3D Brain Segmentation Using Active Appearance Models and Local Regressors

K.O. Babalola, T.F. Cootes, C.J. Twining, V. Petrovic, and C.J. Taylor

Division of Imaging Science and Biomedical Engineering,
The University of Manchester, Manchester, UK
kola.babalola@manchester.ac.uk
http://www.isbe.man.ac.uk/~kob

Abstract. We describe an efficient and accurate method for segmenting sets of subcortical structures in 3D MR images of the brain. We first find the approximate position of all the structures using a global Active Appearance Model (AAM). We then refine the shape and position of each structure using a set of individual AAMs trained for each. Finally we produce a detailed segmentation by computing the probability that each voxel belongs to the structure, using regression functions trained for each individual voxel. The models are trained using a large set of labelled images, using a novel variant of 'groupwise' registration to obtain the necessary image correspondences. We evaluate the method on a large dataset, and demonstrate that it achieves results comparable with some of the best published.

1 Introduction

Accurately segmenting structures from 3D medical images is an important step in a wide range of applications, particularly when diagnosing disease, monitoring patients or evaluating the effects of pharmaceutical compounds.

Many approaches to automatic and semi-automatic segmentation are being developed. In this paper we describe a system which uses volumetric Active Appearance Models (AAMs) [4] to locate the shape and positions of each structure, then refines the segmentation with a set of local regressors which estimate the probability that voxels near the boundary belong to the structure.

One of the most effective methods of segmentation of subcortical structures in the brain is that of Aljabar *et al.*[11]. Given a set of labelled training images, a new image can be segmented by registering each of a subset of the training images to the new image, transfering the labels and using these to vote for the most likely label for each voxel. By selecting the subset to be those training images with similar properties to the target image, state of the art performance can be achieved.

This approach is effective, but relatively time-consuming, due to the number of full 3D registrations required. The method described below is inspired by this classifier fusion approach, but is much more efficient. Rather than perform many individual registrations, we use a combination of AAMs to locate the structures

D. Metaxas et al. (Eds.): MICCAI 2008, Part I, LNCS 5241, pp. 401–408, 2008.

rapidly. Rather than retaining all the training set to use to perform label voting, we summarise the information by training local regressors, one per voxel in a model reference image. These are able to correct small boundary errors, leading to a good final result.

The contributions of the paper include

- An improved method of finding correspondences by applying groupwise registration to multi-plane images containing both intensity and label data
- A two layer system for locating deep brain structures using a global AAM followed by the use of individual AAMs for each part.
- A method of refining the solution using simple linear regressors trained in the reference frame
- An evaluation of the performance on multiple structures on a large dataset

In the following we describe the key components of the system in more detail, and present quantitative results on a large dataset.

2 Background

There are many approaches to segmenting structures from 3D volumetric images, including deformable surface meshes, medial representations, segmentation by registration, and shape constrained level set approaches, amongst others e.g. see [12]. However, here we concentrate on methods using statistical shape models, based on the original work of Cootes *et al.*[5]. In particular, the Active Appearance Model, a fast method of matching appearance models to new images, has been found to be effective for analysing medical images [3]. It was extended to volumetric images (2D+time) by Mitchell *et al.*[10]. To construct such models each image in a training set must be annotated with corresponding points. Frangi *et al.*[8] obtained correspondences by using volumetric deformation of binary images. Klemencic *et al.*[9] extended this to build 3D (volumetric) AAMs of the hippocampus. In related work, Duchesne and Collins [7] used an AAM trained on deformation fields to estimate the deformation of a new image. The target image was then segmented by transferring labels from the reference volume.

3 Methodology

To train the system we have access to a set of 3D MR images, for each of which we have voxel label images (i.e. an image of label values indicating which structure each voxel belongs to).

In order to construct statistical models of shape and appearance, it is necessary to find a set of points in each image which define the correspondences.

3.1 Computing Correspondence

The label images allow us to construct binary images for each structure. To find correspondences for a given structure we could use the approach of Frangi

et al.[8] and apply non-rigid registration to such binary images. However, we use a variant of the "groupwise" registration approach of Cootes *et al.*[2]. This seeks to find the deformations of space which allow us to construct the most compact model of a set of images. A key novelty in this work is that rather than registering a set of binary label images, we generate a set of two plane images in which the first plane is the binary label, and the second is a locally normalised intensity image. Groupwise registration involves estimating the volumetric deformations which best match each two-plane image to the group mean. The inclusion of both intensity and labels in the process means that the correspondences are estimated to both match the label boundaries, as well as nearby structures in the grey-level image. This extra information leads to better final shape and appearance models. Yeo *et al.*[13] also register intensity and label images to build probabistic atlases.

The deformations are represented by the movements of a tetrahedral mesh, covering the region of interest. The nodes of the mesh are control points. The positions of the control points on two images define a sparse set of corresponding points across the images, and a dense correspondence can be found efficiently using affine interpolation within each tetrahedra of the mesh.

3.2 Model Construction

By aligning the sets of control points on each image and applying Principal Component Analysis we can generate a statistical shape model [5].

By warping each grey-level image into the mean reference image (using the tetrahedral mesh) we can apply PCA to the resulting textures to generate a statistical texture model. We can combine the shape and texture models to build a combined statistical model of appearance [4], with the form

$$\begin{aligned} \mathbf{x} &= \bar{\mathbf{x}} + \mathbf{Q}_s \mathbf{c} \\ \mathbf{g} &= \bar{\mathbf{g}} + \mathbf{Q}_g \mathbf{c} \end{aligned} \tag{1}$$

where $\bar{\mathbf{x}}$, $\bar{\mathbf{g}}$ are vectors of the mean shape and mean texture, $\mathbf{x}$, $\mathbf{g}$ are the shape and texture vectors in the reference frame, $\mathbf{Q}_s$, $\mathbf{Q}_g$ are matrices describing the modes of variation derived from the training set, and $\mathbf{c}$ is a vector of parameters controlling both shape and texture. For a given $\mathbf{c}$ we can generate a texture, then warp this using the generated shape to create a synthetic image of the modelled object. For instance, Figure 1 shows a slice through the 3D image generated by the model of the region around the left ventricle.

Such a model can be matched to a new image rapidly using the Active Appearance Model algorithm [4]. This seeks to minimise a sum-of-squares problem of the form

$$F(\mathbf{p}) = |\mathbf{r}(\mathbf{p})|^2 = \mathbf{r}^T \mathbf{r} \tag{2}$$

where $\mathbf{p}$ contains the t model parameters (the appearance parameters $\mathbf{c}$, together with global pose parameters - see [4]), and $\mathbf{r} = \mathbf{r}(\mathbf{p})$ is a function returning the n_g residual differences between model and data for parameters $\mathbf{p}$. By making assumptions about the Jacobian, a fast updating algorithm can be derived which can match the model to a new image in a few iterations.

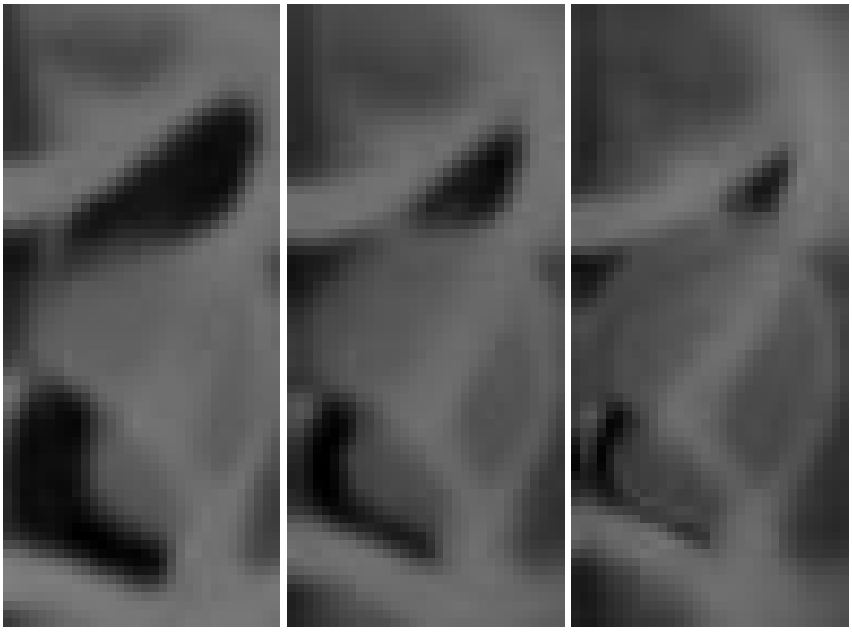

Fig. 1. Slice through images synthesized by varying the first mode of the appearance model of the left ventricle between -2.5 and +2.5 standard deviations

3.3 Segmenting Voxels

Given the correspondences estimated by the groupwise registration, we can warp each binary label image into the reference frame. If we compute the mean of these warped images, across the training set, we derive a mean probability image. Each voxel of this gives the probability that it belongs to the object of interest in the reference frame. For instance, Figure 2a shows a slice through the mean probability image for the left ventricle.

One approach to segmenting a new image is to first match the appearance model using the AAM, then to project the mean probability image into the target image using the resulting model points. The resulting warped probability image can be thresholded at 0.5 to obtain a hard estimate of voxels inside and outside of the object of interest.

However, in some cases the result of the AAM search may not be accurate enough to give a good segmentation. Mis-matches can occur because of poor initialisation, unmodelled image structures or because the model does not contain enough degrees of freedom to accurately deform to a particular image.

Such sub-optimal matching may mean that the object boundaries are incorrectly delineated. However, where such mis-matches are relatively small, we can correct for them by analysing the local image structure.

Rather than use the mean probability image, we generate a new probability image based on pixel intensities warped into the model frame.

a) Mean Probability Image b) Difference in probability due to regression

Fig. 2. Probability images in reference frame for the left ventricle. The image on the left is the mean probability image. That on the right is a difference image of the probabilities computed from the mean for one subject before and after application of regression.

During a training phase we warp each normalised intensity image into the model reference frame using the known correspondences. For any given voxel in this frame we then have a set of probabilities, p_i, $i = 1..n_{images}$, that it is inside the object (by warping the binary label images)[1] together with corresponding vectors of intensity values $\mathbf{g}_i$ sampled in the region around the voxel.

We then perform linear regression to learn a function to estimate the probability given the intensity pattern

$$p = f(\mathbf{g}) = \mathbf{a}^T\mathbf{g} + d \tag{3}$$

This is repeated for every voxel near the boundary. Voxels away from the boundary are assumed to have either p=0 (outside) or p=1 (inside).

Given a new image, we can then segment it using the following steps

- Match a volumetric AAM to estimate correspondence with the model frame
- Use the resulting correspondences to warp the normalised intensities into the model frame
- Use the voxel probability estimators (Eq.3) to compute the probability image in the reference frame
- Use the correspondences to warp this into the image frame
- Obtain a labelled image by selecting the structure with maximum probability ≥ 0.5 at each voxel

Figure 2b shows the difference in the resulting probability image before and after regression.

[1] We use trilinear interpolation during warping, leading to values of p in the range $(0, 1)$ for voxels near boundaries.

4 Experiments

We were provided with 270 T1 MR images and corresponding manually labelled images which included the labels of 18 subcortical structures – the brain stem, fourth ventricle and the left and right pairs of the accumbens, amygdala, caudate, hippocampus, lateral ventricles, thalamus, pallidum and putamen (see figure 3). There was a wide variation in the subject pool age (4.5 years – 83 years), and disease (controls, Alzheimer's disease, Schizophrenia, Attention Deficit Hyperactivity Disorder and prenatal drug exposure).

We established correspondence using the groupwise registration method described in section 3.1. We then constructed both a global AAM (containing all structures) and a set of individual AAMs, one for each structure. The voxel regressors were constructed as described above.

To evaluate the performance of the method a set of 27 leave-10-out experiments were performed, in which the models are trained on 260 images, then tested on the remaining 10. The search is fully automatic, involving three stages:

i) a global search with the AAM of all structures,
ii) local search with each individual structure AAM to refine the matches
iii) segmentation by combining the generated probability images into a single label image

The quality of the resulting segmentations were quantitatively evaluated using the corresponding manually labelled image as a gold standard and the Dice overlap coefficient [6] as the evaluation metric.

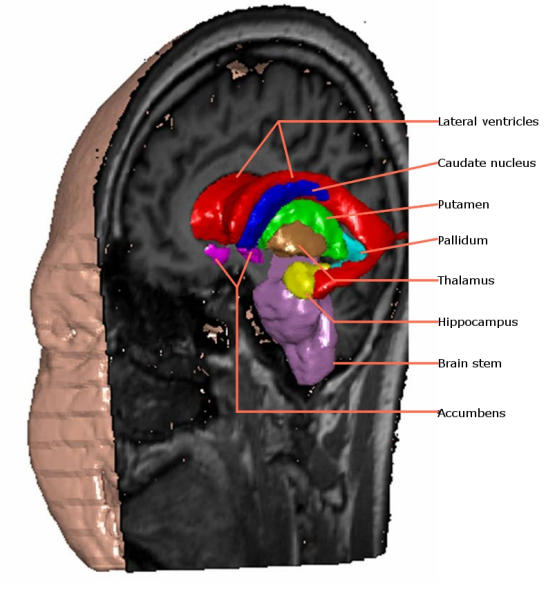

Fig. 3. Surface rendering of an MR image showing subcortical structures used in model building

5 Results

Columns 2 and 3 of Table 1 shows the Dice coefficients for the results of the leave 10 out segmentations using the AAM, both with and without the final regressor estimate of the probability image. It demonstrates the the regressors are able to significantly improve the Dice overlap results for almost every structure.

Table 1. Mean Dice overlap coefficients for segmentation results. Standard errors are shown in parenthesis ($n = 270$ for fourth ventricle and brain stem, $n = 540$ for other structures). The last column shows results for the method of Aljabar *et al.*[11] on the same dataset (figures obtained from [1]).

Structure	AAM + mean prob	AAM + regressor	Classifier Fusion
Accumbens	73.1% (0.4)	73.7% (0.4)	75.8%
Amygdala	75.0% (0.4)	76.1% (0.4)	77.7%
Brain Stem	91.3% (0.4)	92.5% (0.4)	94.2%
Caudate	85.9% (0.2)	88.7% (0.2)	88.1%
Fourth ventricle	78.7% (0.4)	82.9% (0.3)	83.2%
Hippocampus	80.9% (0.2)	83.7% (0.2)	83.5%
Pallidum	81.7% (0.2)	83.1% (0.2)	81.9%
Putamen	88.6% (0.1)	90.8% (0.1)	89.8%
Thalamus	89.9% (0.1)	90.8% (0.1)	90.8%
Lateral ventricles	83.2% (0.4)	90.2% (0.3)	91.3%

6 Discussion and Conclusions

We have described a system for segmenting sub-cortical structures from 3D MR images of the brain. The method first finds the approximate positions of all structures using a global AAM. It refines the positions using individual AAMs of each structure, then produces a final segmentation by computing a labelled image from probability images of each structure. These probability images can be derived from the mean of the training set. However, we have demonstrated that better results can be obtained by using simple regressors at each voxel to estimate the probability of object occupancy based on the pattern of grey level intensities nearby. Because this is able to correct for some mis-matching of the shape models, it generally leads to better overall results.

The results for most structures are comparable to the best published results for these structures, though most other results are on much smaller datasets. Direct comparison with four different algorithms from four different groups [1] shows that the method is achieving results similar to the best known method on this dataset (the fusion of predictions from registering many atlases [11]).

The proposed method is relatively swift, taking less than 20 minutes to segment all structures in a single image using a standard single processor PC.

The tool and models will be made freely available for research purposes.

We anticipate it will be possible to improve the results still further by tuning the model building and search parameters for each structure individually, and

by exploring more sophisticated regression models for predicting the probability images.

Acknowledgements

This work was funded by the EPSRC. David Kennedy of the Center for Morphometric Analysis, Boston, provided the MR images used.

References

1. Babalola, K., Patenaude, B., Aljabar, P., Schnabel, J., Kennedy, D., Crum, W., Smith, S., Cootes, T.F., Jenkinson, M., Rueckert, D.: Comparison and evaluation of segmentation techniques for subcortical structures in brain MRI. In: Proc. MICCAI. LNCS, vol. 5241 (2008)
2. Cootes, T., Twining, C., Petrović, V., Schestowitz, R., Taylor, C.: Groupwise construction of appearance models using piece-wise affine deformations. In: 16th British Machine Vision Conference, vol. 2, pp. 879–888 (2005)
3. Cootes, T.F., Beeston, C., Edwards, G.J., Taylor, C.J.: A unified framework for atlas matching using Active Appearance Models. In: Proc. Information Processing in Medical Imaging, pp. 322–333 (1999)
4. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active Appearance Models. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(6), 681–685 (2001)
5. Cootes, T.F., Taylor, C.J., Cooper, D., Graham, J.: Active Shape Models - their training and application. Computer Vision and Image Understanding 61(1), 38–59 (1995)
6. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology 26, 297–302 (1945)
7. Duchesne, S., Pruessner, J.C., Collins, D.L.: Appearance-based modelling and segmentation of the hippocampus from MR images. In: Proceedings of the 23rd Annual International Conference of the IEEE on Engineering in Medicine and Biology Society, vol. 3, pp. 2677–2680 (2001)
8. Frangi, A.F., Rueckert, D., Schnabel, J.A., Niessen, W.J.: Automatic 3D ASM construction via atlas-based landmarking and volumetric elastic registration. In: Insana, M.F., Leahy, R.M. (eds.) IPMI 2001. LNCS, vol. 2082, pp. 78–91. Springer, Heidelberg (2001)
9. Klemencic, J., Pluim, J., Viergever, M., Schnack, H., Valencic, V.: Non-rigid registration based Active Appearance Models for 3D medical image segmentation. Journal of Imaging Science and Technology 48(2), 166–171 (2004)
10. Mitchell, S., Boudewijn, P., Lelievedt, P.F., van der Geest, R., Bosch, H., Reiber, J., Sonka, M.: Time continuous segmentation of cardiac MR image sequences using active appearance motion models. In: SPIE Medical Imaging (February 2001)
11. Aljabar, P., Heckemann, R., Hammers, A., Hajnal, J., Rueckert, D.: Classifier selection strategies for label fusion using large atlas databases. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 523–531. Springer, Heidelberg (2007)
12. Pizer, S.M., Fletcher, P., Joshi, S., et al.: Deformable M-Reps for 3D medical image segmentation. International Journal of Computer Vision 2-3(55), 85–106 (2003)
13. Yeo, B., Sabuncu, M., Desikan, R., Fischl, B., Golland, P.: Effects of registration regularization and atlas sharpness on segmentation accuracy. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 683–691. Springer, Heidelberg (2007)