can predict where the latter should be when the right bone is deformed. The remaining atoms of the left bone are then deformed by an appropriate transformation so that the overall shape change of the left bone is minimized. In other words, the prediction function (a) reflects the effect of right bone deformation through the deformations of the augmented atoms, and (b) propagates this effect to all other atoms of the left bone in a way such that the overall geometry (modulo rigid or similarity transforms) of the left bone has minimal change. The advantage of m-reps for these purposes is that it provides the inter-relation among atoms not only in terms of local translations, but also in terms of local differences in orientation.

In summary, the k-th step deformation $F_k : A^{k-1}_{\mathcal{I}} \to A^{k}_{\mathcal{I}}$ is characterized as follows:

$$\begin{aligned} F_1 :\quad & A^0_{\mathcal{I}} \mapsto A^1_{\mathcal{I}}, \qquad \text{for } k = 1,\\ F_k : \begin{cases} A^{k-1}_{\mathcal{L}^k} & \mapsto A^{k-1}_{\mathcal{L}^k},\\ A^{k-1}_{\mathcal{M}^k} & \mapsto f_k\big(A^{k-1}_{\mathcal{M}^k}\big), \qquad \text{for } k \geq 2.\\ A^{k-1}_{\mathcal{N}^k} & \mapsto \text{Pred}_k\big(A^{k-1}_{\mathcal{N}^k}, A^{k-1}_{\mathcal{M}^k}, f_k\big(A^{k-1}_{\mathcal{M}^k}\big)\big), \end{cases} \end{aligned} \tag{3}$$

Because of the way we define each F_k, the sequence $A^1_{\mathcal{I}}, A^2_{\mathcal{I}}, \ldots, A^{\text{K}}_{\mathcal{I}}$ has the following Markov property: given the geometry of an object ensemble at all previous steps, the geometric description at any particular step only depends on that of the immediate previous step. In other words,

$$\Pr\big(A^{\text{k}}_{\mathcal{I}} \big| A^{\text{k}-1}_{\mathcal{I}}, \ldots, A^1_{\mathcal{I}}\big) = \Pr\big(A^{\text{k}}_{\mathcal{I}} \big| A^{\text{k}-1}_{\mathcal{I}}\big), \qquad \text{for } k \geq 2,$$

and therefore

$$\begin{aligned}&\Pr\big(A^1_{\mathcal{I}}, A^2_{\mathcal{I}}, \ldots, A^{\text{K}}_{\mathcal{I}}\big)\\ &\quad = \Pr\big(A^1_{\mathcal{I}}\big) \prod_{k=2}^{K} \Pr\big(A^{\text{k}}_{\mathcal{I}} \big| A^{\text{k}-1}_{\mathcal{I}}, \ldots, A^1_{\mathcal{I}}\big)\\ &\quad = \Pr\big(A^1_{\mathcal{I}}\big) \prod_{k=2}^{K} \Pr\big(A^{\text{k}}_{\mathcal{I}} \big| A^{\text{k}-1}_{\mathcal{I}}\big).\end{aligned} \tag{4}$$

In light of Eq. (4), we will focus on defining $\Pr(A^1_{\mathcal{I}})$ and $\Pr(A^{\text{k}}_{\mathcal{I}} | A^{\text{k}-1}_{\mathcal{I}})$ for $k \geq 2$. This set of probabilities describe (a) the geometric variability of each object at different scale levels; and (b) the inter-object relations, using augmentation and prediction discussed earlier.

4.2. *Residues and Residue Deformations*

Our approach is to define $\Pr(A^{\text{k}}_{\mathcal{I}} | A^{\text{k}-1}_{\mathcal{I}})$ in terms of the difference between $A^{\text{k}}_{\mathcal{I}}$ and $A^{\text{k}-1}_{\mathcal{I}}$. Recall that a medial atom is an element of the space $G = \mathbb{R}^3 \times \mathbb{R}^+ \times S^2 \times S^2$. In fact, the space G is a *symmetric space*. A more thorough discussion about symmetric spaces and their properties can be found in Fletcher and Joshi (2004) and the references therein. Given any two medial atoms $A_1, A_2 \in G$, their difference can be described by using the following operator:

$$\begin{aligned}&\ominus : G \times G \to G,\\ &(A_1, A_2) \mapsto \Big(x_1 - x_2,\ \frac{r_1}{r_2},\ R_{u_2}(u_1),\ R_{v_2}(v_1)\Big),\end{aligned} \tag{5}$$

where $A_i = (x_i, r_i, u_i, v_i), i = 1, 2$, and for any $w = (w_1, w_2, w_3) \in S^2$, $R_w \in SO(3)$ denotes the rotation around the axis that passes through $(0, 0, 0)$ and $(w_2, -w_1, 0)$, with the rotation angle θ being the spherical distance between $p = (0, 0, 1)$ and w, i.e., $\theta = \arccos(w_3)$. Also, Let R_w^{-1} be the inverse of R_w. It is easy to see that $R_w(w) = p$ and that R_w maps a neighborhood of w to a neighborhood of p. R_w^{-1} does just the opposite.

Let $\Delta A = A_1 \ominus A_2$. It is also an element of G and is called the *residue* of A_1 relative to A_2. It is a measure of the difference between A_1 and A_2, *relative to* A_2*'s coordinates*. When $A_1 = A_2$, we get $\Delta A = (0, 1, p, p)$, which is the identity of G (if we choose p to be the identity of S^2). On the other hand, given an atom $A = (x, r, u, v)$ and a residue $\Delta A = (\Delta x, \Delta r, \Delta u, \Delta v)$, we can obtain the atom that has residue ΔA relative to A using the operator $\oplus$ defined as

$$\begin{aligned}&\oplus : G \times G \to G,\\ &(A, \Delta A) \mapsto \big(x + \Delta x,\ r \cdot \Delta r,\ R_u^{-1}(\Delta u),\ R_v^{-1}(\Delta v)\big).\end{aligned} \tag{6}$$

Table 1. Multi-step representation of the two-bone shape.

Step	Representation	Residue	Residue deformation
Shape template	$A_{\mathcal{I}}^0$	$R_{\mathcal{I}}^0$	—
Two-bone ensemble	$A_{\mathcal{I}}^1$	$R_{\mathcal{I}}^1$	$\Delta A_{\mathcal{I}}^1$: global deformation of both bones
Right bone object	$A_{\mathcal{I}}^2$	$R_{\mathcal{I}}^2$	$\Delta A_{\mathcal{I}}^2$: right bone object level deformation, plus prediction of left bone
Right bone atom	$A_{\mathcal{I}}^3$	$R_{\mathcal{I}}^3$	$\Delta A_{\mathcal{I}}^3$: right bone atom level deformation, plus prediction of left bone
Left bone object	$A_{\mathcal{I}}^4$	$R_{\mathcal{I}}^4$	$\Delta A_{\mathcal{I}}^4$: left bone object level deformation
Left bone atom	$A_{\mathcal{I}}^5$	$R_{\mathcal{I}}^5$	$\Delta A_{\mathcal{I}}^5$: left bone atom level deformation

Notice that this operator is neither commutable nor associative. The second operand is always understood as a residue measured relative to the coordinates of the first operand.

For two sets of corresponding atoms with a common index set $\mathcal{J}$, say $A_{\mathcal{J}}^1$ and $A_{\mathcal{J}}^2$, the difference between them is defined to be

$$A_{\mathcal{J}}^1 \ominus A_{\mathcal{J}}^2 := \left\{A_j^1 \ominus A_j^2 : j \in \mathcal{J}\right\}.$$

In our multi-scale representation, the k-th step representation $A_{\mathcal{I}}^{\mathrm{k}}$ is an approximation to the "true" representation $A_{\mathcal{I}}$. The *residue* at the k-th step is defined to be

$$R_{\mathcal{I}}^k := A_{\mathcal{I}} \ominus A_{\mathcal{I}}^{\mathrm{k}}, \quad k \geq 1. \tag{7}$$

It represents the residual geometric information that needs to be described at later steps. On the other hand, the effect of the k-th step deformation F_k is given by the difference

$$\Delta A_{\mathcal{I}}^{\mathrm{k}} := A_{\mathcal{I}}^{\mathrm{k}} \ominus A_{\mathcal{I}}^{\mathrm{k}-1}, \quad k \geq 1. \tag{8}$$

Therefore, $\Delta A_{\mathcal{I}}^{\mathrm{k}}$ can be regarded as an approximation to $R_{\mathcal{I}}^{k-1}$. We will call F_k (or equivalently, $\Delta A_{\mathcal{I}}^{\mathrm{k}}$) the *residue deformation* at step k.

In the two-bone shape example, let $A_{\mathcal{I}}$ be a random two-bone shape, and $A_{\mathcal{I}}^0$ be the shape template. At each step $k(1 \leq k \leq 5)$, the residue is $R_{\mathcal{I}}^k = A_{\mathcal{I}} \ominus A_{\mathcal{I}}^{\mathrm{k}}$, and the residue deformation is given by $\Delta A_{\mathcal{I}}^{\mathrm{k}} = A_{\mathcal{I}}^{\mathrm{k}} \ominus A_{\mathcal{I}}^{\mathrm{k}-1}$. We emphasize again that each residue deformation is described relative to the geometry of the previous step, not relative to the shape template. For instance, at step 4, the residue deformation is a further object level deformation of the left bone after step 3.

In general, the conditional distribution $\Pr(A_{\mathcal{I}}^{\mathrm{k}} | A_{\mathcal{I}}^{\mathrm{k}-1})$ describes the residual transformation $\Delta A_{\mathcal{I}}^{\mathrm{k}}$, which is an approximation to the residue $R_{\mathcal{I}}^k$. We can construct the sequence $A_{\mathcal{I}}^1, A_{\mathcal{I}}^2, \ldots, A_{\mathcal{I}}^{\mathrm{K}}$ in such a way that $\Pr(A_{\mathcal{I}}^{\mathrm{k}} | A_{\mathcal{I}}^{\mathrm{k}-1})$ only depends on $\Delta A_{\mathcal{I}}^{\mathrm{k}}$ and not explicitly on $A_{\mathcal{I}}^{\mathrm{k}-1}$, so long as we can capture most of the variation among the population. As a result, the original joint distribution $\Pr(A_{\mathcal{I}}^1, A_{\mathcal{I}}^2, \ldots, A_{\mathcal{I}}^{\mathrm{K}})$ can be induced by $A_{\mathcal{I}}^0$ and the product

$$\prod_{k=1}^{K} \Pr\left(\Delta A_{\mathcal{I}}^{\mathrm{k}}\right). \tag{9}$$

In stochastic process terminology, we have designed our approach so that the sequence of representations $A_{\mathcal{I}}^0, A_{\mathcal{I}}^1, A_{\mathcal{I}}^2, \ldots, A_{\mathcal{I}}^{\mathrm{K}}$ has "independent increments". The multi-step representation of the two-bone shape $A_{\mathcal{I}}$ is summarized in Table 1. In this case, we need to estimate $A_{\mathcal{I}}^0$ as well as the (independent) probabilities $\Pr(\Delta A_{\mathcal{I}}^1), \ldots, \Pr(\Delta A_{\mathcal{I}}^5)$.

4.3. Atom Level Residue Deformations

As mentioned earlier, if $\Delta A_{\mathcal{I}}^{\mathrm{k}}$ corresponds to an atom residue transformation step (e.g., if $k = 3$ or $k = 5$ in the previous example), it actually represents a collection of individual atom residue deformations, each of which characterizes the deformation of an atom given those of all other atoms and is described by the probability $\Pr(\Delta A_i^{\mathrm{k}} | \Delta A_{\mathcal{M}^k \setminus \{i\}}^{\mathrm{k}})$, where $A_{\mathcal{M}^k}$ contains (a) the atoms that are subject to atom residue deformations at step k, and (b) the augmenting atoms (on neighboring objects). At this finer scale level, we assume that the direct long range dependency among $A_{\mathcal{M}^k}$ have already been described at the (larger) object level, so that the atom residue deformations can be described by the following Markov random field (MRF) model:

$$\Pr\left(\Delta A_i^{\mathrm{k}} | \Delta A_{\mathcal{M}^k \setminus \{i\}}^{\mathrm{k}}\right) = \Pr\left(\Delta A_i^{\mathrm{k}} | \Delta A_{N(i)}^{\mathrm{k}}\right), \quad i \in \mathcal{M}^k, \tag{10}$$

where $N(i)$ denotes the index set of the atoms that are neighbors of atom i and is relatively small. Because the medial sheet is sampled by a quadrilateral array of atoms, we use the neighborhood structure induced by the 4-adjacency graph of the quad-mesh (see Fig. 5.)

We can define these MRF models by specifying potentials on the cliques in the 4-adjacency graph and estimating parameters using standard MRF techniques such as

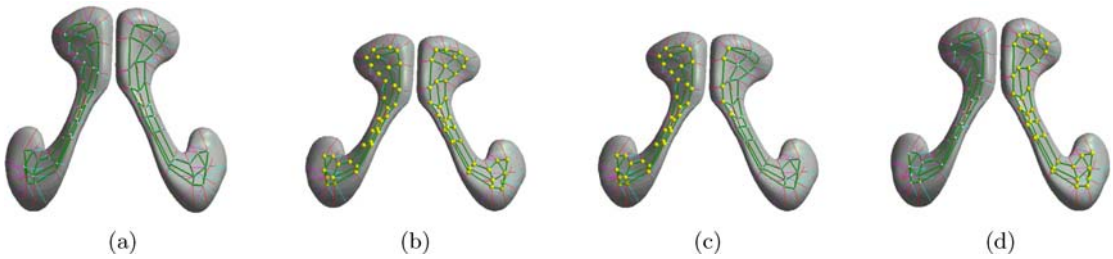


Figure 4. Multi-step representation of the two-bone shape. The deformation at each stage is applied to highlighted atoms. (a) The template. (b) The ensemble stage deformation is applied to all atoms. (c) The right bone deformation is applied to atoms on the right bone and the augmenting atoms on the left bone (highlighted). The rest of the atoms on the left bone are deformed by the prediction function. (d) The left bone deformation is applied to all atoms on the left bone. The right bone does not change.

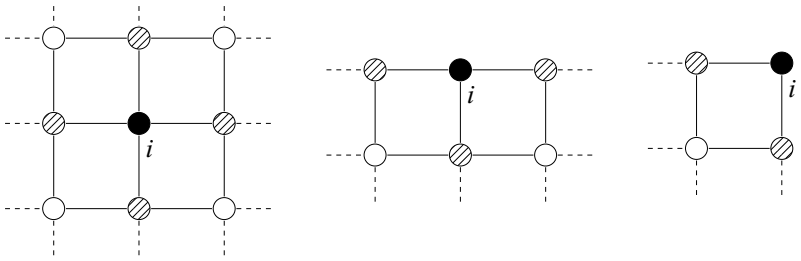


Figure 5. The 4-neighbor structure for quad-mesh. A typical atom i has 4 neighbors (shaded), whereas an atom on an edge or at a corner has fewer neighbors.

Monte Carlo Markov Chain (MCMC) methods. An example of this approach can be found in Lu et al. (2003). In many situations, this type of models still yield a large number of parameters that need to be estimated simultaneously. An alternative approach is to select an order in which we go through the atoms one by one and decompose the joint distribution $\Pr(\Delta A_{\mathcal{M}^k})$ into a product of marginal and conditional distributions, similar to Eq. (2). However, the quadrilateral structure of the atom mesh does not provide a natural traversing order, making this method somewhat ad hoc and hard to analyze. A third method, which we adopt here, is to estimate the conditional distributions $\Pr(\Delta A_i^{\mathrm{k}} | \Delta A_{N(i)}^{\mathrm{k}})$ directly. To deal with the low sample size problem, we assume that the conditional probability distribution $\Pr(\Delta A_i^{\mathrm{k}} | \Delta A_{N(i)}^{\mathrm{k}})$ does not explicitly depend on $\Delta A_{N(i)}^{\mathrm{k}}$, but rather is induced by a probability distribution on the difference $\Delta A_i^{\mathrm{k}} \ominus \Delta A_{i,p}^{\mathrm{k}}$, where for each i, the term $\Delta A_{i,p}^{\mathrm{k}}$ represents a residue deformations of atom i that is *predicted* by the deformations of its neighbors and is calculated as a weighted average of $\Delta A_{N(i)}^{\mathrm{k}}$. The problem then is to estimate the probabilities

$$\Pr\left(\Delta A_i^{\mathrm{k}} \ominus \Delta A_{i,p}^{\mathrm{k}}\right), \quad i \in \mathcal{M}^k. \tag{11}$$

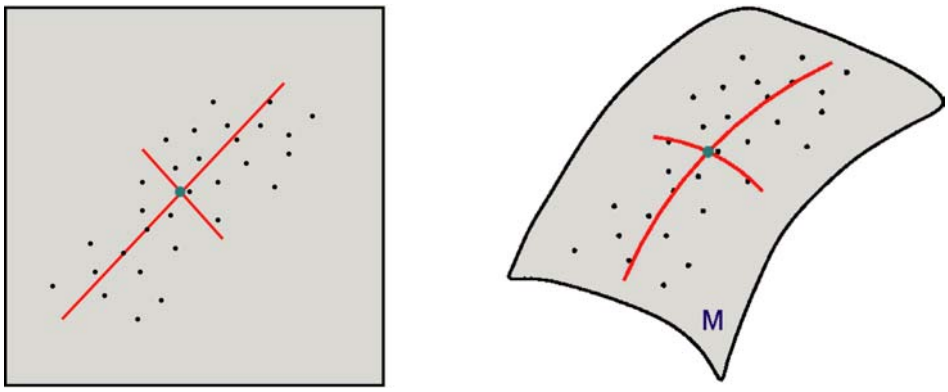
From (9) and (11), we see that the probability distribution of the object ensemble is induced from a product of inter-scale and intra-scale residue distributions, each of which can be estimated separately. In the next section we discuss how to learn these distributions from a training set.

5. Estimation of the Probabilistic Model

Suppose the training set $\{A_{1,\mathcal{I}}, A_{2,\mathcal{I}}, \ldots, A_{N,\mathcal{I}}\}$ contains N instances of the random object ensemble $A_{\mathcal{I}}$. Let $\mathcal{T} = \{1, 2, \ldots, N\}$. We use $A_{n,\mathcal{I}}$ to denote an instance. The first subscript n denotes the index in the training set, i.e., $n \in \mathcal{T}$, and the second one denotes the index for medial atoms as before. Therefore, the whole training set can be denoted by $A_{\mathcal{T},\mathcal{I}}$. We assume that the m-rep representation provides correspondence atom-to-atom among the training set. In particular, the corresponding atoms across the population have the same index in $\mathcal{I}$. As before, we assume that the object ensemble has K single-figure objects, and that we use the depth-first order to traverse the objects as shown in Fig. 3.

5.1. *M-rep Alignment and Principal Geodesic Analysis (PGA)*

The first step in training is to align the shape instances. We do this by an m-rep alignment procedure that is similar to Procrustes analysis (Goodall, 1991). In the standard Procrustes analysis, a rigid or similarity transformation is applied to each instance so that the total sum-of-squared Euclidean distances between corresponding



(a) PCA (b) PGA

Figure 6. Comparison of PCA in linear spaces and PGA in curved spaces. Each small dot represents a sample point. The big dot indicates the Fréchet mean. The thick lines/curves represent principal components/geodesics. Both surfaces represent n-dimensional spaces.

points is minimized. In contrast, in the m-rep alignment, described in Fletcher et al. (2004), the sum-of-squared geodesic distances between corresponding medial atoms is minimized. In what follows we assume the shape instances are aligned.

The probabilities $\Pr(\Delta A_{\mathcal{I}}^1), \ldots, \Pr(\Delta A_{\mathcal{I}}^{\mathrm{K}})$ are defined on m-rep parameter spaces, which are nonlinear Riemannian symmetric spaces. In these spaces, one can define the so-called exponential map, Exp_x, and log map, Log_x, which establish correspondence between tangent vectors and geodesics at any given point x (Fletcher et al., 2003). Based on the properties of these maps, Fletcher et al. have developed a method called principal geodesic analysis (PGA), which is a generalization of principal component analysis to nonlinear Lie groups and symmetric spaces. The details are laid out in Fletcher et al. (2004, 2003). Here we briefly summarize the basic ideas.

Suppose M is a symmetric space or Lie group. Given a set of samples $x_1, \ldots, x_N \in M$, we want to estimate the underlying probability distribution from which they are drawn. The PGA of $\{x_i\}_{i=1}^N$ essentially involves finding the mean μ of $\{x_i\}_{i=1}^N$, projecting $\{x_i\}_{i=1}^N$ to $T_\mu M$ (the tangent space at μ) by the log map, and doing a standard principal component analysis (PCA) in that tangent space, which yields a set of principal directions $\{v_k\}$ and corresponding eigenvalues $\{\lambda_k\}$. The "principal geodesics" are determined by applying the exponential map to the principal components obtained from PCA. Analogous to principal components, the principal geodesics describe the major modes of variation among $\{x_i\}_{i=1}^N$ in the curved space M (see Fig. 6).

The space M can be approximated by a submanifold $H \subset M$ that is generated by the first, say h, principal geodesics. Elements in H can be represented as

$$\mathrm{Exp}_\mu\left(\sum_{k=1}^{h} \alpha_k v_k\right), \tag{12}$$

where $\alpha_k \in \mathbb{R}$. $\forall x \in M$, the projection of x onto H is

$$\begin{aligned}\pi_H(x) &= \arg\min_{y \in H} \mathrm{dist}^2(x, y) \\ &\approx \mathrm{Exp}_\mu\left(\sum_{k=1}^{h} \langle \mathrm{Log}_\mu(x), v_k\rangle \cdot v_k\right),\end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes inner product in $T_\mu M$.

The density

$$\frac{1}{Z}\exp\left(-\frac{1}{2}\sum_{k=1}^{h}\frac{\langle v_k, y\rangle^2}{\lambda_k}\right), \quad y \in N(\mu) \subset T_\mu M, \tag{13}$$

where Z is a normalizing constant and $N(\mu)$ is a neighborhood of μ in the tangent plane, defines a probability distribution on $N(\mu)$. It induces a probability distribution on M by the exponential map Exp_μ. This induced distribution is an approximation to the probability distribution on M from which $\{x_i\}_{i=1}^N$ are drawn.

In what follows we describe the process for estimating $\Pr(\Delta A_{\mathcal{I}}^1), \ldots, \Pr(\Delta A_{\mathcal{I}}^{\mathrm{K}})$. The two-bone shape will be used as an example. The ensemble has a total of 72 medial atoms, with 36 atoms on each bone.

5.2. *Estimation of the Shape Model*

The shape model $A_{\mathcal{I}}^0$ is set to be the centroid of the training population $A_{\mathcal{T},\mathcal{I}}$. This is obtained by computing the Fréchet mean of each atom. In other words,

$$A_{\mathcal{I}}^0 = \underset{\mathcal{T}}{\mathrm{mean}}(A_{\mathcal{T},\mathcal{I}}) = \left\{\underset{\mathcal{T}}{\mathrm{mean}}(A_{\mathcal{T},i}) : i \in \mathcal{I}\right\}. \tag{14}$$

The Fréchet mean of the set of atoms $A_{\mathcal{T},i}$ is defined to be an atom A_i^0 that satisfies

$$A_i^0 = \arg\min_{A \in G} \sum_{n \in \mathcal{T}} \text{dist}^2(A, A_{n,i}),$$

where $\text{dist}(\cdot,\cdot)$ is the geodesic distance on G, the atom parameter space. The Fréchet mean can be computed using a gradient descent method. For details refer to Fletcher et al. (2003, 2004).

The residues at this step are the differences between the instances and the model, i.e.,

$$R^0_{n,\mathcal{I}} := A_{n,\mathcal{I}} \ominus A^0_{\mathcal{I}}, \quad n \in \mathcal{T}.$$

5.3. Object Ensemble Residue Statistics

The ensemble deformation $\Delta A^1_{\mathcal{I}}$ is an approximation to $R^0_{n,\mathcal{I}}$. We estimate $\Pr(\Delta A^1_{\mathcal{I}})$ by doing a principal geodesic analysis (PGA) on $R^0_{\mathcal{T},\mathcal{I}}$. In the two-bone example, this amounts to a PGA of a population of tuples consisting of 72 atoms, or equivalently, a 576-dimensional PGA since each atom has 8 parameters.

Let H be the m-rep parameter space (in fact, if there are a total of d medial atoms in the object ensemble, then $H = G^d$. Recall that $G = \mathbb{R}^3 \times \mathbb{R}^+ \times S^2 \times S^2$.) $\Pr(\Delta A^1_{\mathcal{I}})$ is approximated by the first h_1 principal geodesics, which generate a sub-manifold $H^1 \subset H$. The ensemble deformation for each instance is described by the projection of $R^0_{\mathcal{T},\mathcal{I}}$ onto H^1, i.e.,

$$\Delta A^1_{n,\mathcal{I}} := \text{Proj}_{H^1}\big(R^0_{n,I}\big), \quad n \in \mathcal{T}.$$

The representations and residues at this step are

$$\begin{aligned} A^1_{n,\mathcal{I}} &:= A^0_{\mathcal{I}} \oplus \Delta A^1_{n,\mathcal{I}}, \\ R^1_{n,\mathcal{I}} &:= A_{n,\mathcal{I}} \ominus A^1_{n,\mathcal{I}}, \end{aligned} \qquad n \in \mathcal{T},$$

where $\oplus$ is defined by Eq. (6). The above procedure is described by Algorithm 1.

As with PCA, we need to decide the value for h_1, which is the number of principal geodesics to use for approximation. While too few geodesics might not be adequate, too many may be unstable due to the small sample size. In principle, we want to choose the smallest value such that the residues possess certain Markov properties at the (next) object level. In practice, this is not always easy to verify, and we use some heuristics to determine h_1. For example, we can choose h_1 to be the smallest number such that (a) the percentage of variance explained by the first h_1 principal geodesics is above a certain threshold, and/or (b) the inter-object correlations among residues are below a certain threshold.

5.4. Object Residue Statistics for Object 1

First, suppose object 1's residue deformation does not have an effect on neighboring objects. Let the index set for the atoms in object 1 be $\mathcal{I}_1$. Then at this step, the probability $\Pr(\Delta A^2_{\mathcal{I}})$ actually describes the object residue deformation $\Delta A^2_{\mathcal{I}_1}$. We estimate this probability by performing a PGA on $R^1_{\mathcal{T},\mathcal{I}_1}$ and using the first h_2 principal geodesics to approximate it. The procedure is similar to Algorithm 1 and is summarized in Algorithm 2. For the two-bone case, this corresponds to a $36 * 8 = 288$ dimensional PGA on the right bone.

To capture the effect of object 1's deformation on other objects, we need to update all other objects so that the inter-object relationship is reflected. This is achieved with an augmentation-prediction mechanism. We use the two-bone shape to illustrate how it works. In this case object 1 is the right bone, and object 2 is the left bone. The index set for the atoms of these objects are $\mathcal{I}_1, \mathcal{I}_2$, respectively. We pick a subset of object 2 (left bone)'s atoms that are most highly correlated with the atoms in object 1 (right bone); these typically correspond to the section of object 2 that is closest to object 1 (see Fig. 4). Let $\mathcal{U}^2$ be the index set for the augmenting atoms and $\mathcal{M}^2$ denote the union of $\mathcal{I}_1$ and $\mathcal{U}^2$. We can then perform an *augmented* PGA analysis on $R^1_{\mathcal{T},\mathcal{M}^2}$ instead of $R^1_{\mathcal{T},\mathcal{I}_1}$, using Algorithm 2 with $\mathcal{J} = \mathcal{M}^2$. In the two-bone example, if there are 5 augmented atoms on the left bone, then this leads to a $(36 + 5) \cdot 8 = 328$ dimensional PGA.

Input: $A_{\mathcal{T},\mathcal{I}}, \mathbf{A}^0_{\mathcal{I}}$
Output: P^1 =estimates of $\Pr(\Delta \mathbf{A}^1_{\mathcal{J}})$, as well as $A^1_{\mathcal{T},\mathcal{J}}$
$R^0_{\mathcal{T},\mathcal{J}} \leftarrow A_{\mathcal{T},\mathcal{J}} \ominus \mathbf{A}^0_{\mathcal{I}}$
$P^1 \leftarrow$ PGA results on $R^0_{\mathcal{T},\mathcal{I}}$ (principal geodesics and eigenvalues)
$H^1 \leftarrow$ sub-manifold generated by a chosen number of principal geodesics of $R^0_{\mathcal{T},\mathcal{I}}$
$\Delta A^1_{\mathcal{T},\mathcal{I}} \leftarrow \text{Proj}_{H^1}(R^0_{\mathcal{T},\mathcal{I}}), A^1_{\mathcal{T},\mathcal{I}} \leftarrow \mathbf{A}^0_{\mathcal{I}} \oplus \Delta A^1_{\mathcal{T},\mathcal{I}}$
Output $P^1, A^1_{\mathcal{T},\mathcal{I}}$

Algorithm 1. Object ensemble statistics.

Input: $A_{\mathcal{T},\mathcal{J}}, A^{k-1}_{\mathcal{T},\mathcal{J}}$, where $\mathcal{J}$ is the index set for the atoms in the object
Output: P^k = estimates of $\Pr(\Delta \mathbf{A}^{\mathrm{k}}_{\mathcal{J}})$, as well as $A^k_{\mathcal{T},\mathcal{J}}$
$R^{k-1}_{\mathcal{T},\mathcal{J}} \leftarrow A_{\mathcal{T},\mathcal{J}} \ominus A^{k-1}_{\mathcal{T},\mathcal{J}}$
$P^k \leftarrow$ PGA results on $R^{k-1}_{\mathcal{T},\mathcal{J}}$ (Fréchet mean $\mu^k_{\mathcal{J}}$, principal geodesics and eigenvalues)
$H^k \leftarrow$ sub-manifold generated by the principal geodesics of $R^{k-1}_{\mathcal{T},\mathcal{J}} \ominus \mu^k_{\mathcal{J}}$
$\Delta A^k_{\mathcal{T},\mathcal{J}} \leftarrow \mathrm{Proj}_{H^k}(R^{k-1}_{\mathcal{T},\mathcal{J}}), A^k_{\mathcal{T},\mathcal{J}} \leftarrow A^{k-1}_{\mathcal{T},\mathcal{J}} \oplus \mu^k_{\mathcal{J}} \oplus \Delta A^k_{\mathcal{T},\mathcal{J}}$
Output $P^k, A^k_{\mathcal{T},\mathcal{J}}$

Algorithm 2. Object residue statistics for a single-figure object at step k.

The deformation of the augmenting atoms $\Delta A^{2'}_{\mathcal{U}^2} : A^1_{\mathcal{U}^2} \to A^{2'}_{\mathcal{U}^2}$ also predict how the rest of the atoms on object 2 will deform. Our approach is as follows. We find that member of the principal geodesic shape space of the remaining objects that agrees best on the augmenting atoms, and we remove that member from the remaining objects' atoms. In detail, first, we perform a PGA on $R^1_{\mathcal{T},\mathcal{I}_2}$, the residue of object 2, and denote the mean and PGA submanifold by ρ^2 and $\mathcal{H}^2$, respectively. The *predicted* residual deformation of object 2 is then defined to be

$$\begin{aligned}&\mathrm{Pred}_2\big(\Delta A^2_{\mathcal{I}_2}; \Delta A^{2'}_{\mathcal{U}^2}\big)\\ &= \mathrm{Proj}_{\mathcal{H}^2}\big(\Delta A^{2'}_{\mathcal{U}^2}\big) = \exp_{\rho^2}\left(\sum_{j=1}^{h} \big\langle \mathrm{Log}_{\rho^2}\big(\Delta A^{2'}_{\mathcal{U}^2}\big), v_j\big\rangle \cdot v_j\right),\end{aligned} \tag{15}$$

where $\{v_j\}_{j=1}^h$ are principal directions in the tangent space of ρ^2 corresponding to the principal geodesics in $\mathcal{H}^2$ and the dimension of $\mathrm{Log}_{\rho^2}(\Delta A^{2'}_{\mathcal{U}^2})$ is adjusted to match with that of v_j by adding zeros to $\mathrm{Log}_{\rho^2}(\Delta A^{2'}_{\mathcal{U}^2})$ for parameters corresponding to $A_{\mathcal{I}_2 \setminus \mathcal{U}^2}$.

It is straightforward to adapt these procedures to situations where the object ensemble has more than two objects. The net effect of these predictions is summarized by a prediction function $\mathrm{Pred}_2()$[1], which is determined using statistics from a training set. The general algorithm is described in Alorithm 3.

In the two-bone example, after the object 1 (right bone) deformation step the we have

$$\begin{aligned}A^2_{\mathcal{T},\mathcal{I}_1} &= A^1_{\mathcal{T},\mathcal{I}_1} \oplus \mu^2_{\mathcal{I}_1} \oplus \mathrm{Proj}_{H^2}\big(R^1_{\mathcal{T},\mathcal{I}_1}\big),\\ A^2_{\mathcal{T},\mathcal{I}_2} &= A^1_{\mathcal{T},\mathcal{I}_2} \oplus \rho^2_{\mathcal{I}_2} \oplus \mathrm{Proj}_{\mathcal{H}^2}\big(\Delta A^{2'}_{\mathcal{T},\mathcal{U}^2}\big),\\ R^2_{\mathcal{T},\mathcal{I}} &= A_{\mathcal{T},\mathcal{I}} \ominus A^2_{\mathcal{T},\mathcal{I}}.\end{aligned} \tag{16}$$

5.5. *Atom Residue Statistics for Object 1*

The probabilities to be estimated at this step are $\Pr(\Delta A^3_i \ominus \Delta A^3_{i,p})$, $i \in \mathcal{J}^{3,1}$. Recall that $\Delta A^3_{i,p}$ is the residue deformation of atom i that is predicted by those of its neighbors. For each i, we estimate $\Pr(\Delta A^3_i \ominus \Delta A^3_{i,p})$ by doing a PGA on $R^2_{\mathcal{T},i} \ominus R^2_{\mathcal{T},\,i,p}$, where $R^2_{\mathcal{T},\,i,p}$ denotes the residue of atom i that is predicted by those of its neighbors. Let H^3_i be the submanifold generated by a chosen number of stable principal geodesics. Then

$$A^3_{\mathcal{T},i} = A^2_{\mathcal{T},i} \oplus R^2_{\mathcal{T},\,i,p} \oplus \mu^3_i \oplus \mathrm{Proj}_{H^3_i}\big(R^2_{\mathcal{T},i} \ominus R^2_{\mathcal{T},\,i,p}\big). \tag{17}$$

This process is summarized in Algorithm 4 (with $\mathcal{J}$ being the augmented set). For the two-bone model, this step corresponds to estimating atom residue deformations for the 36 atoms on the right bone and 5 augmenting atoms on the left bone. Each estimation involves an 8-dimensional PGA.

Input: $A_{\mathcal{T},\mathcal{J}}, A^{k-1}_{\mathcal{T},\mathcal{J}}, A^k_{\mathcal{T}\mathcal{U}'}$, where $\mathcal{J}$ is the index set for atoms in objects that need to be predicted and $\mathcal{U} \subset \mathcal{J}$ is the index set for augmenting atoms at step k
Output: $A^k_{\mathcal{T},\mathcal{J}}$ = predicted objects
$R^{k-1}_{\mathcal{T},\mathcal{J}} \leftarrow A_{\mathcal{T},\mathcal{J}} \ominus A^{k-1}_{\mathcal{T},\mathcal{J}}$
$\mathcal{P}^k \leftarrow$ PGA results on $R^{k-1}_{\mathcal{T},\mathcal{J}}$ (Fréchet mean $\rho^k_{\mathcal{J}}$, principal geodesics and eigenvalues)
$\mathcal{H}^k \leftarrow$ sub-manifold generated by the principal geodesics of $R^{k-1}_{\mathcal{T},\mathcal{J}} \ominus \rho^k_{\mathcal{J}}$
$A^k_{\mathcal{T},\mathcal{J}} \leftarrow A^{k-1}_{\mathcal{T},\mathcal{J}} \oplus \rho^k_{\mathcal{J}} \oplus \mathrm{Proj}_{\mathcal{H}^k}(\Delta A^k_{\mathcal{T}\mathcal{U}'})$
Output $\mathcal{P}^k, A^k_{\mathcal{T},\mathcal{J}}$

Algorithm 3. Statistical prediction at step k.

Input: $A_{\mathcal{T},\mathcal{J}}, A^{k-1}_{\mathcal{T},\mathcal{J}}$, where $\mathcal{J}$ is the index set for the atoms in the current object
Output: P^k = estimates of $\left\{\Pr\left(\Delta\mathbf{A}^{\mathrm{k}}_i \ominus \Delta\mathbf{A}^{\mathrm{k}}_{i,p}\right)\right\}_{i\in\mathcal{I}}$, as well as $A^k_{\mathcal{T},\mathcal{J}}$
$R^{k-1}_{\mathcal{T},\mathcal{J}} \leftarrow A_{\mathcal{T},\mathcal{J}} \ominus A^{k-1}_{\mathcal{T},\mathcal{J}}$, $R^{k-1}_{\mathcal{T},\mathcal{J},p} \leftarrow A_{\mathcal{T},\mathcal{J},p} \ominus A^{k-1}_{\mathcal{T},\mathcal{J},p}$
for all $i \in \mathcal{J}$ **do**
$P^k_i \leftarrow$ PGA results on $R^{k-1}_{\mathcal{T},i} \ominus R^{k-1}_{\mathcal{T},i,p}$ (mean μ^k_i, principal geodesics and eigenvalues)
$H^k_i \leftarrow$ sub-manifold generated by the first few principal geodesics
$A^k_{\mathcal{T},i} \leftarrow A^{k-1}_{\mathcal{T},i} \oplus R^{k-1}_{\mathcal{T},i,p} \oplus \mu^k_i \oplus \mathrm{Proj}_{H^k_i}\left(R^{k-1}_{\mathcal{T}\cdot i} \ominus R^{k-1}_{\mathcal{T},i,p}\right)$
end for
Output $P^k := \{P^k_i\}_{i\in\mathcal{J}}, A^k_{\mathcal{T},\mathcal{J}}$

Algorithm 4. Atom residue statistics for a single-figure object at step k.

The atom residue deformations for object 1 will have an effect on other objects. This can be described by the augmentation-prediction scheme similar to the one described in the previous step. For the two-bone example we have

$$\begin{aligned} A^3_{\mathcal{T},i} &= A^2_{\mathcal{T},i} \oplus R^2_{\mathcal{T},i,p} \oplus \mu^3_i \oplus \mathrm{Proj}_{H^3_i}\left(R^2_{\mathcal{T},i} \ominus R^2_{\mathcal{T},i,p}\right), \\ &\qquad i \in \mathcal{I}_1 \\ A^3_{\mathcal{T},\mathcal{I}_2} &= A^2_{\mathcal{T},\mathcal{I}_2} \oplus \rho^3_{\mathcal{I}_2} \oplus \mathrm{Proj}_{\mathcal{H}^3}\left(\Delta A^{2'}_{\mathcal{T},\mathcal{U}^3}\right) \\ &= A^2_{\mathcal{T},\mathcal{I}_2} \oplus \rho^3_{\mathcal{I}_2} \oplus \mathrm{Pred}_3\left(\Delta A^3_{\mathcal{I}_2}; \Delta A^{3'}_{\mathcal{U}^3}\right) \qquad (18) \\ R^3_{\mathcal{T},\mathcal{I}} &= A_{\mathcal{T},\mathcal{I}} \ominus A^3_{\mathcal{T},\mathcal{I}} \end{aligned}$$

5.6. *Object and Atom Residue Statistics for Other Objects*

Suppose we describe object 2 next. Let $\mathcal{I}_2$ be the index set for its atoms. The object residue statistics can be obtained using Algorithm 2, with $k = 4$ and $\mathcal{J} = \mathcal{I}_2 \cup \mathcal{U}^4$, where $\mathcal{U}^4$ is the index set for augmenting atoms. The difference is that in this step, there is no deformation on object 1. Therefore

$$\begin{aligned} A^4_{\mathcal{T},\mathcal{I}_1} &= A^3_{\mathcal{T},\mathcal{I}_1} \\ A^4_{\mathcal{T},\mathcal{I}_2} &= \text{results from Algorithm 2} \\ &\qquad (\text{with } k = 4 \text{ and } \mathcal{J} = \mathcal{I}_2 \cup \mathcal{U}^4) \qquad (19) \\ A^4_{\mathcal{T},\mathcal{I}'} &= \text{results from Algorithm 3} \\ &\qquad (\text{with } k = 4 \text{ and } \mathcal{J} = \mathcal{I}' = \mathcal{I} \setminus (\mathcal{I}_1 \cup \mathcal{I}_2)) \\ R^4_{\mathcal{T},\mathcal{I}} &= A_{\mathcal{T},\mathcal{I}} \ominus A^4_{\mathcal{T},\mathcal{I}}. \end{aligned}$$

For the two-bone example, the left bone object residue deformation is estimated by a $36 * 8 = 288$ dimensional PGA. Since there are only two objects, there is no augmentation-prediction at this step ($\mathcal{U}^4$ and $\mathcal{I}'$ are empty).

The atom residue statistics for object 2 is computed by Algorithm 4, with $k = 5$ and $\mathcal{J} = \mathcal{I}_2 \cup \mathcal{U}^5$. The prediction works the same way as in the previous step. Therefore,

$$\begin{aligned} A^5_{\mathcal{T},\mathcal{I}_1} &= A^4_{\mathcal{T},\mathcal{I}_1} \\ A^5_{\mathcal{T},i} &= A^4_{\mathcal{T},i} \oplus R^4_{\mathcal{T},i,p} \oplus \mu^5_i \ominus \mathrm{Proj}_{H^5_i}\left(R^4_{\mathcal{T},i} \ominus R^4_{\mathcal{T},i,p}\right), \\ &\qquad i \in \mathcal{I}_2 \\ A^5_{\mathcal{T},\mathcal{I}'} &= A^4_{\mathcal{T},\mathcal{I}'} \oplus \rho^5_{\mathcal{I}'} \oplus \mathrm{Proj}_{\mathcal{H}^5}\left(\Delta A^{4'}_{\mathcal{T},\mathcal{U}^5}\right), \\ &\qquad \text{where } \mathcal{I}' = \mathcal{I} \setminus (\mathcal{I}_1 \cup \mathcal{I}_2) \\ R^5_{\mathcal{T},\mathcal{I}} &= A_{\mathcal{T},\mathcal{I}} \ominus A^5_{\mathcal{T},\mathcal{I}}. \qquad (20) \end{aligned}$$

Again, in the two-bone case, the left bone atom residue deformations are described by 36 individual 8-dimensional PGA's. There is no augmentation-prediction at this step ($\mathcal{U}^5$ and $\mathcal{I}'$ are empty).

If there are more than two objects in the object complex, we can perform the same procedures and go through each of them one by one to generate the statistics we need.

5.7. *Summary*

The entire estimation process is summarized in the Algorithm 5.

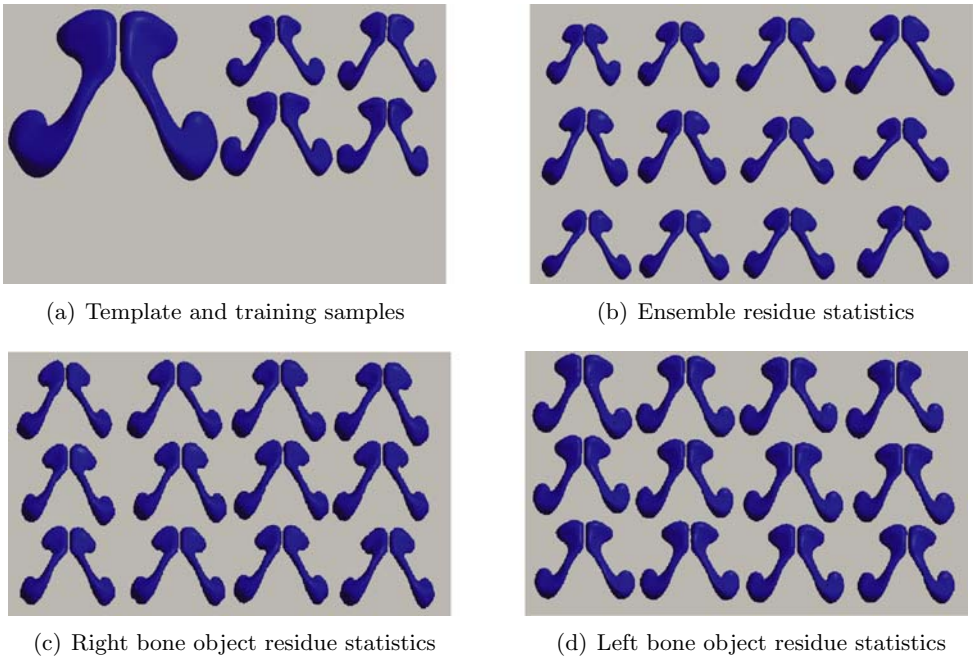
5.8. *Examples of Multi-Object Statistics*

We now present some statistics of a pair of pubic bones using the training procedure just described. The training population has 15 shape instances from different patients. For this illustration, we omit the atom residue stages and describe each instance at the ensemble, right bone object, and left bone object stages. Each bone is represented by 36 medial atoms on a 3×12 grid. The results are shown in Fig. 7.

At each stage, the base shape is the shape obtained by applying the mean residue at this stage to the shape template. In other words, the base shape at step k is $A^0_{\mathcal{I}} \oplus \mu^k$. The principal modes are shown relative to the corresponding base shape. These modes typically have intuitive interpretations. For instance, Fig. 7(b) shows the first three PGA modes for the ensemble residues. The first mode is mainly a size variation; the second mode reflects a

Input: A training set $A_{\mathcal{T},\mathcal{I}}$, where $\mathcal{I}$ is the index set for all atoms in the object ensemble, and $\mathcal{T}$ is the index set for the training samples
Output: estimates of mean $\mathbf{A}_{\mathcal{I}}^0$ and $\Pr(\Delta\mathbf{A}_{\mathcal{J}}^k), k = 1,\ldots,K$
Determine an order to traverse the objects (see Fig. 3). $K \leftarrow$ number of steps
$\mathbf{A}_{\mathcal{I}}^0 \leftarrow \text{mean}(A_{\mathcal{T},\mathcal{I}})$
$\{P^1, A^1_{\mathcal{T},\mathcal{I}}\}$ $\leftarrow$Algorithm 1
for $k = 2$ to K **do**
  $\mathcal{M}^k \leftarrow$ index set for atoms in the current object and its augmenting atoms
  $\mathcal{N}^k \leftarrow$ index set for atoms in the rest of the objects that have not been visited
  if step k is an object residue step **then**
    $\{P^k, A^k_{\mathcal{T},\mathcal{M}^k}\}$ $\leftarrow$Algorithm 2
  else {step k is an atom residue step}
    $\{P^k, A^k_{\mathcal{T},\mathcal{M}^k}\}$ $\leftarrow$Algorithm 4
  end if
  $\{\mathcal{P}^k, A^k_{\mathcal{T},\mathcal{N}^k}\}$ $\leftarrow$Algorithm 3
end for
Output $\mathbf{A}_{\mathcal{I}}^0$ and $\{P^k : k = 1,\ldots,K\}$

Algorithm 5. Statistical training for multi-object ensemble.



(a) Template and training samples (b) Ensemble residue statistics

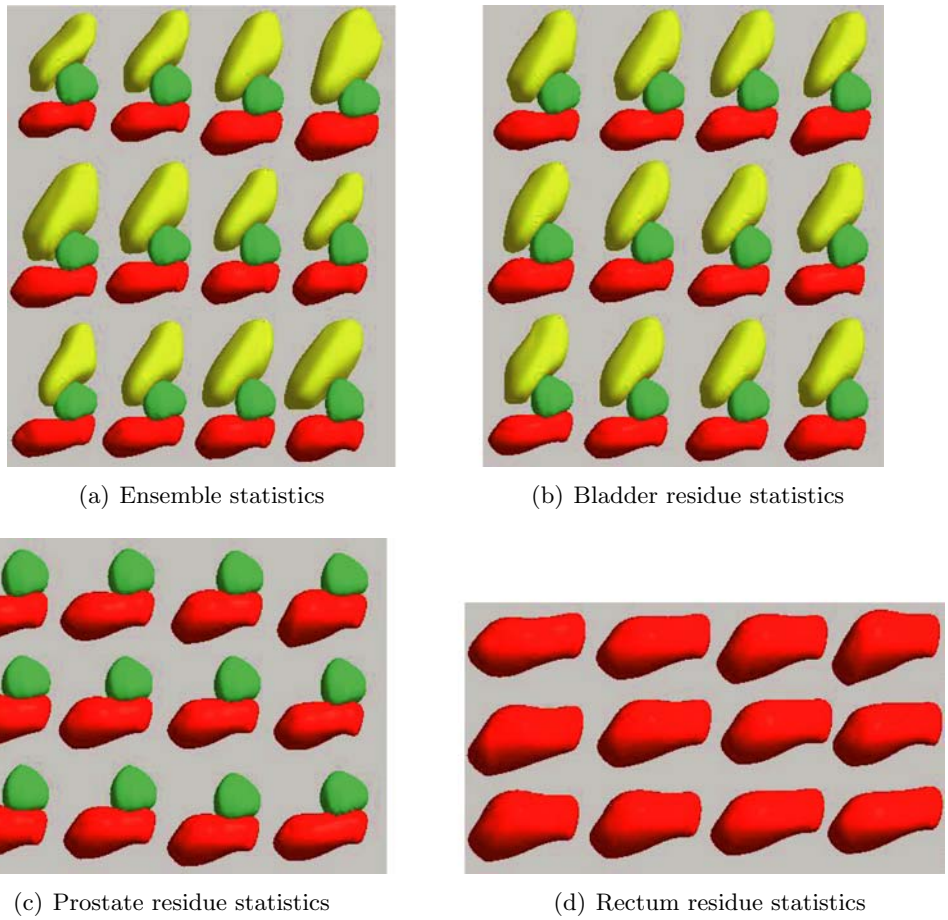
(c) Right bone object residue statistics (d) Left bone object residue statistics

Figure 7. Statistics of the two-bone ensemble. (a) Some training samples (right) and the estimated shape template (left). (b)–(d) The first three PGA modes of variation for the different stages. In each picture, from top to bottom: the first, second, and the third mode. From left to right, PGA deformations corresponding to -2, -1, $+1$, and $+2$ standard deviations.

simultaneous bending of the two bones; and the third mode corresponds to a twisting of the lower part of the two bones. In Fig. 7(c), the residual deformation of the right bone takes place mostly at both ends, especially at the one that is closer to the left bone. In contrary, the left bone deformation mostly happens at the end that is further from the right bone. This is to be expected because the deformation at the other end has been mostly described in the previous stage through the augmenting atoms.

In the second example, we show statistics for a bladder-prostate-rectum complex across different days within a particular patient. The medial atom grids used to fit the organs are: 4×6 for the bladder, 3×4 for the prostate, and 3×7 for the rectum, and we described the objects in the above order. The training set consists of 12 cases of one patient at different days. Figure 8 shows the ensemble statistics of the complex (a) and the residue statistics for each object (b–d).

The statistics can be used to generate random samples of the shape. We assume each principal mode follows the standard Gaussian distribution after we scale each principal direction by the square root of the corresponding eigenvalue in the tangent space. Although this is not exactly true, it is a reasonable approximation which gives us a way of quickly generating samples (refer to Section



(a) Ensemble statistics (b) Bladder residue statistics

(c) Prostate residue statistics (d) Rectum residue statistics

Figure 8. Statistics of the bladder-rectum-prostate complex. (a)–(d) The first three PGA modes of variation for the different stages. In each picture, from top to bottom: the first, second, and the third mode. From left to right, PGA deformations corresponding to -2, -1, $+1$, and $+2$ standard deviations.

5.1). Thus, random shape samples can be obtained by starting from the template and successively going through each stage as follows: at step k, generate multivariate normal samples on the tangent space about the corresponding mean μ^k, take the exponential map to produce m-reps residues, and concatenate them to the previous-step m-reps using the $\oplus$ operation. Figure 9 shows a few samples for the two-bone shape. We have found that when we stay within $[-2, +2]$ standard deviations most of the samples preserve topology, do not have geometric singularities and have very little inter-penetration among objects. In other words, with high likelihood the samples are geometrically proper (Pizer et al., 2005). Some quantitative evidence of this can be found in Pizer et al. (2006).

6. Segmentation Using the Multi-Object Statistics

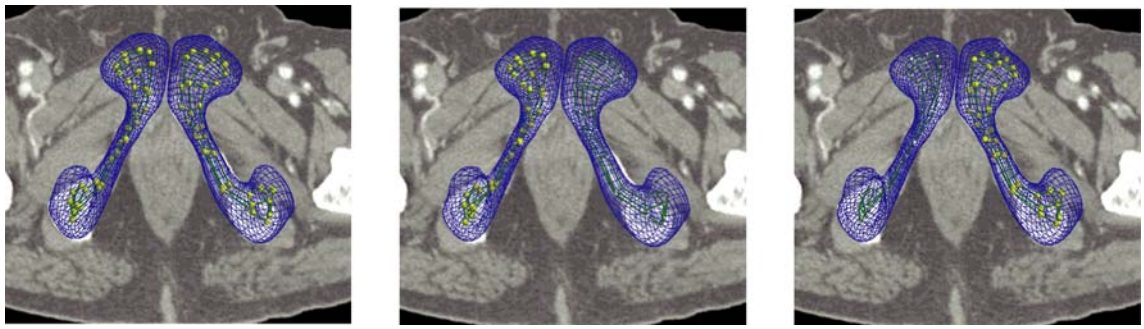
The above framework can be use in image segmentation, where we seek an m-rep model that best fits a given image I. In the Bayesian framework, this can be done by maximizing the log-posterior

$$\log p(A|I) \propto \log p(I|A) + \log p(A).$$

The prior distribution $p(A)$ is given by Eq. (9), and we use PGA statistics $\{P^k : k = 1, \ldots, K\}$ generated by the training process to approximate this prior. Here we assume a likelihood distribution $p(I|A)$ is available. For details on estimating this probability refer to Ho and Gerig (2004) and Stough et al. (2004).

To segment an image, we first apply a similarity transformation on $A^0_{\mathcal{I}}$ so that it roughly fits the image. Then we go through the object complex in the same order as when we obtained the statistics. At the object ensemble step, the ensemble geometry is deformed simultaneously by an overall similarity transformation and a transformation in the ensemble principal geodesic space H^1. At each subsequent step, we deform a particular object either at the object level or at the atom level, using the appropriate PGA statistics as prior, until all objects have been deformed.

Figure 9. Random samples of the two-bone shape.



(a) ensemble stage (b) right bone stage (c) left bone stage

Figure 10. Segmentation of the two-bone shape.

Since each atom residue step is described by a set of conditional distributions

$$\left\{\Pr\left(\Delta A_i^{\mathrm{k}} \middle| \Delta A_{N(i)}^{\mathrm{k}}\right)\right\} \sim \left\{\Pr\left(\Delta A_i^{\mathrm{k}} \ominus \Delta A_{i,p}^{\mathrm{k}}\right)\right\},$$

the optimization process at this step uses an Iterative Conditional Modes (ICM) algorithm (Besag, 1986).

Figure 10 shows an example of segmenting a two-bone shape from a CT image. Although this is a relatively simple application, it does illustrate how the method works and provides us with a basic test case. Our methodology has also been applied to some other more challenging problems, in particular segmentation of male pelvis, and have produced promising results (Pizer et al., 2005; Chaney et al., 2004).

7. Discussion

We illustrated a new method of describing shapes that consist of multiple geometric objects. Based on our m-reps representation, the geometric features of various scales and inter-object relationships are characterized by a multi-scale probabilistic model, which is in turn characterized by a series of probability distributions on residues. This method has the advantage of being able to provide accurate, intuitive geometric information and efficiency in statistical training. The shape model can be effectively used in a variety of image analysis applications, such as segmentation, shape discrimination, and atlas building.

As we have mentioned, the method described in the previous sections depends on the order we choose to traverse the multi-object complex. Now we briefly describe an alternative that does not specifically requires a predetermined order.

To explain, let us consider how we might use the probability model for segmentation. The idea is to maximize the joint posterior distribution of all objects, $\Pr(O_1, O_2, \ldots, O_M|I)$. In the method described earlier, we break this distribution into a product of distributions (see Eq. (4)). An alternative approach is to use the iterative conditional mode (ICM) method (Besag, 1986), where one iteratively maximizes the conditional distributions $\Pr(O_i|\{O_{j\neq i}\})$. It has been shown that under suitable conditions, the solution to this approach provides a good approximation to the maximizer of the original joint distribution. A method based on such iteration and methods of augmentation and prediction is described in Jeong et al. (2006).

Therefore, we can focus on learning the conditional distributions $\Pr(O_i|\{O_{j\neq i}\})$, which can be described by the difference between the actual deformation of object i and the predicted deformation of it based on those of its neighbors. Clearly, this approach does not depend on the order in which we choose to traverse the objects.

The basic assumptions in our methodology are that geometric entities can be described by a series of successive residual deformations from a common template and that these residual deformations can be treated as conditionally independent. Ultimately, the questions of in which

order we describe objects and how many PGA modes to use at each step come down to whether our choices make those assumptions valid. Depending on the situation, it may make sense to have schemes that have only global deformations or only local deformations, or have schemes going back and forth between scale levels.

The ideas of residual deformation and prediction can also be applied to objects made up of multiple m-reps figures, in which case the hinge atoms on the subfigures are the natural choices for augmenting atoms. For details refer to Han et al. (2004).

The methods described in this paper are applicable to any shape representation of objects (Pizer et al., 2005). In particular, boundary-based representations and atlas diffeomorphism representations can be used. However, the primitives of these representations typically directly describe position or displacements only, and thus inter-scale residues describe displacements only. Probability distributions on the primitives or their residues can access local orientational and size relations indirectly but not directly. It remains to be determined if the benefits of m-reps that they do access such relations directly provide more stable estimation of the probability distributions described in this paper, given the limited number of training samples that are frequently available.

Our tests have suggested that the statistical shape models learned from training sets produce visually valid samples. However, we need more rigorous ways of evaluating these probability models and comparing them with other shape models. This is still work in progress. Other questions that require further investigation include the following: method for determining augmenting atoms, prediction algorithms, and object-based image match models.

Acknowledgments

We thank Gregg Tracton, Derek Merck, Julian Rosenman for providing image data and preparing some of the m-reps segmentations of them. We also thank P. Thomas Fletcher, Qiong Han, Keith Muller, Guido Gerig, and Sean Ho for their help. This work was done with support from NIH grant P01 EB02779.

Notes

1 The subscript 2 here corresponds to the step number, not the object index.

References

Besag, J. 1986. On the statistical analysis of dirty pictures. *J. Royal Stat. Soc. B*, 48(3):259–302.

Blum, H., and Nagel, R. 1978. Shape description using weighted symmetric axis features. *Pattern Recognition*, 10(3):167–180.

Chaney, E., Pizer, S.M., Joshi, R., Broadhurst, S., Flether, P.T., Gash, G., Han, Q., Jeong, J., Lu, C., Merck, D., Stough, J., Tracton, J., Rosenman, J., Chi, Y., and Muller, K. 2004. Automatic male pelvis segmentation from ct images via statistically trained multi-object deformable m-rep models. In *American Society for Therapeutic Radiology and Oncology (ASTRO)*.

Cootes, T.F., Beeston, C., Edwards, G.J., and Taylor, C.J. 1999. A unified framework for atlas matching using active appearance models. *Information Processing in Medical Imaging*. *LNCS* 1613:322–333.

Cootes, T.F., Edwards, G.J., and Taylor, C.J. 1998. Active appearance models. In *Fifth European Conference on Computer Vision*, pp. 484–498.

Cootes, T.F., Taylor, C.J., Cooper, D.H., and Graham, J. 1995. Active shape models—their training and application. *Computer Vision and Image Understanding*, 61(1):38–59.

Davatzikos, C., Xiaodong, T., and Shen, D. 2003. Hierarchical active shape models, using the wavelet transform. *IEEE Transactions on Medical Imaging*, 22(3):414–423.

Davies, R.H., Twining, C.J., Cootes, T.F., Waterton, J.C., and Taylor, C.J. 2002. A minimum description length approach to statistical shape modeling. *IEEE Transactions on Medical Imaging*, 21(5):525–537.

Fletcher, P.T. and Joshi, S. 2004. Principal geodesic analysis on symmetric spaces: statistics of diffusion tensors. In *ECCV 2004 Workshop on Computer Vision Approaches to Medical Image Analysis (CVAMIA)*.

Fletcher, P.T., Lu, C., and Joshi, S. 2003. Statistics of shape via principal component analysis on Lie groups. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, Los Alamitos, CA, pp. 95–101.

Fletcher, P.T., Lu, C., Pizer, S.M., and Joshi, S. 2004. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans. Medical Imaging*, 23(8):995–1005.

Gerig, G., Styner, M., Shenton, M.E., and Lieberman, J.A. 2001. Shape versus size: improved understanding of the morphology of brain structures. In *Proc. MICCAI 2001*, volume 2208 of *Springer LNCS* Springer-Verlag, pp. 24–32.

Goodall, C. 1991. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society*, 53(2):285–339.

Grenander, U. 1995. *Elements of Pattern Theory*. Johns Hopkins University Press.

Grenander, U., Chow, Y., and Keenan, D.M. 1991. *HANDS: A Pattern Theoretic Study of Biological Shapes*. Springer-Verlag, New York.

Han, Q., Lu, C., Liu, G., Pizer, S., Joshi, S., and Thall, A. 2004. Representing multi-figure anatomical objects. In *International Symposium on Biomedical Imaging*, Germany, pp. 1251–1254.

Ho, S., and Gerig, G. 2004. Profile scale-space for multi-scale image match. In C. Barillot, D. Haynor, and P. Hellier (eds.), *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Germany, pp. 176–183.

Jeong, J., Pizer, S., and Ray, S. 2006. Statistics on anatomic objects reflcting inter-object relations. In *MICCAI Conference: Workshop on Mathematical Foundations of Computational Anatomy*, pp. 136–145.

Joshi, S. 1997. Large deformation diffeomorphisms and Gaussian random fields for statistical characterization of brain submanifolds. PhD thesis, Washington University.

Joshi, S., Pizer, S., Fletcher, P.T., Yushkevich, P., Thall, A., and Marron, J.S. 2002. Multi-scale deformable model segmentation and statistical shape analysis using medial descriptions. *IEEE-TMI*, 21(5).

Kapur, T., Beardsley, P.A., Gibson, S.F., Grimson, W.E.L., and Wells, W.M. 1998. Model based segmentation of clinical knee MRI. Model-based 3D Image Analysis workshop (in conjuction with ICCV).

Kass, M., Witkin, A., and Terzopoulos, D. 1987. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331.

Kelemen, A., Szekely, G., and Gerig, G. 1999. Three-dimensional model-based segmentation. *IEEE-TMI*, 18(10):828–839.

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

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

Lu, C., Pizer, S.M., and Joshi, S. 2003. A markov random field approach to multi-scale shape analysis. In L.D. Griffin and M. Lillholm (eds.), *Scale Space Methods in Computer Vision*, volume 2695 of *LNCS*, Springer-Verlag, pp. 416–431.

Mallat, S.G. 1989. Multifrequency channel decompositions of images and wavelet models. *IEEE Trans. Acoust. Speech, Signal Processing*, 37(12):2091–2110.

Merck, D., Tracton, G., Pizer, S.M., and Joshi, S. A methodology for constructing geometric priors and likelihoods for deformable shape models. http://midag.unc.cs.edu/pubs/tech-rpts/Merck06_submission.pdf.

Pizer, S.M., Broadhurst, R., Jeong, J., Han, Q., Saboo, R., Stough, J., Tracton, G., and Chaney, E. 2006. Intra-patient anatomic statistical models for adaptive radiotherapy. In *MICCAI Workshop From Statistical Atlases to Personalized Models: Understanding Complex Diseases in Populations and Individuals*, pp. 43–46.

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., and Chaney, E.L. 2003a. Deformable m-reps for 3D medical image segmentation. *International Journal of Computer Vision*, 55(2):85–106.

Pohl, K.M., Fisher, J., Kikinis, R., Grimson, W.E.L., and Wells, W.M. 2005. Shape based segmentation of anatomical structures in magnetic resonance images. *International Conference on Computer Vision, LNCS* 3765:489–498.

Pohl, K.M., Fisher, J., Levitt, J.L., Shenton, M.E., Kikinis, R., Grimson, W.E.L., and Wells, W.M. 2005. A unifying approach to registration, segmentation, and intensity correction. *Medical Image Computing and Computer Assisted-Intervention, LNCS* 3749:310–318.

Pizer, S.M., Fletcher, P.T., Thall, A., Styner, M., Gerig, G., and Joshi, S. 2003c. Object models in multi-scale intrinsic coordinates via m-reps. *Image and Vision Computing, Special Issue on Generative Model-based Vision*, 21(1):5–15.

Pizer, S.M., Siddiqi, K., Szekely, G., Damon, J.N., and Zucker, S.W. 2003b Multiscale medial loci and their properties. *Int. J. Computer Vision*, 55(2):155–179.

Pizer, S.M., Jeong, J., Lu, C., Muller, K., and Joshi, S. 2005. Estimating the statistics of multi-object anatomic geometry using inter-object relationships. In O.F. Olsen, L. Florack, and A. Kuijer, (eds.), *International Workshop on Deep Structure, Singularities and Computer Vision (DSSCV)*, volume 3753 of *LNCS*, Springer-Verlag, pp. 60–71.

Siddiqi, K., Bouix, S., Tannenbaum, A.R., and Zuker, S.W. 2002. Hamilton-Jacobi skeletons. *Int. J. Computer Vision*, 48(3):215–231.

Stough, J., Pizer, S.M., Chaney, E., and Rao, M. 2004. Clustering on image boundary regions for deformable model segmentation. In *International Symposium on Biomedical Imaging (ISBI)*, Piscataway, NJ, pp. 436–439.

Styner, M., and Gerig, G. 2001. *IPMI '01*, volume 2082 of *LNCS*, chapter Medial models incorporating object variability for 3D shape analysis, Springer, pp. 502–516.

Thall, A. 2004. Deformable Solid Modeling via Medial Sampling and Displacement Subdivision. PhD thesis, University of North Carolina, Chapel Hill.

Tsai, A., Wells, W., Tempany, C., Grimson, E., and Willsky, A. 2003. Coupled multi-shape model and mutual information for medical image segmentation. *Information Processing in Medical Imaging*. *LNCS* 2732:185–197.

Tsai, A., Yezzi, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, E., and Willsky, A. 2003. A shape-based approach to curve evolution for segmentation of medical imagery. *IEEE T-MI*, 22(2):137–154.

Unser, M. 1996. A review of wavelets in biomedical applications. *Proceedings of the IEEE*, 84(4):626–638.

Vaillant, M. and Davatizikos, C. 1999. Hierarchical matching of cortical features for deformable brain image registration. *Information Processing in Medical Imaging*, *LNCS* 1613:182–195.

Yushkevich, P., Pizer, S.M., Joshi, S., and Marron, J.S. 2001. Intuitive, localized analysis of shape variability. In M.F. Insana and R.M. Leahy (eds.), *Information Processing in Medical Imaging (IPMI)*, pp. 402–408,

Zhu, S.C. 1999. Embedding Gestalt laws in the Markov random fields—a theory for shape modeling and perceptual organization. *IEEE Trans. Pattern Anal. Mach. Intell.*, 21(11).