

An Adaptive-Focus Statistical Shape Model for Segmentation and Shape Modeling of 3-D Brain Structures

Dinggang Shen*, Edward H. Herskovits, and Christos Davatzikos

Abstract—This paper presents a deformable model for automatically segmenting brain structures from volumetric magnetic resonance (MR) images and obtaining point correspondences, using geometric and statistical information in a hierarchical scheme. Geometric information is embedded into the model via a set of affine-invariant attribute vectors, each of which characterizes the geometric structure around a point of the model from a local to a global scale. The attribute vectors, in conjunction with the deformation mechanism of the model, warranty that the model not only deforms to nearby edges, as is customary in most deformable surface models, but also that it determines point correspondences based on geometric similarity at different scales. The proposed model is adaptive in that it initially focuses on the most reliable structures of interest, and gradually shifts focus to other structures as those become closer to their respective targets and, therefore, more reliable. The proposed techniques have been used to segment boundaries of the ventricles, the caudate nucleus, and the lenticular nucleus from volumetric MR images.

Index Terms—Active contour, adaptive focus model, attribute vectors, brain image segmentation, deformable model, deformable registration, snake, statistical shape model.

I. INTRODUCTION

DEFORMABLE contour and surface models [1], [2] have been extensively used as segmentation tools in medical or biological imaging applications, where the objects (or structures) to be analyzed undergo deformations [24]. An excellent review of deformable models can be found in [3].

Most boundary-based deformable models adapt to nearby edges under forces emanating from the immediate neighbors and from image gradients. This can cause unrealistic deformations as individual points are pulled toward noisy or fragmented edges. Moreover, it will make these models very sensitive to initialization. Most importantly, deforming the model toward a nearby boundary provides a means for segmenting and labeling structures, but not necessarily for establishing point-correspondences among different individuals.

Manuscript received June 12, 2000; revised December 22, 2000. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was N. Ayache. *Asterisk indicates corresponding author.*

*D. Shen is with the Department of Radiology, JHOC 3230, Johns Hopkins University, 601 N. Caroline Street, Baltimore, MD 21287 USA (e-mail: dgshen@cbmv.jhu.edu).

E. H. Herskovits is with the Department of Radiology, Johns Hopkins University, Baltimore, MD 21287 USA.

C. Davatzikos is with the Department of Radiology, the Department of Computer Science, and the Center for Computer-Integrated Surgical Systems and Technology, Johns Hopkins University, Baltimore, MD 21287 USA.

Publisher Item Identifier S 0278-0062(01)02776-8.

To remedy the aforementioned undesired situations, shape preserving deformable models [30]–[32] have been proposed. In an alternative formulation, the deformation of a model is restricted by prior knowledge in the form of normal statistical variation [4], [5]. Chen and Kanade [27] used the statistics of anatomical variations as prior knowledge to guide the process of registering the statistical atlas with a particular subject. Cootes *et al.* [6], [7] have developed a technique for building compact models of the shape and appearance of variable structures in two-dimensional (2-D) images, based on the statistics of labeled images containing examples of the objects. Each model consists of a flexible shape template describing how the relative locations of important points on the shapes can vary, and a statistical model of the expected gray-levels in a region around each point. An extended version of this technique was presented in [28]. Following the seminal work of Cootes *et al.*, a flexible parametric surface model [8], [9] was proposed based on a hierarchical parametric object description rather than a point distribution model. This model has been used to segment 2-D and three-dimensional (3-D) structures from tomographic brain images. Finally, some researchers consider capturing statistical information from the covariance matrix that has been preprocessed by prior smoothness constraints [10].

The aforementioned statistical models can also be implemented hierarchically, since hierarchical implementation usually increases the likelihood of finding the globally optimal match by performing the calculations of finer details of the deformation after matching on a more global scale. A hierarchical statistical modeling framework, which combines the advantage of a compact description of global deformations along with the accurate description of local deformations, has also been developed for representation, segmentation, and tracking of 2-D deformable structures in image sequences [11]. A review of the hierarchical strategies can be found in [12].

In this paper we present a deformable model for segmentation and definition of point correspondences in brain images, which incorporates geometric as well as statistical information about the shapes of interest, in a hierarchical fashion. Our deformable model is comprised of several interconnected surfaces, each corresponding to a structure of interest. Our methodology has three novel aspects, which are briefly described next.

First, an *attribute vector* is attached to each point of the model, and it is used to characterize the geometric structure of the model around that point, from a local to a global scale. Our model has the tendency to preserve its geometric shape while deforming, based on an energy term that favors similar attribute

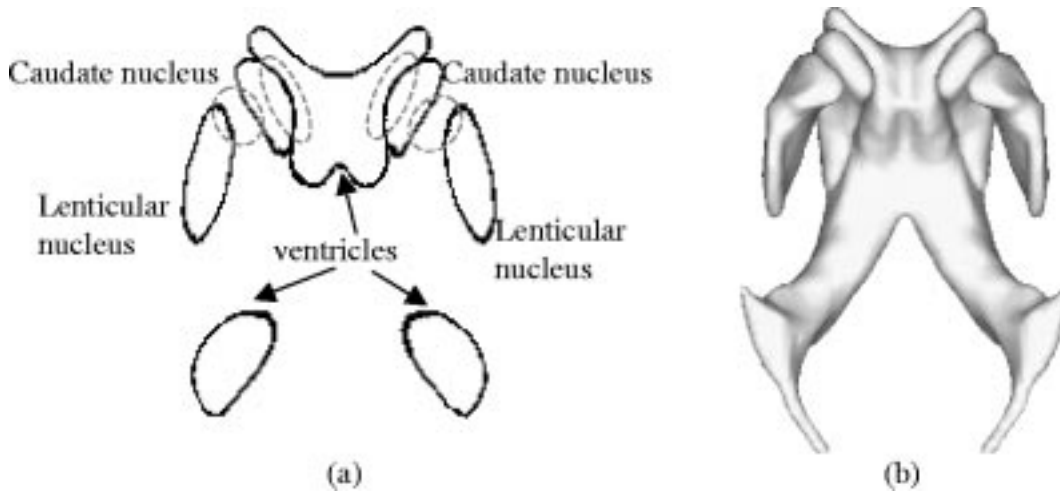


Fig. 1. A model with five surfaces. (a) Cross section of the 3-D model of (b). In (a), any vertex in the dotted ellipses is connected with another vertex of the model.

vectors. The attribute vectors are essential in our formulation, since they distinguish different parts of a boundary according to their shape properties. If rich enough, attribute vectors can uniquely characterize their respective boundary segments.

The second contribution of our model is in its adaptive formulation. In particular, the deformable model can shift focus from one structure to another, or potentially from one part of a structure to another. For example, the model can initially focus on the most reliable parts of the shape and subsequently shift focus to other parts as they become closer to their respective targets and, therefore, more reliable. The model adaptivity is accomplished both via its hierarchical deformation strategy and via an *adaptive focus* statistical shape model. The latter is in contrast to previous statistical shape models [6], in which all landmarks in the training samples have the same level of importance. In those models, larger features of a shape dominate over relatively smaller, yet important features, merely because their large sizes and, therefore, influence the measures of shape variability disproportionately. Furthermore, unreliable features, if they are large, will dominate over relatively smaller reliable and important features. The proposed adaptive focus statistical model accounts for size differences between different structures when determining the parameters of shape variation. Moreover, it allows the algorithm to selectively focus on certain structures, by biasing the statistics of the model by the statistics of the structures of interest.

Finally, our third contribution is in the training of the statistical shape model. In particular, we build our surface models in a way that point correspondences are determined by the algorithm for the training samples, based on hand-segmented images. Establishing point correspondences for the training samples is readily done in two dimensions, but it is very difficult in three dimensions. Consequently, previous work in three dimensions has relied largely on parametric representations that are not necessarily based on underlying point correspondences [5], [8], [9].

The paper is organized as follows. Section II describes an adaptive-focus deformable model (AFDM) for the case that no training set is available. In Section III, AFDM is extended to be adaptive focus deformable statistical shape model (AFDSM),

by incorporating information about the statistical variation of the model, obtained from a training set. Experiments on testing the performances of both AFDM and AFDSM are demonstrated in Section IV. Finally, conclusion and future work are presented in Section V.

II. ADAPTIVE-FOCUS DEFORMABLE MODEL

In our approach, we first construct a model that represents the typical anatomy of a number of structures (e.g., Fig. 1). Since several structures can be included in the model, we incorporate several interconnected surfaces, as in Fig. 1. In Section II-A, we describe the details pertaining to the surface construction and interconnection. In Section II-B, we describe an attribute vector associated with each vertex of these surfaces, which reflects the geometric structure of each surface from a global to a local level. In Sections II-C and II-D, we describe the mechanism that deforms the model to an individual magnetic resonance (MR) volumetric image, and the corresponding energy function being minimized. In Section II-E, we present a hierarchical and adaptive-focus strategy for the model deformation. Finally, in Section II-F, we use a learning algorithm to extract adaptive-focus knowledge from the training samples.

A. Model Description

The generalized model used in this paper consists of several separate surfaces that can be open or closed. Each surface is represented by its own set of vertices and triangles. In the following, we will give definitions for vertices and their neighborhood layers on each surface, and as well as on the interconnections of the surfaces of the model.

Let us assume that there are M separate surfaces V^1, \dots, V^M in the model, and that the i th surface V^i has N_i vertices. Let V_j^i be the j th vertex on the i th surface, where $1 \leq i \leq M$ and $0 \leq j \leq (N_i - 1)$. In the 3-D space, the coordinates of vertex V_j^i can be denoted as $V_j^i = [x_j^i \ y_j^i \ z_j^i]^T$.

For each vertex V_j^i in the model, the first *neighborhood layer* includes its immediate neighbors. The second neighborhood layer includes the immediate neighbors of the immediate neighbors, and so forth. The neighborhood layers

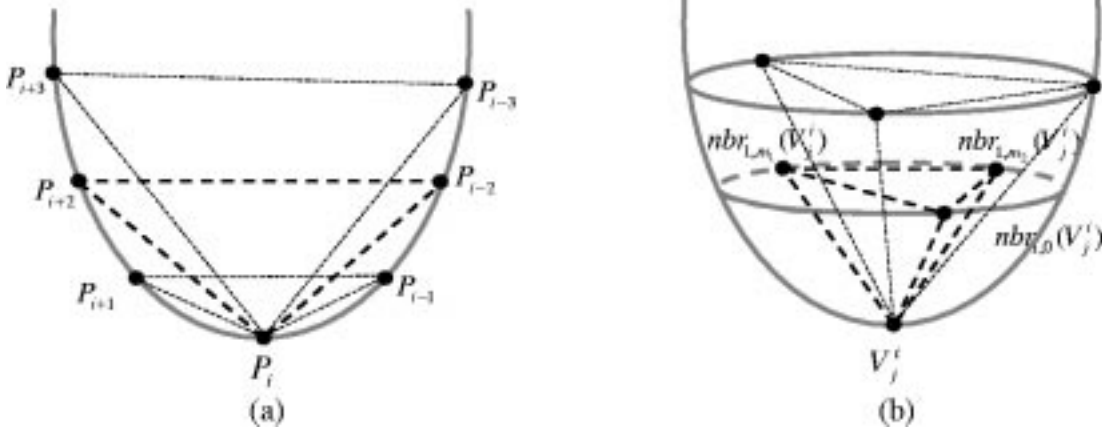


Fig. 2. The attribute vector in two dimensions and three dimensions. (a) The area of a triangle formed by three points P_{i-v_s} , P_i , and P_{i+v_s} , is used as the v th element of the attribute vector. (b) Attribute vector for the 3-D vertex V_j^i . The volume of a tetrahedron, formed by four vertices V_j^i , $nbr_{l,0}(V_j^i)$, $nbr_{l,m_1}(V_j^i)$, and $nbr_{l,m_2}(V_j^i)$, is used as the l th element of the attribute vector $f_l(V_j^i)$. Here, m_1 and m_2 are selected as $m_1 = \lfloor (S_l(V_j^i))/3 \rfloor$ and $m_2 = \lfloor (S_l(V_j^i))/3 \rfloor \times 2$.

are constructed so that no vertex is repeated twice in the neighborhood of another vertex. For example, vertex V_j^i is assumed to have $R(V_j^i)$ neighborhood layers, and its l th neighborhood layer $nbr_l(V_j^i)$ contains $S_l(V_j^i)$ vertices. The neighborhood layer l is in the range $1 \leq l \leq R(V_j^i)$. Let $nbr_{l,m}(V_j^i)$ denote the m th vertex in the l th neighborhood layer, then the l th neighborhood layer can be represented by $nbr_l(V_j^i) = \{nbr_{l,m}(V_j^i) | 0 \leq m < S_l(V_j^i)\}$.

In certain applications, some surfaces in the model are in proximity to each other. Therefore, we impose additional constraints that prevent these surfaces from intersecting during the deformation procedure. In particular, if the 3-D Euclidean distance between vertices belonging to two different surfaces is below a threshold, the vertices are connected as first-layer neighbors. For example, in Fig. 1 the caudate nucleus is very close to, or in contact with, the ventricular surface. Then, for each vertex of the boundary of the caudate nucleus that is in contact with the ventricular surface, an additional neighbor is selected as its closest vertex in the ventricular surface. Fig. 1(a) shows a slice of the 3-D model given in Fig. 1(b). If two vertices belonging to the same surface are close enough, based on their 3-D Euclidean distance, they are also joined with each other. This helps prevent self-intersections of the surfaces of the model.

B. Affine-Invariant Attribute Vector

1) *The Attribute Vector in Two Dimensions:* As we mentioned earlier, an attribute vector is attached to each vertex of the model, which reflects the model's geometric structure from a local to a global level. For clarity, we first describe the attribute vector in two dimensions. The attribute vector is, in some ways, an extension to curvature, which has been widely used in the recognition of objects from the digital images. As a local characteristic able to carry information at multiple resolutions, curvature has been applied in shape matching [13], shape analysis [14], and object recognition [15]. Curvature is invariant under rotation and translation of the shape, and can easily be normalized with respect to scale changes. However, curvature is not invariant under affine transformation [26].

In order to describe shape characteristics of various scales, an affine-invariant vector of geometric attributes was used in [17], [21]. Each attribute is the area of a triangle formed by a point, P_i , of the model and its two neighboring points, P_{i-v_s} and P_{i+v_s} , which correspond to the v s-th neighborhood layer [see Fig. 2(a)]. Two-dimensional attribute vectors have been successfully applied to a number of problems, such as shape matching and indexing [18], model-based adaptive image segmentation [17], skewed symmetries detection [19], and affine-invariant detection of perceptually parallel curves [20]. In the following, we will extend the definition of the 2-D affine-invariant attribute to the 3-D. The volume of a tetrahedron will be used as an affine-invariant attribute in the 3-D case.

2) *The Attribute Vector in Three Dimensions:* In three dimensions, each attribute is the volume of a tetrahedron [see Fig. 2(b)]. The volume of the tetrahedron, formed by the nearest neighbors of vertex V_j^i reflects the local structure of the surface around vertex V_j^i . The volumes of larger tetrahedrons represent more global properties of the surface around vertex V_j^i . It is not hard to see that the attribute vector corresponding to, say, a high-curvature region is completely different from attribute vectors of flat segments of the surface. More importantly, even vertices of similar curvatures might have very different attribute vectors, depending on the number of neighborhood layers, or equivalently the number of components of the attribute vector. We note that additional attributes, such as depth or tissue membership functions, can also be incorporated into the attribute vector, depending on the application.

Any four points in the 3-D space can establish a tetrahedron. Let us generally denote these four points by $\{W_j | j = 0, 1, 2, 3\}$, where $W_j = [x_j \ y_j \ z_j]^T$. The volume of the tetrahedron formed by four points $\{W_j | j = 0, 1, 2, 3\}$ is given by the determinant of a 4×4 matrix

$$Volume(W_0, W_1, W_2, W_3) = \begin{vmatrix} x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \\ z_0 & z_1 & z_2 & z_3 \\ 1 & 1 & 1 & 1 \end{vmatrix}.$$

If these four points are linearly transformed by a 4×4 matrix A , then the volume of the new tetrahedron is equal to $|A| \cdot Volume(W_0, W_1, W_2, W_3)$. The value of volume is relatively

invariant to the linear transformation and can be made absolutely invariant after an appropriate normalization [17].

The definition of the volume of a tetrahedron can be used to design an attribute vector for each vertex on the model surface. For a particular vertex V_j^i , we can select any three points from the l th neighborhood layer [see Fig. 2(b)]. The volume of the tetrahedron formed by these four vertices is defined by $f_l(V_j^i)$. We compile the volumes calculated for different neighborhood layers into an attribute vector for vertex V_j^i

$$F(V_j^i) = [f_1(V_j^i) \quad f_2(V_j^i) \quad \dots \quad f_{R(V_j^i)}(V_j^i)]$$

where $R(V_j^i)$ is the number of neighborhood layers around vertex V_j^i . When l is small, the value of the attribute element $f_l(V_j^i)$ reflects local shape information of vertex V_j^i . As l increases, the value of the attribute element $f_l(V_j^i)$ begins to capture more global shape information of vertex V_j^i . Therefore, the attribute vector $F(V_j^i)$ captures different levels of shape information around vertex V_j^i .

The definition of the attribute vector can be made affine-invariant, by normalizing it on the whole model, i.e.,

$$\hat{F}(V_j^i) = \frac{F(V_j^i)}{\sum_{i=1}^M \sum_{j=1}^{N_i} \sum_{l=1}^{R(V_j^i)} |f_l(V_j^i)|}$$

where $\hat{F}(V_j^i) = [\hat{f}_1(V_j^i) \quad \hat{f}_2(V_j^i) \quad \dots \quad \hat{f}_{R(V_j^i)}(V_j^i)]$. Unlike curvature, the normalized attribute vectors are affine-invariant.

C. Energy Definition

The goal of our deformable model is to define point correspondences, in addition to segmenting structures of interest. Our premise is that the attribute vector, if rich enough, uniquely characterizes different parts of a boundary of a structure. Therefore, in the definition of the energy function to be minimized, we include a term that reflects the difference between the attribute vectors of the model and individual surface. An obvious difficulty in this approach is that the attribute vector of an individual surface cannot be obtained directly from the corresponding MR images, since it is based on a triangularized surface. We overcome this difficulty by deforming our model via a sequence of global and local transformations. Since the attribute vectors are invariant to linear transformation, they remain relatively unchanged in this deformation process. Hence, the model tends to gradually adapt to an individual boundary, yet it does so in a way that its attribute vectors remain relatively unchanged. Effectively, this defines a segmentation, but also a set of point correspondences based on a similarity between attribute vectors. Our deformation strategy is very robust to local minima, since it deforms surface segments at a time, and not individual vertices.

The energy that our deformable model minimizes is defined as follows:

$$E = \sum_{i=1}^M \sum_{j=1}^{N_i} \omega_{i,j} E_{i,j} = \sum_{i=1}^M \sum_{j=1}^{N_i} \omega_{i,j} (E_{i,j}^{\text{model}} + E_{i,j}^{\text{data}}). \quad (1)$$

The weighting parameter $\omega_{i,j}$ determines the relative weight given to the local energy term $E_{i,j}$. The local energy term $E_{i,j}$, defined for the j th vertex on the i th surface (V_j^i), is composed of two terms: $E_{i,j}^{\text{model}}$ and $E_{i,j}^{\text{data}}$. The term $E_{i,j}^{\text{model}}$ defines the degree of difference between the model and its deformed configuration, around vertex V_j^i . The term $E_{i,j}^{\text{data}}$ defines the external energy, aiming at deforming the segment around the vertex V_j^i toward a boundary of the image.

As we elaborated earlier in this section, the term $E_{i,j}^{\text{model}}$ reflects the difference between the attribute vectors of the model and its deformed configuration at V_j^i , and it is given by

$$E_{i,j}^{\text{model}} = \sum_{l=1}^{R(V_j^i)} \delta_l \left(\hat{f}_l^{\text{Def}}(V_j^i) - \hat{f}_l^{\text{Mdl}}(V_j^i) \right)^2 \quad (2a)$$

where $\hat{f}_l^{\text{Def}}(V_j^i)$ and $\hat{f}_l^{\text{Mdl}}(V_j^i)$ are, respectively, the components of the (normalized) attribute vectors of the deformed model configuration and the model at vertex V_j^i . The parameter δ_l denotes the degree of importance of the l th attribute element $\hat{f}_l^{\text{Def}}(V_j^i)$ (or the l th neighborhood layer) in the surface segment under consideration. Notice that $R(V_j^i)$ is the number of geometric attributes, and is equal to the number of the neighborhood layers around vertex V_j^i .

The data energy term, $E_{i,j}^{\text{data}}$, is usually designed to move the deformable model toward an object boundary. Accordingly, for every vertex V_j^i , we require that in the position of V_j^i , the magnitude of image gradient should be high, and the direction of image gradient should be similar to the normal vector of the deformed surface. Since our deformation mechanism, which is defined in Section II-D, deforms a surface segment around each vertex V_j^i at a time, and not just the vertex itself, we want to design an energy term that reflects the fit of the whole segment, rather than a single vertex, with image edges. A surface segment is defined by vertex V_j^i and its neighbors from $R(V_j^i)$ neighborhood layers, where $R(V_j^i)$ can vary throughout the deformation procedure, as detailed in Section II-E. The l th neighborhood layer has $S_l(V_j^i)$ vertices, $\text{nbr}_l(V_j^i) = \{\text{nbr}_{l,m}(V_j^i) | 0 \leq m < S_l(V_j^i)\}$, where $\text{nbr}_{l,m}(V_j^i)$ denotes the m th vertex in the l th neighborhood layer. The data energy term $E_{i,j}^{\text{data}}$ is designed as follows.

$$E_{i,j}^{\text{data}} = \sum_{l=1}^{R(V_j^i)} \delta_l \sum_{m=0}^{S_l(V_j^i)} \left(1 - |\nabla I(\text{nbr}_{l,m}(V_j^i))| \cdot \left| \vec{H}(\text{nbr}_{l,m}(V_j^i)) \cdot \vec{n}(\text{nbr}_{l,m}(V_j^i)) \right| \right) \quad (2b)$$

where $|\nabla I(\text{nbr}_{l,m}(V_j^i))|$, valued between zero and one, is the normalized magnitude of the gradient at vertex $\text{nbr}_{l,m}(V_j^i)$; $\vec{h}(\text{nbr}_{l,m}(V_j^i))$ is the direction of the gradient; $\vec{n}(\text{nbr}_{l,m}(V_j^i))$ is the normal vector of the deformed model for vertex $\text{nbr}_{l,m}(V_j^i)$, directed toward the model's interior. The parameter δ_l denotes the weight of the l th neighborhood layer.

The whole energy function of the snake is then given by (1), (2a) and (2b).

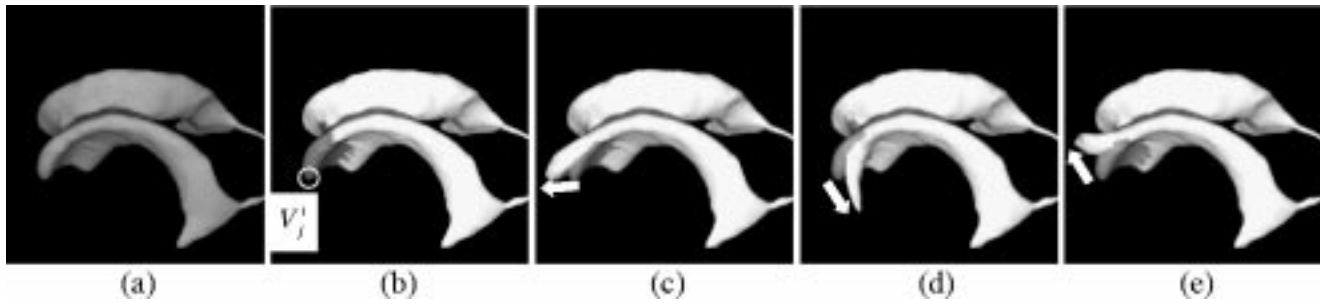


Fig. 3. Demonstration of the deformation mechanism used in AFDM. A ventricular model is shown as gray surface on the left. Tentative deformations of a surface segment around V_j^i are shown in the rest of the image, with the original model shown as gray and the deformed model shown as white surface.

D. Local Deformation Mechanism

We now describe a greedy deformation algorithm for the minimization of the energy function in (1). For the standard snake, the greedy algorithm moves the vertex to the position in its neighboring region that makes the snake energy minimal. For our method, we suggest considering the deformation of a surface segment as whole, which greatly helps the snake avoid local minima.

At each iteration, a surface segment around a vertex can be deformed by an affine transformation. Recall that, under affine transformations the value of the energy term $E_{i,j}^{\text{model}}$ remains relatively unchanged. Accordingly, the new configuration of a particular surface segment can be determined directly by minimizing an energy term $E_{i,j}^{\text{data}}$ [see (2b)]. In the 2-D case, the transformation matrix can be easily calculated from the motion of the studied point [21]. However, in three dimensions, this calculation becomes much more complicated, as described next.

Since we deform only one piece of the model surface at a time, we can introduce discontinuities at the boundary of the segment being deformed. In our 2-D model [21], we solved this issue by restricting the local affine transformation so that it leaves the end-points of a deforming segment unchanged, thereby maintaining continuity of the deformable contour. However, for a surface model this is not possible, because the vertices belonging to the $R(V_j^i)$ th neighborhood layer, which is the layer farthest away from V_j^i , do not necessarily lie on the same plane. Therefore, we cannot necessarily find a local affine transformation that preserves the position of the end-vertices of a deforming segment. In order to remedy this situation, we used a different form of transformation for each deforming surface segment, which is described next.

Let V_j^i be the vertex whose neighborhoods form the surface segment to be deformed at a particular iteration (see Fig. 3). The $R(V_j^i)$ th neighborhood layer forms the boundary of the surface segment. Consider a tentative position, $V_j^i + \Delta V$, to which V_j^i is to move during the greedy algorithm. Then, the new position of each vertex, $nbr_{l,m}(V_j^i)$, in the segment is defined as

$$nbr_{l,m}(V_j^i) + \Delta V \cdot \exp\left(-\frac{l^2}{2\sigma^2}\right),$$

where σ is a parameter determining the locality of the transformation. We use values of σ that make $\exp(-(R(V_j^i))^2)/(2\sigma^2)$ close to zero, effectively leaving the bounding curve of a deforming segment unchanged and, hence, maintaining continuity.

The new configuration of the surface segment is then determined by finding ΔV that minimize the sum of two energy terms $E_{i,j}^{\text{data}}$ and $E_{i,j}^{\text{model}}$.

Fig. 3 demonstrates some tentative positions of vertex V_j^i and the corresponding deformations of the surface segment. The gray surface in Fig. 3 (left), is a ventricular model. The rest of the images in Fig. 3 show tentative deformations of the ventricular model (white surfaces) overlaid on the undeformed model (gray surface).

E. Adaptive-Focus Deformation Strategy

Brain images contain several boundaries. Prior knowledge, in conjunction with the quality of image information (e.g., edge strength), is used in AFDM to guide the deformation of the model in a hierarchical fashion. In particular, surfaces for which we have relatively higher confidence are deformed first. As other surfaces follow this deformation and get closer to their respective targets, they become more reliable features for driving the model's deformation. We demonstrate this scheme using the example of the caudate nucleus (CN), the lenticular nucleus (LN) and the ventricular boundaries. In Fig. 4 we show cross sections of the initial (automatic) placement of a model containing these five surfaces, and the deformation of that model after ten iterations. The one in Fig. 4(b) is the result of AFDM, with the ventricular boundaries deforming first, and the CN and LN boundaries following. In fact, there was a continuous blending in the deformation of the CN and LN as iteration number increased. The result of Fig. 4(c) was obtained via the same model but with a nonadaptive deformation mechanism, i.e., with forces applied to all components of the model simultaneously. In the adaptive focus scheme, the ventricles first pulled the LN close enough to its corresponding boundary in the MR image, before the LN model started deforming. In the nonadaptive scheme, however, the LN deformed toward the wrong (the cortical) boundary.

In addition to its cross-component hierarchical formulation, our approach is also hierarchical within-components of the model. In particular, the parameter $R(V_j^i)$ that determines the locality of the deformation transformation is typically chosen large in the initial iterations, and is gradually reduced to one. Therefore, initially, relatively more vertices are involved in the surface segment around vertex V_j^i , and the resulting transformation is of relatively global form. In later stages, the transformation affects the deformable model more locally.

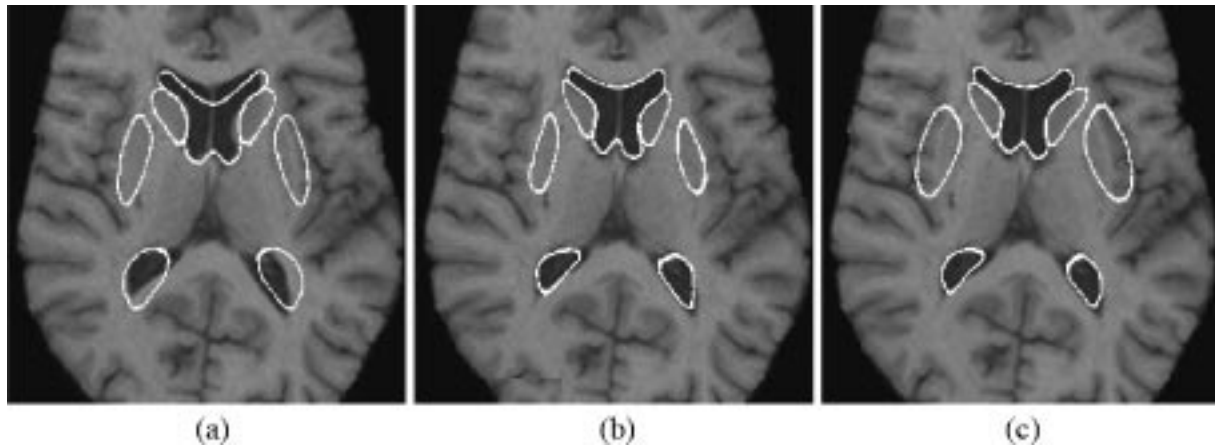


Fig. 4. Demonstration of the adaptive-focus deformation strategy. Cross sections of a 3-D model are shown overlaid on the MR images. (a) The model is automatically initialized, based on the center of mass of the brain. (b) An intermediate stage using the adaptive-focus strategy, in which the ventricles at first had a much stronger influence on the model's deformation, while the other structures followed the ventricular deformation. (c) An intermediate state obtained via the nonadaptive strategy, with all structures deforming simultaneously (nonadaptive scheme). In (b), the LN followed the contraction of ventricular component of the model, thereby avoiding the adjacent cortical edges. In (c), the LN has been trapped by the cortical edges.

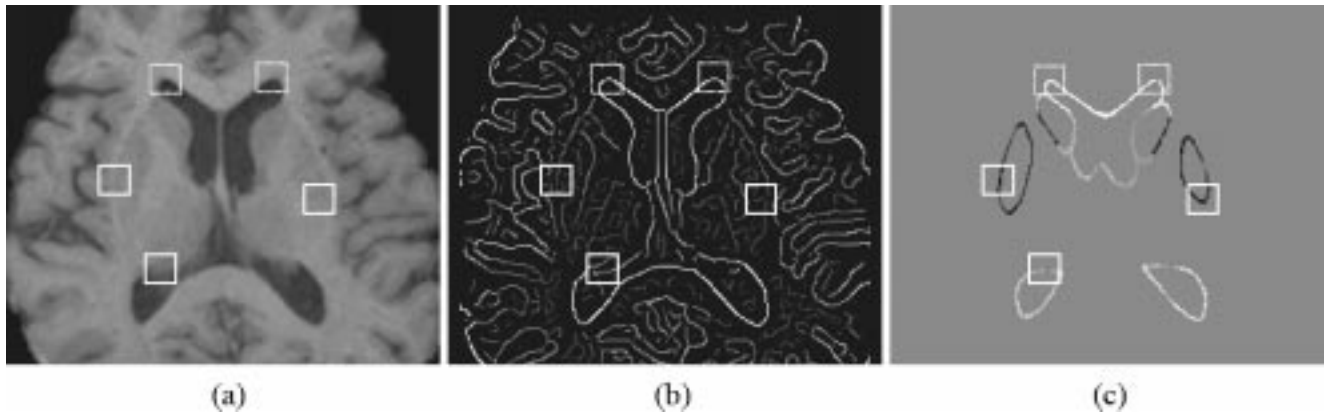


Fig. 5. Level of confidence obtained from one sample. (a) Cross-section of a brain image. (b) Its edge map. (c) The confidence level, where white denotes the highest confidence, black denotes the lowest confidence, and uniform gray denotes background.

F. Learning Adaptive-Focus Knowledge

The design of adaptive-focus strategy depends on our knowledge (or the level of our confidence) on each component of the object of interest. In this section, a learning algorithm is chosen to extract adaptive-focus knowledge from a set of aligned samples. Usually, we have higher confidence level for the model vertices that are located on the strong and isolated edges. Particularly, for a certain model vertex, the level of our confidence can be defined as the difference between its edge strength and the mean edge strength in its neighborhood. We detail it in the following.

In Section II-C, we use $|\nabla I(V_j^i)|$ as the magnitude of the gradient at vertex V_j^i . Here, we also use $|\nabla I(V_j^i)|$ as the edge strength at vertex V_j^i . In the neighborhood of vertex V_j^i , i.e., $11 \times 11 \times 11$, we can calculate the mean edge strength, which is denoted as $mean(V_j^i)$. The difference between $|\nabla I(V_j^i)|$ and $mean(V_j^i)$ is defined as the level of our confidence on vertex V_j^i , $conf(V_j^i) = |\nabla I(V_j^i)| - mean(V_j^i)$. The confidence level $conf(V_j^i)$ could be negative. It represents the case of a weaker edge at vertex V_j^i and lots of stronger confusing edges in the neighborhood. By finding the minimal and the maximal confi-

dence levels from all model vertices, we can normalize all confidence levels to a domain of $[0, 1.0]$.

We applied the above algorithm to the model with the boundaries of the ventricles, CN and LN. In the following, we give two examples where the confidence levels on the model vertices are calculated from one sample (Fig. 5) and 16 aligned samples (Fig. 6), respectively. For visualization purposes, the level of confidence is represented by the intensities of the model vertices, where white corresponds to confidence level 1.0 and black corresponds to confidence level 0. In Fig. 5, we provide a cross section of a brain image [Fig. 5(a)], its edge map [Fig. 5(b)], and its confidence level map [Fig. 5(c)]. For clearness, the background in Fig. 5(c) has been set to gray, level 0.5. It is shown that the ventricular corners have relatively higher confidence levels, while partial CN and most LN have very lower confidence levels. Fig. 6 gives the confidence levels calculated from the 16 aligned samples.

III. ADAPTIVE-FOCUS DEFORMABLE STATISTICAL SHAPE MODEL

In this section, we extend AFDM to incorporate information about the statistical variation of the model, in a form analogous

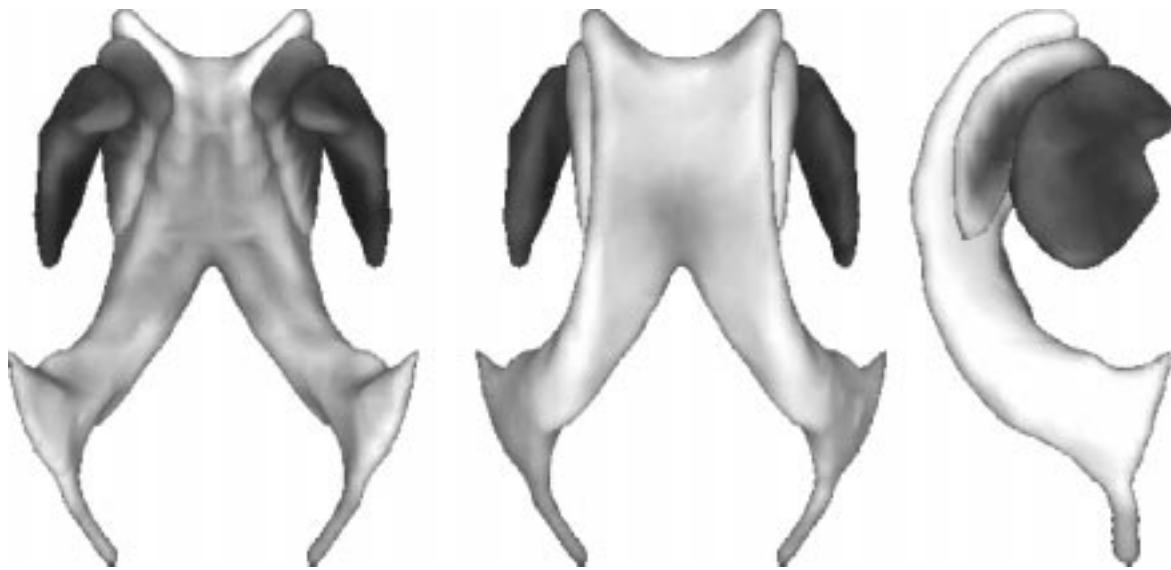


Fig. 6. The levels of confidence, obtained from the 16 aligned samples and displayed as intensities of the model vertices. White denotes confidence level 1.0 while black represents confidence level zero. The model is displayed at three different views, top, down, and side.

to the work in [6], [8]–[10]. That is, the output of AFDM at each iteration is constrained by the currently focused statistical information. In the following, we first describe how we construct models of the training samples using AFDM, while simultaneously establishing point-correspondences, via the attribute vector similarity criteria embedded in AFDM. We then extend the statistical shape modeling paradigm of [6], [8]–[10] to the adaptive focus framework of AFDM. The resulting model is called adaptive focus deformable statistical shape model (AFDSM).

A. Training Set Construction

Statistical shape models have gained popularity in the medical image analysis community after they were first introduced in [6], because they elegantly incorporate prior knowledge about the expected shape and variation of a structure of interest. One of the difficulties, however, associated with these models is their training, which depends on defining point correspondences in a training sample. This task is fairly straightforward in two dimensions, although efforts to automate landmark definition have also been made [16]. However, definition of point correspondences in three dimensions is a very difficult task. To overcome this difficulty, some investigators have assumed that approximate correspondences can be defined by placing parametric grids on the structure of interest [8], [22]. Although this is a convenient way to define correspondences and train a statistical shape model, it is based on only a rough approximation of point correspondences.

In our work we train the deformable model on samples whose point correspondences are defined via AFDM. In particular, images of each training sample are first hand-segmented on a section-by-section basis to the structures of interest (ventricles, LN, and CN in this paper). After hand-segmentation, each interesting class has its own intensity. AFDM is then applied to the hand-labeled images, resulting in a surface representation

of each boundary. Notably, since AFDM is based on a similarity between attribute vectors, it determines point-correspondences while deforming, and does not merely rely on approximate point correspondences defined via parametric grids. We found that AFDM worked very well on these hand-labeled images. In very few cases we had to manually help the algorithm by “pulling” the surface to the boundary. The resulting point-correspondences were used to calculate the covariance matrix and, hence, build the principal components of shape variation, as in [6], [9].

Figs. 10 and 11 show the deformation of the model (initially obtained from a single subject) to a hand-labeled training sample. Cross sections of the deformed model are shown in black and are overlaid on (gray) sections of the target boundary. Three stages of the process are shown: the initialization, an intermediate result obtained after the model has focused primarily on the ventricles, and the final result.

B. Adaptive-Focus Statistical Information

In previous statistical shape models [6], [8]–[10], all landmarks in the training samples was given equal weights when calculating shape statistical parameters. However, in this equal weighting scheme, larger features of a shape dominate over relatively smaller, yet important features, merely because their large size influences the measures of shape variability. Furthermore, relatively unreliable features, if they are large, will dominate over relatively more reliable and important features. To overcome this limitation, in our calculation of the statistical parameters, we weight different vertices of the model differently, with vertices belonging to relatively smaller structures assigned relatively higher weights and vice versa. In order to better explain the importance of variable weighting, we will use the example of a model containing a large structure (ventricles) and a smaller structure (LN). Due to their large size, the ventricles dominantly affect the statistical shape parameters, in that

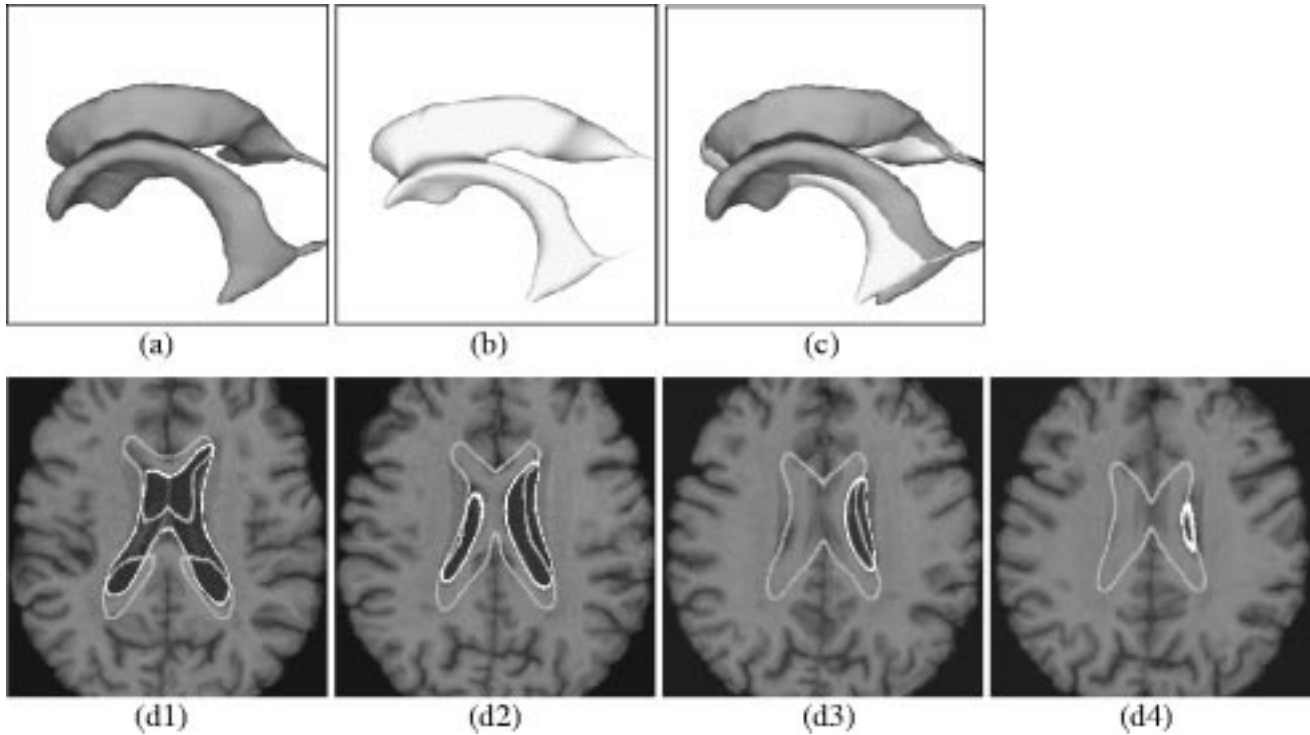


Fig. 7. An example of ventricular segmentation using AFDM. (a) Model (gray) and (b) Segmented ventricles (white). (c) Overlay of segmented ventricles and initial model, showing the deformation imposed by AFDM. (d1–d4) Cross sections of initial (gray) and final (white) model configuration. Initialization of the model was performed automatically.

the dominant eigenvectors primarily reflect the variability of the ventricles. Accordingly, a deformation of the ventricles by image-derived forces induces very little deformation on the LN. This is problematic, since the LN should follow the deformation of the ventricles. Moreover, depending on the number of eigenvectors used, fine details of the LN can be lost.

Our statistical model is analogous to the one in [6] and [8]–[10]. However, the variable weighting of the components of the model effectively zooms each component to the same overall size in the space in which the statistics are calculated, so that each component is represented in the most important eigenvectors of the corresponding covariance matrix. Each component is then scaled back appropriately to its actual size. Fig. 12 provides some examples of training samples after they have been linearly transformed to a template. More details of the algebraic manipulations involved in these transformations can be found in [21]. In addition to accounting for size differences in substructures of an active shape model, AFDSM allows for emphasis to be placed on structures or features that are relatively more reliable (see Figs. 4–6 and Section II-E).

IV. EXPERIMENTS

In the experiments of this section we test the performances of AFDM and AFDSM, first when the model consists of a single surface (the ventricular boundary) and then when multiple surfaces are included. Furthermore, we also test the robustness to noise. For the all experiments, the surface segment lengths $R(V_j^i)$ are initially set to 30, δ_l set to one and $\omega_{i,j}$ set to their own levels of confidence. The computation time depends on the

complexity of the studied subject. It is about 30 min in an SGI OCTANE workstation.

A. Segmentation of Ventricular Boundaries by a Single-Component Model

A ventricular model is shown in Fig. 7(a). This model uses 2646 vertices and 5269 triangles; the initial model was constructed from a single individual using the Matlab isosurface routine. We tested this model on 32 typical brain images, and were able to accurately reconstruct the ventricular surface in 31 out of 32 cases by AFDM. Fig. 7 shows one of the correct results. In Fig. 7(a) and (b), initial model and segmentation result are shown as grey and white surfaces, respectively. They are overlaid in Fig. 7(c). Fig. 7 (d1–d4) shows four different cross sections of the same result, with the initial models (grey contours) and final segmentations (white contours). The final segmentation result accurately fits the expected ventricular boundaries.

To show the importance of using statistical information in the deformable models, we present one case for which AFDM failed (see Fig. 8). The reasons for which AFDM failed in this case is that the local search applied in AFDM, in conjunction with a rather poor initialization and the notable difference between the model and the target, prevented the model from reaching far out to the boundary, especially in the posterior horns of the ventricles. In Fig. 9 we demonstrate a good result on the same case, using AFDSM. The robustness provided by the statistical information allows the algorithm to increase its searching window and avoid local minima, yielding a result far better than AFDM or the statistical model alone.

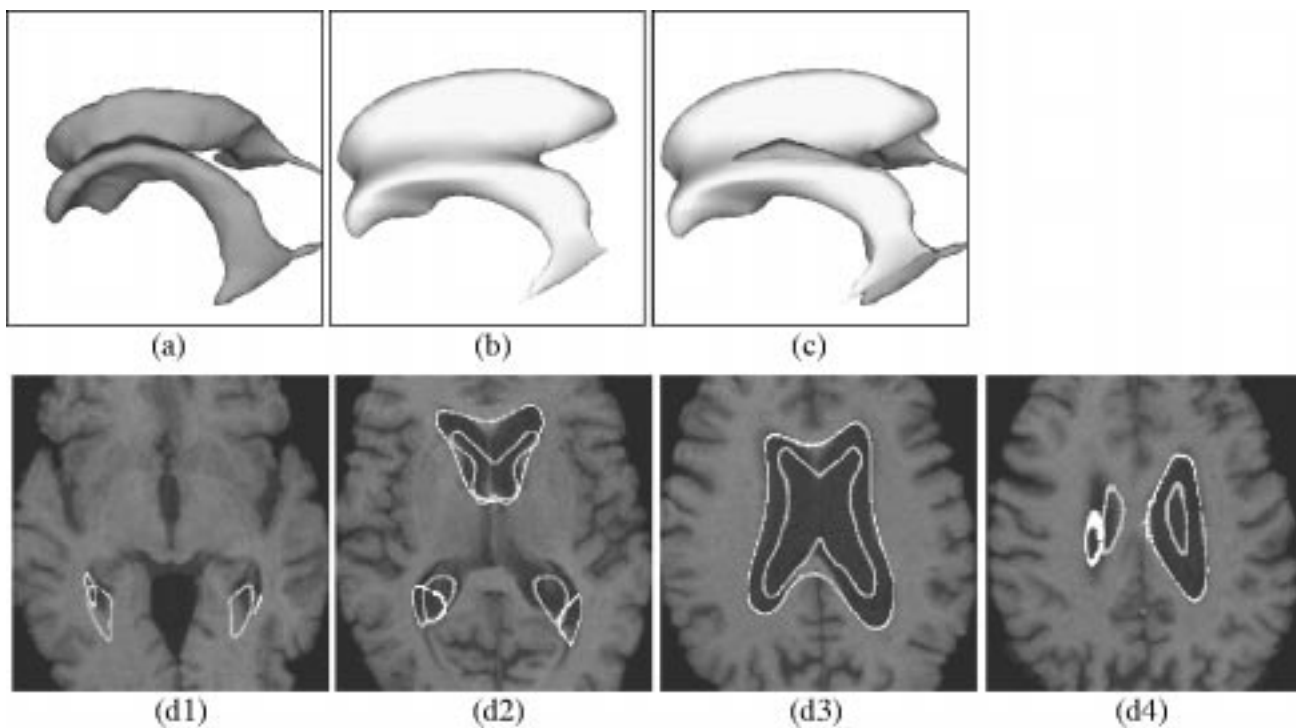


Fig. 8. A bad result obtained with AFDM.

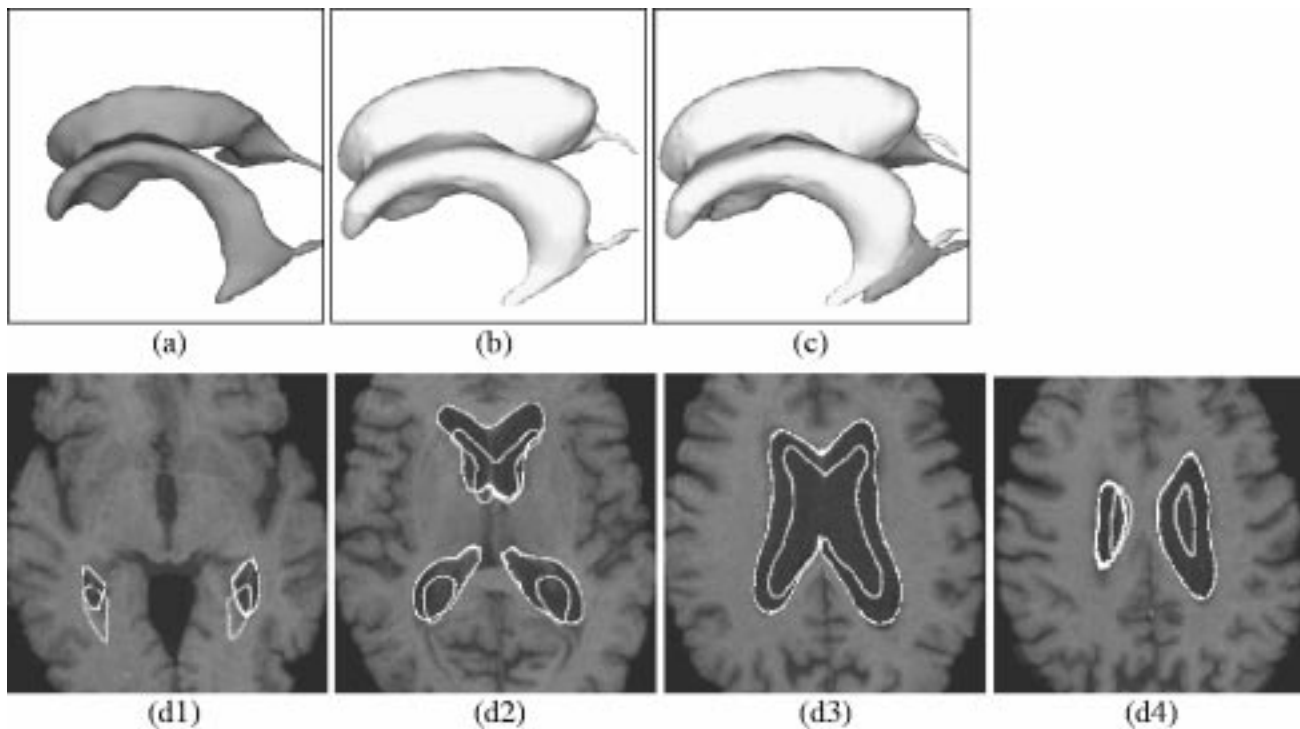


Fig. 9. Result obtained from the example of Fig. 8, using AFDSM. Gray is the (automatic) initialization and white is the result.

B. Segmenting Multiple Structures of the Human Brain with a Multicomponent Model

Fig. 1(b) shows a five-component model containing the boundaries of the ventricles and the left and right CN and LN. The total number of vertices in this model is 3966. The surfaces of the ventricles, CN and LN have 2399, 760, and 807 vertices,

respectively. The total number of triangles in the whole model is 7912. Using the technique in Section II-A, connections among proximal vertices of different components were formed, as in Fig. 1(a).

As we described in Section III-A, AFDM was applied on hand-labeled images in order to construct the training set. An example of this procedure is shown in Figs. 10 and 11, where

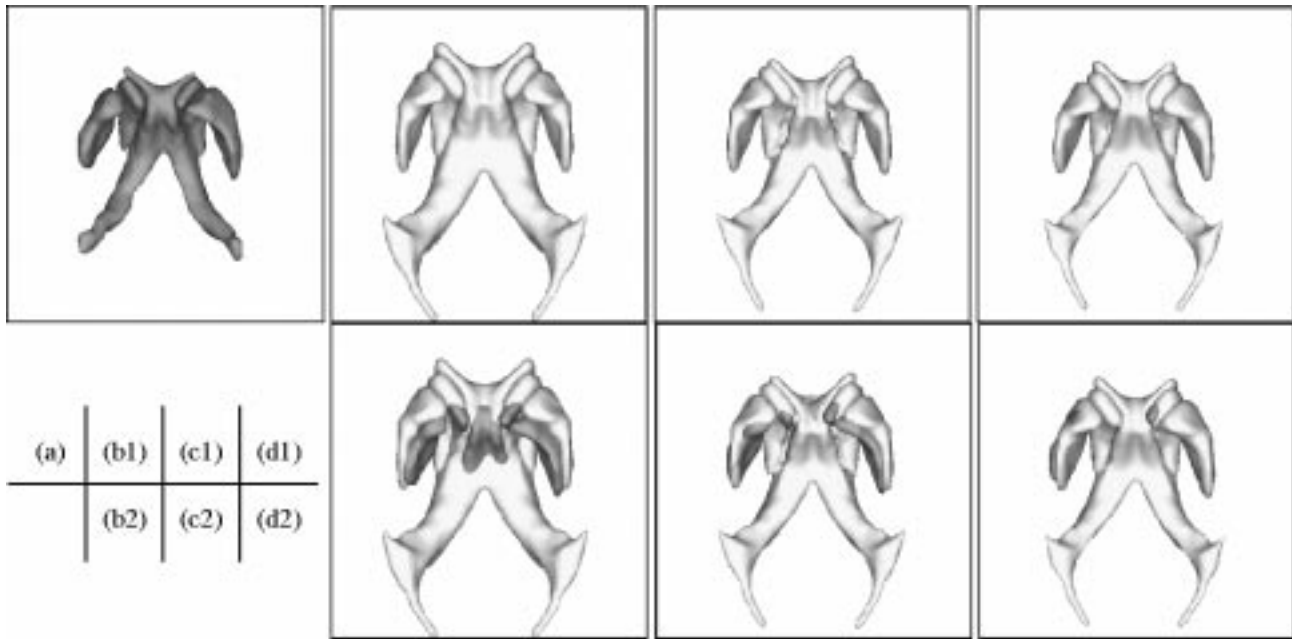


Fig. 10. Deformation of the model to a hand-labeled target image of a training sample, for determining point correspondences. (a) Binary sample data, (b1) model, and (b2) model with the target. (c1) Intermediate result after the phase of focusing on the ventricles. (c2) Intermediate result merged with the target. (d1) Final result. (d2) Final result merged with the target.

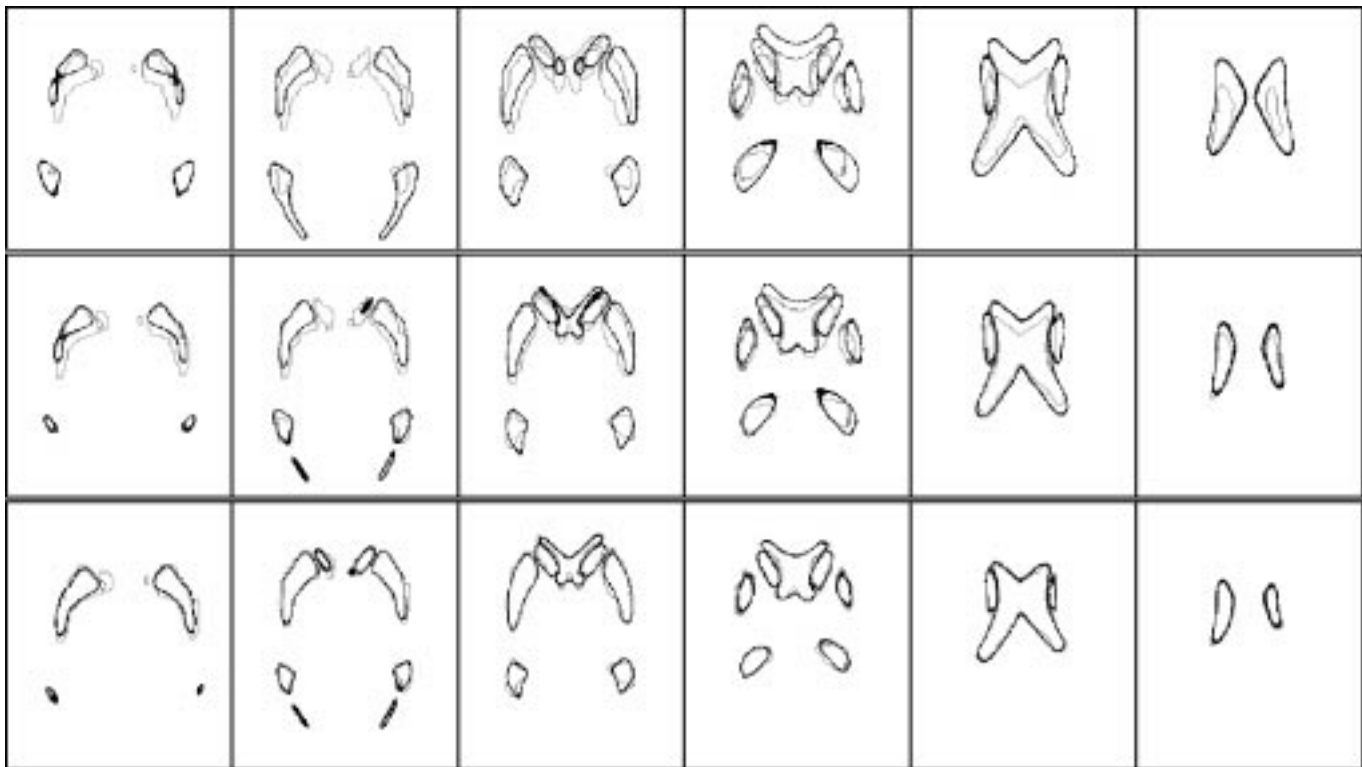


Fig. 11. Cross sections of the example of Fig. 10. Target outlines are shown as gray and deformed model as black. Top row shows initialization and middle row shows an intermediate result during the phase of focusing on the ventricles. The last row shows the final result.

the deformable model is shown as a black curve, and the outline of the hand-labeled image is shown in gray lines. The initial, and intermediate, and the final results are shown in the top, middle, and bottom rows, respectively, in Fig. 11. The intermediate result was obtained primarily by focusing on the ventricles, while the rest of the components of the model were deformed

via a global linear transformation following the ventricular deformation. Some of the resulting training samples are displayed in Fig. 12, which are linearly transformed versions of the AFDM results on the hand-labeled images.

One representative result obtained via AFDSM is shown in Fig. 13, in three dimensions and as cross sections.

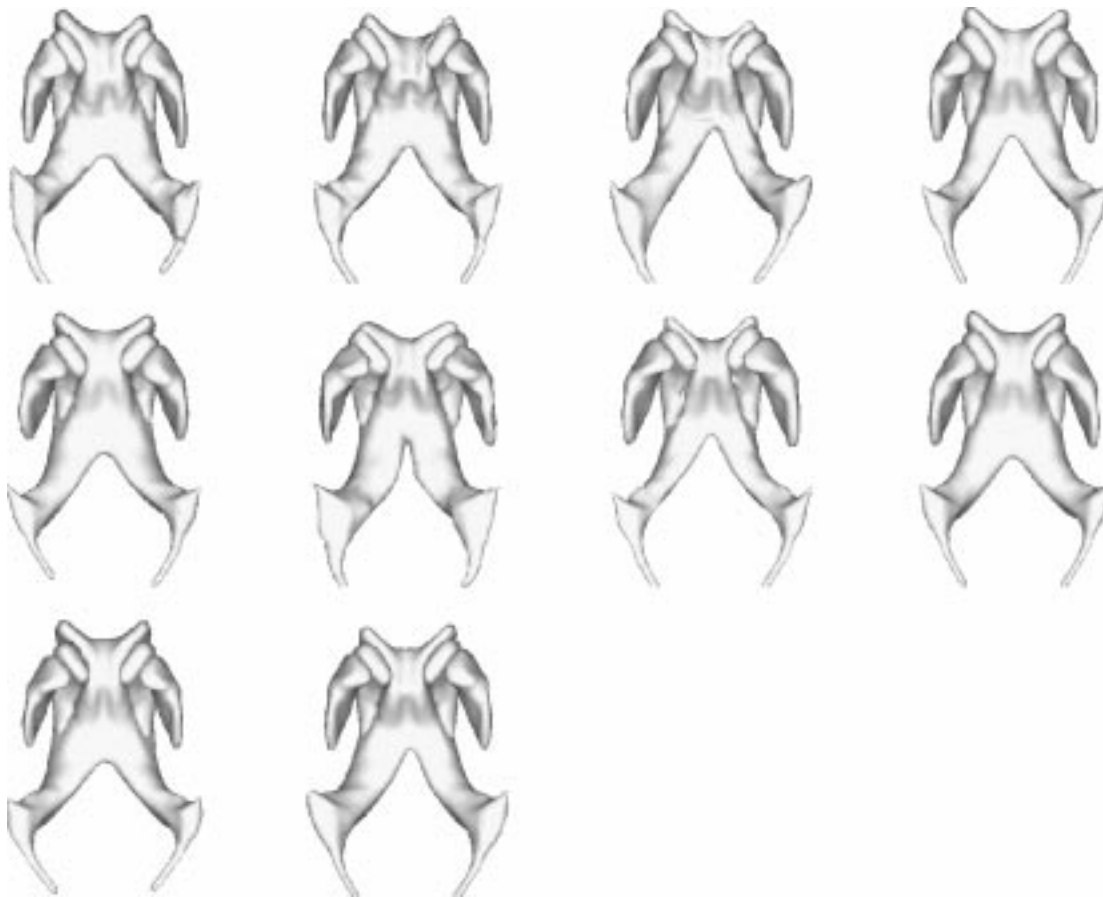


Fig. 12. Some samples used in the multicomponent model. The training set contained initially 12 subjects, and was subsequently extended using correctly segmented images. All samples have been normalized to the same space of the model, via a linear transformation.

In order to validate the algorithm on the ten training samples for which we had available manual segmentations, we applied AFDM and AFDSM on each subject's original MR images, and we compared the result to the manual segmentations. We calculated the average distance between the resulting surfaces and the corresponding hand-labeled boundaries, and we found it to be comparable to the voxel size. Table I shows the results for the two algorithms. The results of Table I show that AFDSM usually has slightly better performance than AFDM, because of additional the use of statistical information. But the improvement is only marginal. In one case (sample 10), the error actually increases when using statistical information, probably due to the relatively small size of the training set.

C. Robustness of AFDSM to Noise

In order to demonstrate the effect of noise on the performance of AFDSM, we generated a 3-D synthetic study image from a hand-labeled image shown in Fig. 10(a). One cross section of the synthetic study image is shown in Fig. 14(a). In order to make this simulation realistic, within each structure we used the average image intensity calculated from the training set. Accordingly, the image intensity in the ventricles was set to the average CSF intensity, 35, while intensities in the left and right CN and LN were set to the average gray-matter intensity, 82. Others are set to average white-matter intensity. We added varying amounts of zero-mean Gaussian noise to this synthetic study image, to

obtain noisy images. Fig. 14(b-f) shows five cross sections of five noisy images that were corrupted by Gaussian noise with standard deviation of 10, 20, 30, 40, and 50, respectively. In our algorithm, the initialization for all these testing images are the same as that used in Figs. 10 and 11. The average boundary error [29], $E_{b,a}$, is defined as the average distance between the actual boundaries of the study image [Fig. 14(a)] and the boundaries that are segmented by our algorithm. Solid lines in Fig. 15 show the variations of the average boundary error, $E_{b,a}$, with standard deviations of zero-mean Gaussian noise. Moreover, the average correspondence error, $E_{c,a}$, is defined and calculated to evaluate the further performance of our algorithm. Here, we define $E_{c,a}$ as the average distance between corresponding points obtained under noise, relative to their counterparts in the absence of noise. The dotted lines in Fig. 15 plot $E_{c,a}$ for different noise levels.

The performance of our algorithm becomes worse with the increase of noise levels. The average boundary error ranges from 0.9 to 1.1 pixels, while the average correspondence error is from 0.15 to 0.78. Overall, the proposed model is robust to noise.

V. CONCLUSION

In this paper, we have presented a deformable model (AFDSM) for automatically segmenting objects from volumetric MR images, as well as for establishing point correspondences, by combining geometric and statistical information.

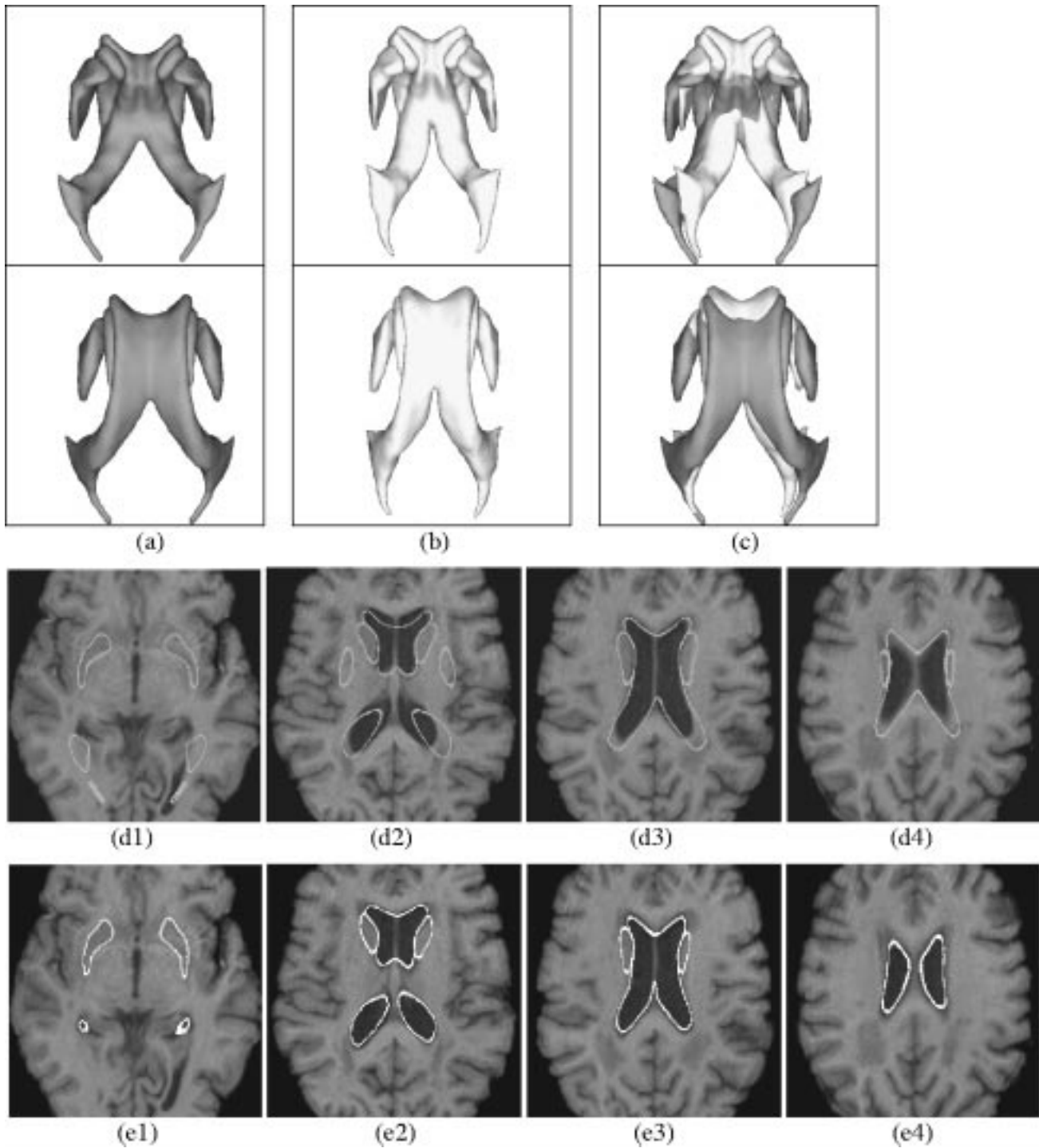


Fig. 13. One example on segmenting multiple structures using AFDSM. (a) Model surface, displayed at two different views and (b) segmentation result. (c) Overlay of the segmentation result and the model, revealing the deformation that the model underwent. (d1–d4) Initial position of the model in the four different slice images. (e1–e4) The final segmentation results corresponding to (d1–d4).

TABLE I

QUANTITATIVE VALIDATION ON BOTH AFDM AND AFDSM ALGORITHMS BY USING TEN HAND-LABELED SAMPLES. THE ERRORS ARE AVERAGE BOUNDARY DISTANCES, AT THE UNIT OF PIXEL

samples	1	2	3	4	5	6	7	8	9	10
AFDM	1.14	1.45	1.04	1.77	1.51	0.97	1.52	1.53	1.66	1.07
AFDSM	0.99	1.42	0.90	1.46	1.35	0.91	1.37	1.32	1.59	1.12

The geometric information is in the form of an attribute vector attached to each point of a deformable model, which carries

information about the geometric structure of the model from a local to global scale. Attribute vectors help differentiate among

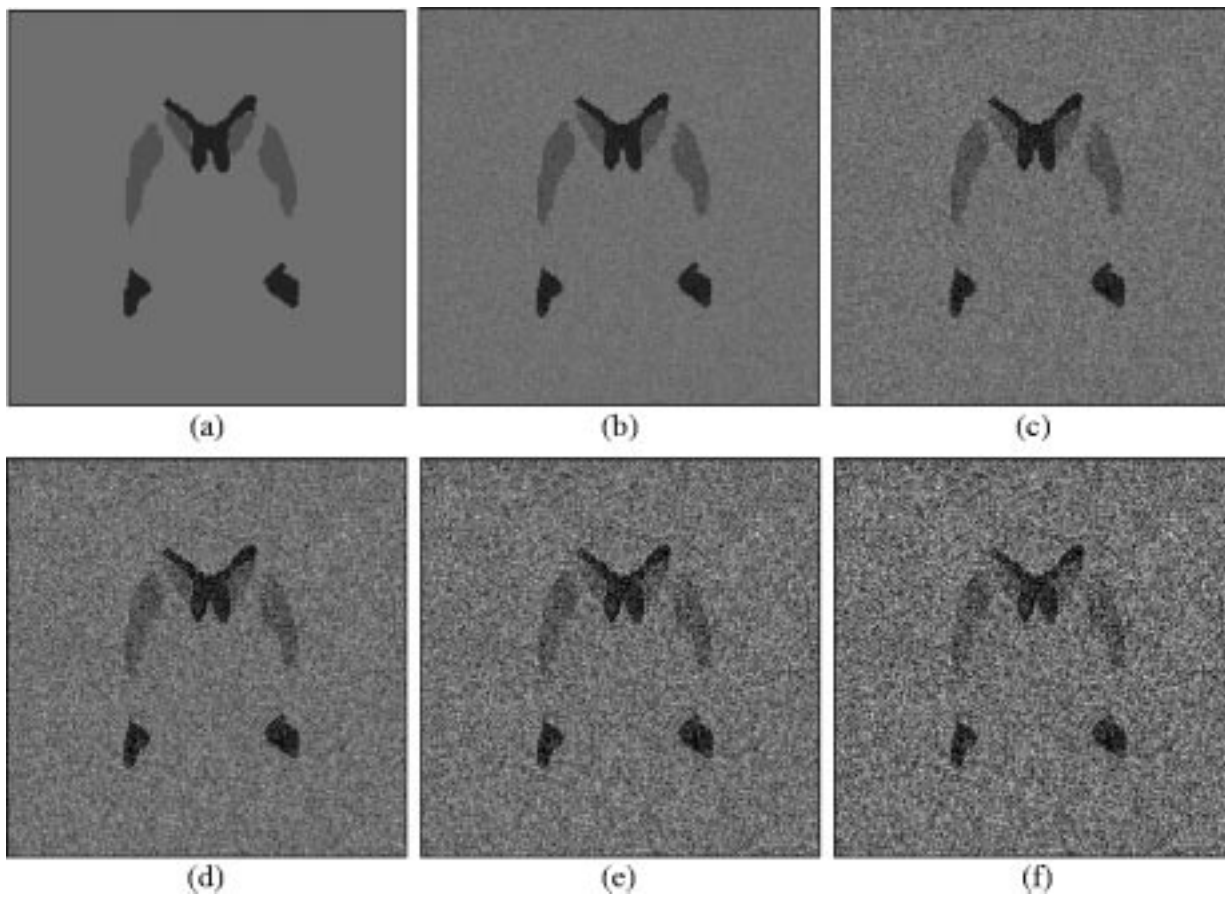


Fig. 14. Images used for demonstrating the robustness of AFDSM against noise. (a) Synthetic study image. (b–f) Noisy images that are corrupted by zero-mean Gaussian noise with standard deviation of 10, 20, 30, 40, and 50, respectively.

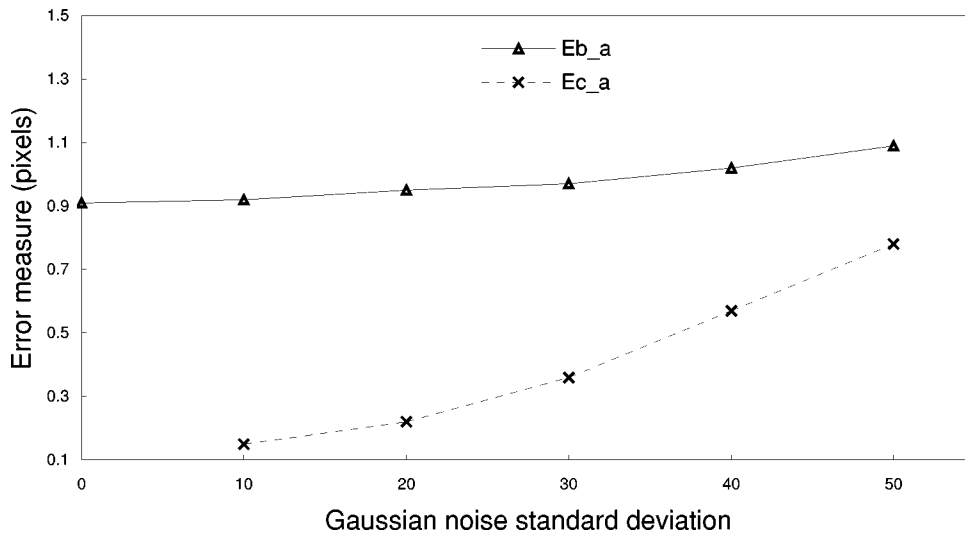


Fig. 15. The effect of noise on the performance of AFDSM. E_{b_a} : average boundary error; E_{c_a} : average correspondence error. Horizontal axis represents the amounts of standard deviation of zero-mean Gaussian noise. Vertical axis denotes the errors at a unit of pixel.

different points along the boundaries of anatomical structures. The statistical information in AFDSM reflects the expected shape variability, while accounting for size differences in the different components constituting the model. A hierarchical deformation mechanism was proposed, which initially focuses on boundaries that are easier to find, and gradually shifts focus to other structures, as those get closer to their respective targets.

A key feature of our model is that it is able to define point correspondences, in addition to segmenting a structure of interest. This is achieved via an energy term that uses similarity of the attribute vectors instead of some arbitrary elastic or other internal energy form, and is based on the fact that attribute vectors characterize the local and global geometric structure of the model around each of its points. We also demonstrated that the adap-

tive focus deformation mechanism helps the model avoid bad solutions, which perhaps correspond to local minima.

Our model has been applied to segmenting the boundaries of the ventricles, the CN, and the LN from volumetric MR images. Current work focuses on extending this model to include more structures, particularly the hippocampus and adjacent cortical structures.

Some extensions of our methodology are possible. In particular, instead of or in conjunction with using the Canny edges to determine external forces, a fuzzy segmentation [23] can be first applied. This approach will render the technique independent of the particular image acquisition protocol.

Currently, the vertices are arranged into tetrahedra that represent the 3-D structure of objects. In this way, the spatially close vertices are possibly distant on the surface. Medial representation [25] can fix this unexpected case. Therefore, the combination of our geometric representation with the medial representation should improve the results, since these two methods have complementary merits and weaknesses. Finally, the creation of our model at multiple resolutions will definitely accelerate the speed of our algorithm.

We finally want to note that all of our experiments have been performed on MR images of elderly individuals, which display reduced white matter/grey matter contrast and often extreme atrophy, which is reflected, in part, by very large ventricles. Despite the difficulties imposed by the nature of the data, we have obtained good and robust results in approximately 90 brain images.

The source code for our deformable model (AFDSM) is freely available from <http://pandora.cbm.v.jhu.edu/~dgshen/3DSnakeCode.htm>.

REFERENCES

- [1] A. K. Jain, Y. Zhong, and S. Lakshamanan, "Object matching using deformable templates," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 3, pp. 267–278, Mar. 1996.
- [2] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1988.
- [3] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Med. Image Anal.*, vol. 1, no. 2, pp. 91–108, 1996.
- [4] K. F. Lai and R. T. Chin, "Deformable contour: Modeling and extraction," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 1084–1090, Nov. 1995.
- [5] H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, pp. 1061–1075, Nov. 1992.
- [6] T. F. Cootes, D. Cooper, C. J. Taylor, and J. Graham, "Active shape models—Their training and application," *Comput. Vis. Image Understanding*, vol. 61, no. 1, pp. 38–59, Jan. 1995.
- [7] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," in *Proc. 5th Eur. Conf. Computer Vision*, H. Burkhardt and B. Neumann, Eds., 1998, vol. 2, pp. 484–489.
- [8] A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-D neuroradiological data sets," *IEEE Trans. Med. Imag.*, vol. 18, pp. 828–839, Oct. 1999.
- [9] G. Szekely, A. Kelemen, C. Brechbuhler, and G. Gerig, "Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models," *Med. Image Anal.*, vol. 1, no. 1, pp. 19–34, 1996.
- [10] Y. Wang and L. H. Staib, "Boundary finding with prior shape and smoothness models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 738–743, July 2000.
- [11] C. Kervrann and F. Heitz, "A hierarchical Markov modeling approach for the segmentation and tracking of deformable shapes," *Graphical Models Image Processing*, vol. 60, no. 3, pp. 173–195, May 1998.
- [12] H. Lester and S. R. Arridge, "A survey of hierarchical nonlinear medical image registration," *Pattern Recogn.*, vol. 1, no. 32, pp. 129–149, 1999.
- [13] E. E. Milios, "Shape matching using curvature process," *Comput. Vis. Graph. Image Processing*, vol. 47, pp. 203–226, 1989.
- [14] M. Baroni and G. Barletta, "Digital curvature estimation for left ventricular shape analysis," *Image Vis. Computing*, vol. 10, no. 7, pp. 485–494, Sept. 1992.
- [15] F. Mokhtarian, "Silhouette-Based isolated object recognition through curvature-scale space," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 539–544, May 1995.
- [16] A. Hill, C. J. Taylor, and A. D. Brett, "A framework for automatic landmark identification using a new method of nongrid correspondence," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 241–251, Mar. 2000.
- [17] H. H. S. Ip and D. Shen, "An affine-invariant active contour model (AI-snake) for model-based segmentation," *Image Vis. Computing*, vol. 16, no. 2, pp. 135–146, 1998.
- [18] D. Shen, W. H. Wong, and H. H. S. Ip, "Affine invariant image retrieval by correspondence matching of shapes," *Image Vis. Computing*, vol. 17, no. 7, pp. 489–499, May 1999.
- [19] D. Shen, H. H. S. Ip, and E. K. Teoh, "Robust detection of skewed symmetries by combining local and semi-local affine invariants," *Pattern Recogn.*, vol. 14, no. 7, pp. 1417–1428, Feb. 2001.
- [20] —, "Affine invariant detection of perceptually parallel 3-D planar curves," *Pattern Recogn.*, vol. 33, no. 11, pp. 1909–1918, Nov. 2000.
- [21] D. Shen and C. Davatzikos, "A adaptive-focus deformable model using statistical and geometric information," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 22, pp. 906–913, Aug. 2000.
- [22] C. Davatzikos, "Spatial transformation and registration of brain images using elastically deformable models," *Comput. Vis. Image Understanding*, vol. 66, no. 2, pp. 207–222, May 1997.
- [23] D. L. Pham and J. L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," *IEEE Trans. Med. Imag.*, vol. 18, pp. 737–752, Sept. 1999.
- [24] T. McInerney and D. Terzopoulos, "Topology adaptive deformable surfaces for medical image volume segmentation," *IEEE Trans. Med. Imag.*, vol. 18, pp. 840–850, Oct. 1999.
- [25] S. M. Pizer, D. S. Fritsch, P. A. Yushkevich, V. E. Johnson, and E. L. Chaney, "Segmentation, registration, and measurement of shape variation via image object shape," *IEEE Trans. Med. Imag.*, vol. 18, pp. 851–865, Oct. 1999.
- [26] H. D. Tagare, "Shape-based nonrigid correspondence with application to heart motion analysis," *IEEE Trans. Med. Imag.*, vol. 18, pp. 570–579, July 1999.
- [27] M. Chen, T. Kanade, D. Pomerleau, and J. Schneider, "3-D deformable registration of medical images using a statistical atlas," in *Proc. MICCAI*, Sept. 1999, pp. 621–630.
- [28] N. Duta and M. Sonka, "Segmentation and interpretation of MR brain images: An improved active shape model," *IEEE Trans. Med. Imag.*, vol. 17, pp. 1049–1062, Dec. 1998.
- [29] Y. Wang and L. H. Staib, "Physical model-based nonrigid registration incorporating statistical shape information," *Int. J. Med. Image Anal.*, vol. 4, no. 1, pp. 7–20, 2000.
- [30] H. Delingette, "General object reconstruction based on simplex meshes," *Int. J. Comput. Vis.*, vol. 32, no. 2, pp. 111–146, Sept. 1999.
- [31] D. S. Fritsch, S. M. Pizer, L. Yu, V. Johnson, and E. L. Chaney, "Localization and segmentation of medical image objects using deformable shape loci," in *Lecture Notes in Computer Science*. Berlin, Germany: Springer, 1997, vol. 1230, Proceedings of Information Processing in Medical Imaging 1997 (IPMI '97), pp. 127–140.
- [32] J. Montagnat and H. Delingette, "Globally constrained deformable models for 3-D object reconstruction," *Signal Processing*, vol. 71, no. 2, pp. 173–186, Dec. 1998.