

Adaptive-Focus Statistical Shape Model for Segmentation of 3D MR Structures

Dinggang Shen¹ and Christos Davatzikos^{1,2,3}

¹ Department of Radiology, ² Department of Computer Science,
³ Center for Computer-Integrated Surgical Systems and Technology,
Johns Hopkins University
Email: dgshen@cbmv.jhu.edu, hristos@rad.jhu.edu

Abstract. This paper presents a deformable model for automatically segmenting objects from volumetric MR images and obtaining point correspondences, using geometric and statistical information in a hierarchical scheme. Geometric information is embedded into the model via an affine-invariant attribute vector, which characterizes the geometric structure around each model point from a local to a global level. Accordingly, the model deforms seeking boundary points with similar attribute vectors. This is in contrast to most deformable surface models, which adapt to nearby edges without considering the geometric structure. The proposed model is adaptive in that it initially focuses on the most reliable structures of interest, and subsequently switches 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.

1 Introduction

Deformable models have been extensively used as segmentation tools in medical imaging applications. An excellent review of deformable models can be found in [1].

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 towards noisy or fragmented edges. Moreover, it will make the deformable model very sensitive to initialization. To remedy this, prior knowledge in the form of normal statistical variation can be incorporated into the formulation of a deformable model [2], in order to constrain the possible deformations. Chen and Kanada [11] 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* [3] have developed an active shape model (ASM), based on the statistics of labeled samples. An improvement of ASM has been presented in [12]. Following the seminal work of Cootes *et al*, a flexible Fourier surface model [4] was proposed based on a hierarchical parametric object description rather than a point distribution model. Finally, some researchers consider capturing statistical information from the preprocessed covariance matrix [5].

The aforementioned statistical models can also be implemented hierarchically, since hierarchical implementation usually increases the likelihood of finding the globally optimal match. A review of the hierarchical strategies can be found in [6].

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 methodology has three novel aspects, which are briefly described next.

First, an attribute vector is attached to each model point, and it is used to characterize the geometric structure of the model around that point, from a local to a global scale. Our model only deforms the surface segments to the image boundaries with similar geometric structure, based on an energy term that favors similar attribute vectors. The attribute vectors are essential in our formulation, since they distinguish different parts of a boundary according to their shape properties, and therefore they guide the establishment of point-correspondences between the model and an individual anatomy. This is in contrast to most deformable surface models.

The second contribution of our model is in its adaptive formulation. In particular, it can shift focus from one structure to another. The model adaptivity is accomplished both via its hierarchical deformation strategy and via an adaptive focus statistical shape model. This addresses the limitations of previous statistical models. Our 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 defined for the training samples. Although this is readily done in 2D, it is very difficult in 3D. Consequently, other investigators relied on parametric representations that are not necessarily based on point correspondences [4].

2 Adaptive-Focus Deformable Model (AFDM)

In our approach, we first construct a model that represents the typical anatomy of a number of structures by interconnected surfaces (c.f. Fig 1). This is provided in Section 2.1. In Section 2.2, we attach an attribute vector to each vertex of these surfaces. In Sections 2.3 and 2.4, we describe the mechanism that deforms the model to an individual MR volumetric image, and the corresponding energy function being minimized. Finally, in Section 2.5 we present a hierarchical and adaptive-focus strategy for the model deformation.

2.1 Model Description

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's assume that there are M separate surfaces V^i ($1 \leq i \leq M$) in the model, and that the i -th surface V^i has N_i vertices $V_j^i = [x_j^i \quad y_j^i \quad z_j^i]^T$, where $0 \leq j \leq (N_i - 1)$.

Vertex V_j^i is assumed to have $R(V_j^i)$ neighborhood layers, and its l -th neighborhood layer has $S_l(V_j^i)$ neighboring vertices, $nbr_l(V_j^i) = \{nbr_{l,m}(V_j^i) | 0 \leq m < S_l(V_j^i)\}$. Here, layer l is in the range $1 \leq l \leq R(V_j^i)$. The neighborhood layers are constructed so that no vertex is repeated twice in the neighborhood of another vertex.

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 in the deformation procedure. In particular, if the 3D Euclidean distance between vertices belonging to two different surfaces is below a threshold, the vertices are connected as first-layer neighbors. For example, 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 (c.f. Fig 1a). If two vertices belonging to the same surface are close enough, based on their 3D Euclidean distance, they are also joined with each other. This helps prevent self-intersections of the surfaces.

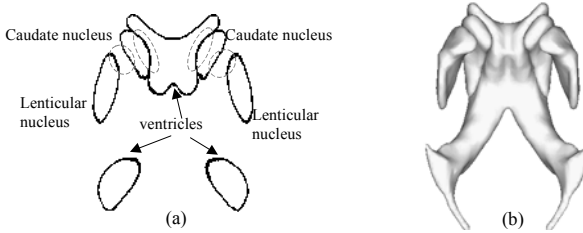


Fig. 1. A model with five surfaces: (a) a cross-section of the 3D model of (b). In (a), any vertex in the dotted ellipses is connected with another vertex of the model.

2.2 Affine-Invariant Attribute Vector

To describe shape characteristics of various scales, an affine-invariant vector of geometric attributes in 2D has been used in [7,8]. Each attribute is the area of a triangle formed by a model point, P_i , and its two neighboring points, P_{i-vs} and P_{i+vs} , in the v_s -th neighborhood layer (c.f. Fig 2a). In the following, we extend the definition of the 2D attribute to the 3D, by using the volume of a tetrahedron (c.f. Fig 2b).

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 . While, the volumes of larger tetrahedrons represent more global properties of the surface around vertex V_j^i . More importantly, even vertices of similar curvatures might have very different attribute vectors, depending on the number of neighborhood layers.

Any four points in the 3D space can establish a tetrahedron. Let's 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 this tetrahedron is:

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

If this tetrahedron is linearly transformed by a 4x4 matrix A , then the volume of the new tetrahedron is $|A| \cdot \text{Volume}(W_0, W_1, W_2, W_3)$. Thus, it is relatively invariant to linear transformation [7].

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 2b). 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 , $F(V_j^i) = [f_1(V_j^i) \ f_2(V_j^i) \ \dots \ f_{R(V_j^i)}(V_j^i)]$, where $R(V_j^i)$ is the number of neighborhood layers around vertex V_j^i . The attribute vector $F(V_j^i)$ captures different levels of shape information around vertex V_j^i .

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

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

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

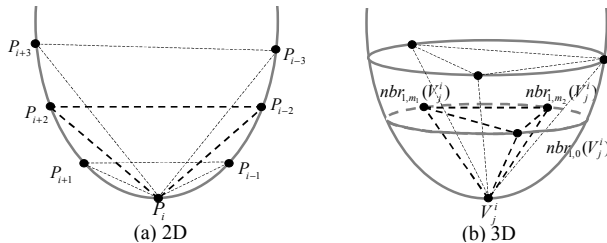


Fig. 2. The attribute vector in 2D and 3D.

2.3 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. Effectively, this defines a segmentation, but also a set of point correspondences based on a similarity between attribute vectors.

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 parameter $\omega_{i,j}$ determines the relative weight for the local energy term $E_{i,j}$. $E_{i,j}$, defined for 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 one of its deformed configurations, around vertex V_j^i . The term $E_{i,j}^{\text{data}}$ defines the external energy.

As we elaborated earlier in this section, the term $E_{i,j}^{\text{model}}$ 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 in the surface segment under consideration. Notice that $R(V_j^i)$ is 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 towards 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 2.4, 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 2.5. The l -th neighborhood layer has $S_l(V_j^i)$ vertices, $nbr_{l,m}(V_j^i)$, where $0 \leq m < S_l(V_j^i)$. The data energy term $E_{i,j}^{\text{data}}$ is designed as follows.

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

2.4 Local Deformation Mechanism

We now describe a greedy deformation algorithm for the minimization of the energy function in equation (1). We suggest considering the deformation of a surface segment as whole, which greatly helps the snake avoid local minima.

Since we deform only one piece of the a model surface at a time, we can introduce discontinuities at the boundary of the segment being deformed. In our 2D model [8], we solved this issue by restricting the local affine transformation so that it leaves the end-points of a deforming segment unchanged. However, for a surface model this is not possible, because the vertices belonging to the $R(V_j^i)$ -th neighborhood layer 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. To remedy this situation, we used a different form of transformation for each deforming surface segment, which is described next.

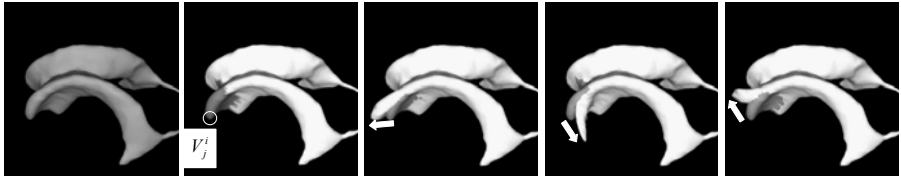


Fig. 3. Demonstration of the proposed deformation mechanism. See text for details.

Let V_j^i be the vertex whose neighborhoods form the surface segment to be deformed at a particular iteration (c.f. 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, $nbr_{i,m}(V_j^i)$, in the segment is defined as $nbr_{i,m}(V_j^i) + \Delta V \cdot \exp(-l^2/2\sigma^2)$, 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 show tentative deformations of the ventricular model (white surfaces) overlaid on the undeformed model (gray surface).

2.5 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

3D model containing these 5 surfaces, and the deformation of that model after 10 iterations. The one in Fig 4b 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 4c was obtained via the same model but with a non-adaptive 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 non-adaptive scheme, however, the LN deformed towards the wrong boundary.

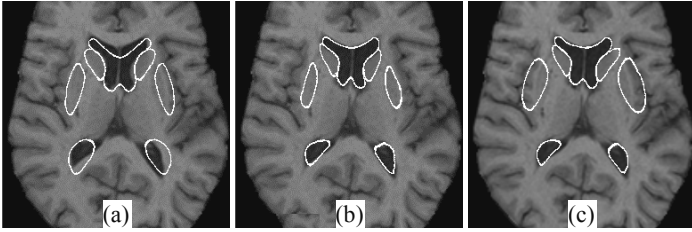


Fig. 4. Demonstration of the adaptive-focus deformation strategy. See text in Section 2.5.

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 1. 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.

3 Adaptive-Focus Deformable Statistical Shape Model (AFDSM)

In this section, we extend AFDM to incorporate information about the statistical variation of the model. We first describe how we construct models of the training samples using AFDM, while simultaneously establishing point-correspondences. We then extend the statistical shape modeling paradigm of [3-5] to the adaptive focus framework of AFDM. The resulting model is called Adaptive Focus Deformable Statistical shape Model (AFDSM).

3.1 Sample Construction

Statistical shape models have gained popularity in the medical image analysis community after they were first introduced in [3]. 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 2D. However, definition of point correspondences in 3D 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 [4,9]. 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, CN, in this paper). AFDM is then applied to the segmented images, resulting in a surface representation of each boundary. Notably, since AFDM is based on a similarity between attribute vectors, it determines point-correspondences based on similarity of the geometric structures of the boundaries of interest. We found that AFDM worked very well on these hand-segmented images. In very few cases we had to manually help the algorithm by “pulling” the surface to the boundary.

Fig 5 shows the deformation of the model (initially obtained from a single subject) to a hand-segmented 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.

3.2 Adaptive-Focus Statistical Information

In the previous statistical shape models [3-5], all landmarks in the training samples was given equal weights when calculating shape statistical parameters. In this way, larger features of a shape dominate over relatively smaller, yet important features, merely because their large size influences the measures of shape variability. Furthermore, unreliable features, if they are large, dominate over relatively reliable and important features. To overcome this limitation, in our calculation of the statistical parameters, we weight different vertices of the model differently. In particular, vertices belonging to relatively smaller structures are assigned relatively higher weights and vice versa.

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 and high variability, the ventricles dominantly affect the statistical shape parameters. In particular, 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.

Our statistical model is analogous to the one in [3-5]. The variable weighting of the components of the model effectively zooms each component to the same overall size in the space where 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. More details of the algebraic manipulations involved in these transformations can be found in [8].

4 Experiments and Conclusion

We test the performance of our algorithm in segmenting multiple structures of the human brain with a multi-component model. Fig 1b shows a 5-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 2.1, connections among proximal vertices of different components were formed, as in Fig 1a.

As described in Section 3.1, AFDM was applied on hand-labeled images in order to construct the training set. An example of this procedure is shown in Fig 5. 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. Fig 6 shows one representative result that was obtained via ADFSM.

We have presented a deformable model for segmenting objects from volumetric MR images. Some extensions of our method are possible. Currently, the vertices are arranged into tetrahedra that represent the 3D structure of objects. An alternative representation could be a medial representation [10]. Combination of our geometric representation with the medial representation should improve the results, since these two methods have complementary merits and weaknesses. Moreover, 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 suffer from reduced white matter/grey matter contrast, and from often extreme atrophy 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 a large set of brain images.

References

1. T. McInerney and D. Terzopoulos: Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1(2): 91-108, 1996.
2. L.H. Staib and J.S. Duncan: Boundary finding with parametrically deformable models. *IEEE Trans. on PAMI*, 14(11):1061-1075, 1992.
3. T.F. Cootes, D. Cooper, C.J. Taylor and J. Graham: Active shape models-their training and application. *Computer Vision and Image Understanding*, 61(1): 38-59, 1995.
4. A. Kelemen, G. Szekely, and G. Gerig: Elastic model-based segmentation of 3-D neuro-radiological data sets. *IEEE Trans. on Medical Imaging*, 18(10): 828-839, 1999.
5. Y. Wang and L.H. Staib: Elastic model based non-rigid registration incorporating statistical shape information. *MICCAI*, pages 1162-1173, Cambridge, MA, October 1998.
6. H. Lester, S.R. Arridge: A survey of hierarchical non-linear medical image registration. *Pattern Recognition*, (32)1, pp.129-149, 1999.
7. Horace H.S. Ip, Dinggang Shen: An affine-invariant active contour model (AI-snake) for model-based segmentation. *Image and Vision Computing* 16(2):135-146, 1998.
8. Dinggang Shen, C. Davatzikos: A adaptive-focus deformable model using statistical and geometric information. To appear in *IEEE Trans. on PAMI*, 2000.

9. C. Davatzikos: Spatial transformation and registration of brain images using elastically deformable models. *Comp. Vis. and Image Understanding*, 66(2):207-222, May 1997.
10. 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. on Medical Imaging*, 18(10): 851-865, October 1999.
11. M. Chen, T. Kanade, D. Pomerleau, J. Schneider: 3-D deformable registration of medical images using a statistical atlas. *MICCAI*, Sept. 1999.
12. N. Duta, M. Sonka: Segmentation and interpretation of MR brain images: An improved active shape model. *IEEE Trans. on Medical Imaging*, 17(6): 1049-1062, 1998

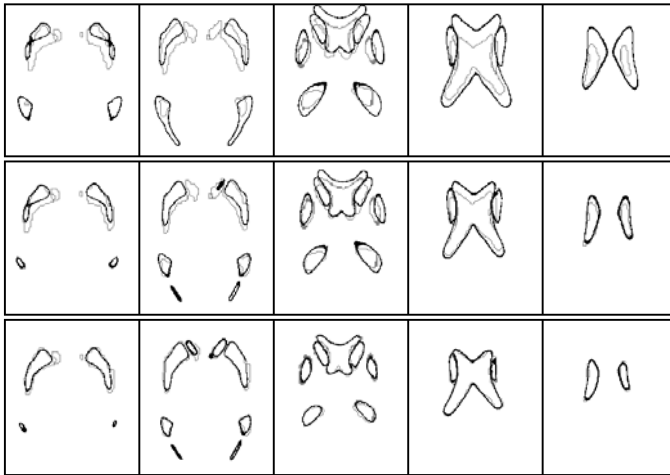


Fig. 5. Deformation of the model to a hand-labeled target image of a training sample, for determining point correspondences. See text in Section 3.1 for details.

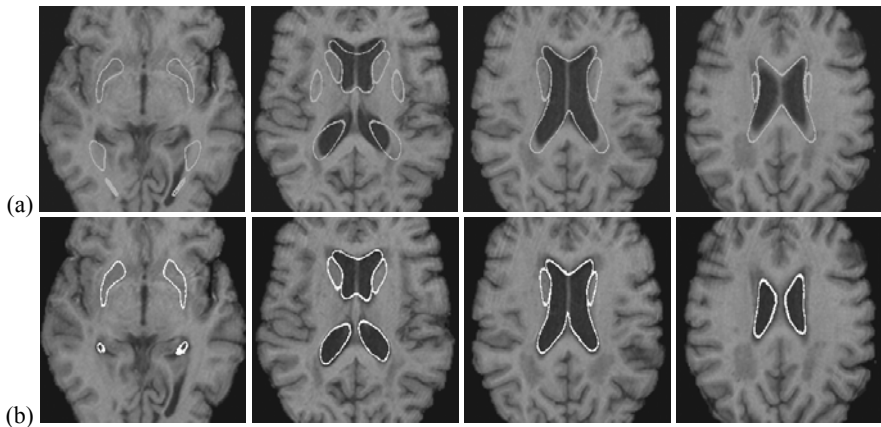


Fig. 6. Example on segmenting multiple structures using AFDSM. (a) Initial position of the model in the four slice images; (b) The final segmentation results corresponding to (a).