STATISTICAL SHAPE MODEL FOR AUTOMATIC SKULL-STRIPPING OF BRAIN IMAGES

Zhiqiang Lao   Dinggang Shen   Christos Davatzikos

Department of Radiology, Johns Hopkins University, Baltimore, Maryland 21287

ABSTRACT

This paper presents a statistical shape model for automatic skull stripping of MR brain images. A surface model of the brain boundary is hierarchically represented by a set of overlapping surface patches, each of which has elastic properties and deformation range that is learned from a training set. The model's deformation is hierarchical which adds robustness to local minima. Moreover, the deformation of the model is constrained and guided by global shape statistics. The model is deformed to the brain boundary by a procedure that matches the local image structures and evaluates the similarity in the whole patch rather than on a single vertex. The experimental results show high agreement between automatic and supervised skull-stripping results.

1. INTRODUCTION

Skull stripping of brain images is usually a necessary and non-trivial procedure for volume measurement, visualization and shape analysis. Currently, most of the available skull-stripping methods are manual, or semi-automatic. Some automatic skull-stripping and related topics have been studied and reported in [1, 2].

We have previously used a semi-automatic procedure [3], which is based on a sequence of morphological operations, thresholding, seeding, region growing and manual editing. In this paper, we use the results from the first automatic part of the above-mentioned method as our initialization, and then deform a surface model [4] to the boundary of the brain in the MR image. There are two characteristics that make the algorithm robust and accurate.

1) A hierarchical model representation and deformation strategy is used. We have determined elasticity variation along the surface model, which is based on the statistical variation of deformation field between initial and final configurations of the surface model (c.f. Fig 1). This is implemented by a hierarchical representation of the model surface using a set of overlapping patches with different elasticities (c.f. Fig 2). In each iteration, instead of deforming each vertex on patch separately, an energy function is evaluated on a whole patch. Each patch deforms as a whole to a position that maximizes its overall fit to the data. After each iteration, each patch is split into two overlapping patches with similar areas until an area threshold is reached (c.f. Fig 3). This surface representation and deformation strategy increases the chances of avoiding local minimum and also makes the deformation adaptive to elasticity variation along the surface model.

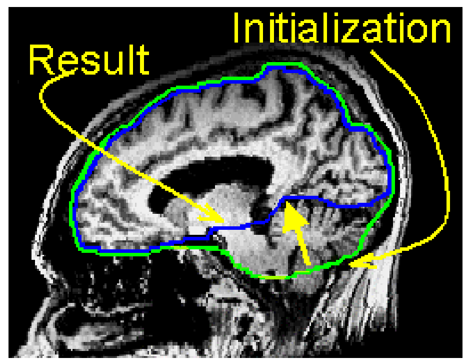

Fig. 1. A typical example showing that some parts of the model have high elasticity than other parts.

2) An attribute vector that characterizes the local image structure is attached to each model point. This is similar to the idea of 'profile' used in Active Shape Model [5]. The attributes in our attribute vector are calculated from both image intensities and edges in an ellipsoidal neighborhood around each model point. The size of the ellipsoid is initially large and decreases gradually. Statistics of each attribute vector, which are determined from a set of aligned brain samples, are also attached to each model point. In the initial deformation stages, the algorithm focuses on attributes that represent relatively coarser aspects of the underlying anatomy, and progressively uses finer attributes based on the edge map.

2. METHODS

2.1. General Formulation

The goal of our automatic skull stripping algorithm is to deform a surface model to the brain boundary in MR images. The deformation of the surface model will be constrained and also guided by prior local shape information that is calculated from a certain surface patch[4], and by global shape statistics that is learned from a set of training samples. Each patch will be deformed as a whole to a pose, based on a similarity criterion that is integrated on the whole patch, and is not evaluated on a single model point.

Suppose a model surface S comprises m surface patches ($S = \{P_i | 1 \le i \le m\}$) in the initial stage, with n_i vertices on the i-th patch P_i, i.e. $P_i = \{v_{i,j} | 1 \le i \le m, 1 \le j \le n_i\}$, where $v_{i,j}$ is the j-th point in the i-th patch. The energy function that our deformable model minimizes is defined as follows:

$$E = \sum_{i=1}^{m} w_i E_i^{patch} \tag{1}$$

$$E_i^{patch} = \sum_{j=1}^{n_i} E_{i,j} = \sum_{j=1}^{n_i} (E_{i,j}^{model} + E_{i,j}^{data}) \tag{2}$$

The weighting parameter w_i determines the relative elasticity given to the i-th patch P_i, with the total energy E_i^{patch}. In the initial stages, large weights are assigned to the patches with high elasticity, in order to make them deform close to their corresponding positions in the image first; these are the patches for which the training set indicates that a large deformation is necessary to reach the brain boundary. A typical example of a patch with large deformation range is the cerebellar boundary. This is because our initialization, which includes region growing and morphological operations, ends up including the cerebellum, which is tightly packed next to the cerebral hemispheres. Our final skull-stripped images do not include the cerebellum. During later stages of the deformation, the same weighting is assigned to all patches in the model, since at that time all the patches are close to their corresponding positions in the image. The local energy term $E_{i,j}$, which is defined for the point $v_{i,j}$, is composed of two terms, $E_{i,j}^{model}$ and $E_{i,j}^{data}$. The model term $E_{i,j}^{model}$ is composed of two constraints [4]: 1) a constraint on local shape similarity between the model surface and the currently deformed configuration; 2) the global shape constraint by prior knowledge of shapes.

The data energy term $E_{i,j}^{data}$ statistically measures the similarity of local image structures around the position $v_{i,j}$ in the model and in the studied image. The local image structure is represented by the attributes that are calculated from both image intensities and edges in an ellipsoidal neighborhood around each model point. The detailed description of $E_{i,j}^{data}$ is given in section 2.3.

2.2. A hierarchical model representation and deformation strategy

From our training set, we determined the elasticity variation along the surface model; this variation is proportional to the statistical variation of deformation field between the initial and final configurations of the surface model. In order to perform an adaptive deformation, different parts on the surface model are deformed according to their elasticity. This is implemented in a hierarchical way. Initially, the surface model is divided into 9 patches according to their elasticity and the topological structure of the surface model. These patches are shown in Fig 2 by different colors. By dividing the model surface into a number of patches, each patch deforms independently according to its elasticity. These patches partly overlap in order to maintain deformation continuity on the surface.

Each patch is deformed as a whole to the positions where the total energy in the i-th patch, E_i^{patch}, is optimal according to a local greedy search. The range of search for each patch is determined by its elastic characteristics. This patch deformation strategy is more robust than the strategy that deforms each model point at each time. However, the proposed deformation strategy cannot accurately follow the details of the brain boundary, since the evaluation of the similarity is performed on the whole patch, and the sizes of the initial patches are too large. To overcome this, we progressively split each patch into 2 overlapping sub-patches with similar areas in the beginning of each iteration. This splitting operation continues until the area of the patch is less than a certain threshold. Fig 3 demonstrates an example of splitting patches. The first figure in Fig 3 is the initial patch that is corresponding to the part of cerebellum in the model, then with the increase of the iterations each patch was split into two sub-patches, according to the arrows in Fig 3.

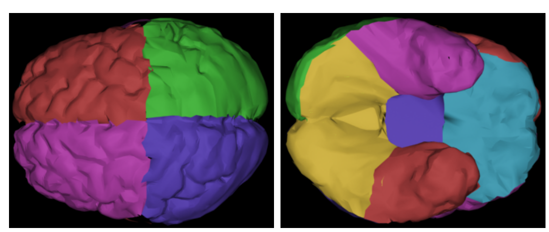

Fig. 2. Initial patches in the model surface.

2.3. Attribute vector

An attribute vector that characterizes local image structure is defined for each model point. Attributes in the attribute vector capture the spatial distribution of image intensities and edges inside of an ellipsoidal neighborhood that is centered on a model point $v_{i,j}$, with its long axis normal to the

Fig. 3. A demostration in splitting patches.

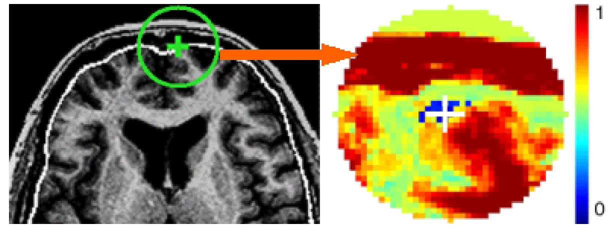

Fig. 4. Demonstration of the attribute vectors in discriminating local image structure. The color coded similarity between the attribute vectors of a point and its neighbors is given in the right, with blue indicating highest similarity. Only the immediate neighborhood around the point indicated by a cross has similiar attribute vector. This implies that this point can be distinguished from others, based on its attribute vector.

model surface. The statistics of each attribute vector can be obtained from a set of the aligned samples, i.e. obtaining the mean $\bar{s}_{i,j}$, eigenvectors $e_{i,j}$ and eigenvalues $\lambda_{i,j}$. We used 100 training brains [6], with supervised skull stripping results. When using the surface model to segment the brain boundary, the attribute vector $s_{i,j}$ from the local image structure around each model point $v_{i,j}$ in an image is statistically compared with the attribute vector that has been learned from the training samples. The similarity between them is used as the data energy term, i.e. $E^{data}_{i,j} = (s_{i,j} - \bar{s}_{i,j})^T e_{i,j} \lambda_{i,j} e^T_{i,j} (s_{i,j} - \bar{s}_{i,j})$.

Two adaptive strategies are used in calculating the attributes and in evaluating the similarity between attribute vectors. First, the attribute vectors are calculated in several scales of the neighborhood. The size of the neighborhood is initially large, and it evolves to be very small as the algorithm proceeds. Second, we focus on different parts of the attributes at different deformation stages, with initially focusing on the attributes calculated from the intensity map and finally focusing on the attributes calculated from the edge map.

The use of the attribute vectors enables us to capture local geometric information, thereby distinguishing one voxel from others in the image. It is demonstrated in Fig 4, where the attribute vector of the point indicated by a cross is compared to the attribute vectors of other points in its neighborhood. The color-coded difference map is shown in the right figure of Fig 4, where the blue color is used for the smallest difference, i.e. the highest similarity. From this figure, we can see the capability of the local image attribute vector to distinguish local structures.

3. EXPERIMENT RESULTS

10 randomly selected subjects have been used to test the performance of our algorithm in automatic skull stripping. Fig 5 shows two typical results, with the white contours representing the locations of the finally deformed surface model in the cross-sectional MR images. Fig 6 gives an initial skull stripping, and the final skull stripping result.

We also measured the accuracy of our algorithm in skull stripping, by comparing our results with supervised skull-stripping results. The differences between two skull stripping results are defined as the maximum distance and the average distance between the points in the deformed surface model and their nearest boundary voxels in the supervised result. The image size is $240mm \times 240mm \times 186mm$. For the 10 testing brains, the maximum distance is ranging from 0.94mm to 2.81mm, with mean 1.34mm and standard deviation 0.74mm. The average distance is ranging from 0.0005mm to 0.074mm, with mean 0.058mm and standard deviation 0.026mm. Compared to the actual size of the image, this error is relatively small and cannot be observed, since most errors are localized in the CSF around the cortex.

For the current version of our skull stripping algorithm, without any code optimization it takes about 20 minutes on an SGI CPU (195MHZ). The algorithm always finishes the skull stripping in less than 10 iterations.

4. CONCLUSION AND FUTURE WORK

In this paper, we have presented a statistical shape model for automatic skull stripping of the brain in MR images. The proposed method used two strategies. First, it uses a hierarchical representation of the model surface, using partly overlapping patches. The size and the number of the patches evolve with time. At each iteration, each patch deforms after an energy function integrated over the whole patch is optimized. The reason of splitting the surface into different overlapping patches is to make our deformation algorithm adaptive to elasticity variation along the surface model. Sec-

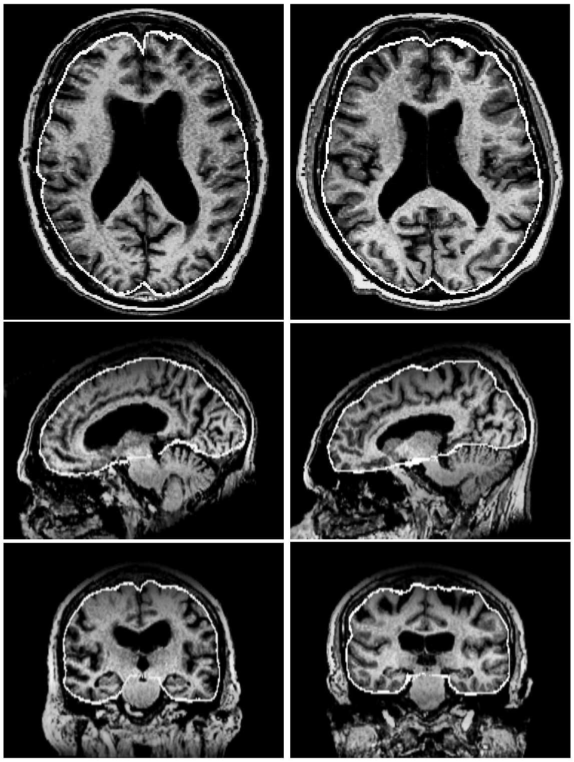

Fig. 5. Two typical skull stripping results; white contours are our results.

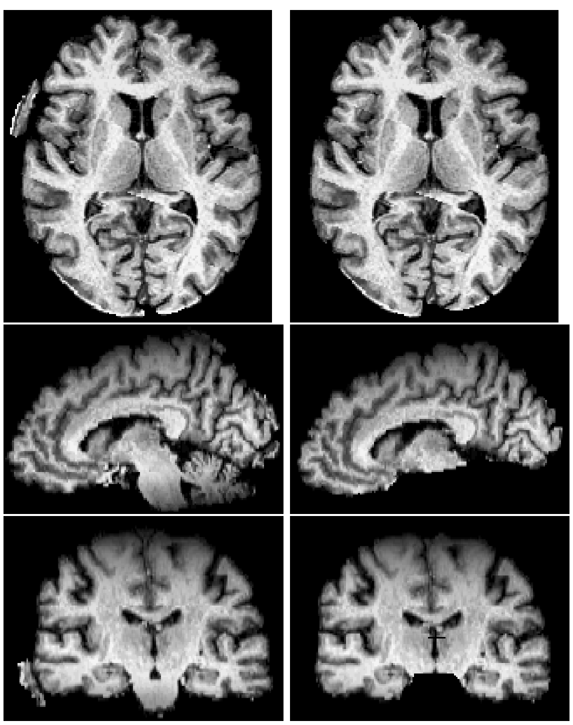

Fig. 6. Comparison on initial and final skull stripping results.

ond, the attribute vector that hierarchically characterizes the local image structure from the image intensity map and the edge map, are employed to deform the surface model to the brain boundary, by focusing on different attributes at different deformation stages. In the final deformation stage, the attributes that are corresponding to the edges are used, thereby increasing the ability of our algorithm to follow the finer details of the brain boundary. Furthermore, both local shape information and global shape statistics are used to constrain and guide the deformations of the surface model.

Our future work includes the development of a fast initialization method for our algorithm. We will design a method to extract the outline of the head first, and then estimate the initialization of our surface model by using the statistical relationships between the outlines of the head and the brain.

Acknowledgments: We would like to thank Dr. Susan Resnick and the BLSA for providing the datasets. This work was supported in part by NIH grant R01 AG14971 and by NIH contract AG-93-07.

5. REFERENCES

[1] J.C.Bezdek, L.O.Hall, and L.P.Clarke, "Review of mr image segmentation techniques using pattern recognition," *Med. Physics*, vol. 20, pp. 1033–1048, 1993.

[2] Y.Hata, S.Kobashi, and S.Hirano, "Medical image segmentation by fuzzy logic techniques," in *IEEE International Conference on Systems, Man and Cybernetics*, 1998, vol. 4, pp. 4098–4103.

[3] A.Goldszal et al, "An image processing system for qualitative and quantitative volumetric analysis of brain images," *J. Comp. Assist. Tomogr.*, vol. 22, no. 5, pp. 827–837, 1998.

[4] Dinggang Shen, Edward H. Herskovits, and C. Davatzikos, "An adaptive-focus statistical shape model for segmentation and shape modeling of 3d brain structures," *IEEE Trans. on Medical Imaging*, vol. 20, no. 4, pp. 257–270, April 2001.

[5] T.F.Cootes, G.J. Edwards, and C.J.Taylor, "Active appearance models," *IEEE PAMI*, vol. 23, no. 6, pp. 681–685, 2001.

[6] S. M. Resnick, A. F. Goldszal, C. Davatzikos, S. Golski, M. A. Kraut, E. J. Metter, R. N. Bryan, and A. B. Zonderman, "One-year age changes in mri brain volumes in older adults," *Cerebral Cortex*, vol. 10, pp. 464–472, 2000.